Loudspeaker Listening
Tabula Rasa or Augmented Reality

DAVID PRIOR

ABSTRACT

Recent developments in virtual and augmented reality technology have stimulated renewed interest in the role of sound and music in these domains. In this article loudspeakers, and the spaces used to listen to them, will be discussed with reference both to the dominant media that have influenced their evolution and in light of emerging media—particularly in augmented reality—which value very different audio-spatial relationships.

In their paper on augmented reality displays, Paul Milgram et al. put forward a useful taxonomic framework for discussing a "Reality-Virtuality" continuum of mixed reality display technologies. On their scale, the real environment lies at one end and virtual reality (VR)—defined as an environment in which "the participant-observer is totally immersed in a completely synthetic world"—lies at the other [1].

Unlike screens, which confine their content to their own perimeter and surface, the behavior of loudspeakers has always been one of *augmenting* the acoustic reality of an existing space. However, while loudspeakers are necessarily located in places that have physical—and therefore acoustic—characteristics of their own, a variety of technical, cultural and economic forces have conspired to make the *ideal* of the loudspeaker listening space one that eradicates its own acoustic signature in favor of being better able to convey the character of the space being rendered—whether live or recorded—by the loudspeaker [2]. Contrary to their inherent "augmented" affordances, then, to this extent loudspeakers have aspired to the ideals of VR.

LOUDSPEAKER ORIGINS

In the pre-electrical era of sound recording, the iconic horn used in Thomas Edison's phonograph and Emile Berliner's gramophone adopted an established form found in brass instruments and "speaking trumpets" to provide a simple means by which to amplify the weak acoustic vibrations cre-

David Prior (artist, academic), Academy of Music and Theatre Arts, Falmouth University, Cornwall, TR10 9EZ, U.K. Email: <david.prior@falmouth.ac.uk>.

See <www.mitpressjournals.org/toc/lmj/-/26> for supplemental files associated with this issue.

ated by the device's stylus. This directional focusing of sound energy was an effective method for maximizing limited sound energy but it also initiated an audio-spatial paradigm that has proved durable long after this means of amplification was an acoustical necessity. With electrical recording in the 1920s came vacuum-tube-based amplification and the introduction of the "dynamic speaker" still used today—a technology that is not inherently directional in the way the acoustic horn was. Although an evolutionary development in loudspeaker design can be observed in the period since, it is perhaps testament to the unchanging ideals of loudspeaker listening that the fundamental principles behind the design of both the speaker driver itself and the cabinet that encloses it have changed relatively little. Indeed, the quest of most loudspeaker manufacturers has been to create devices ever more independent of the rooms in which they are situated, with companies such as Meyer Sound [3] and Kii Audio [4] now boasting true cardioid radiation characteristics down to previously impossibly low frequencies and the widespread adoption of digital signal processing–based technologies designed to "correct" the acoustical shortcomings of a listening space [5].

THE CULTURAL CONTEXT OF THE LOUDSPEAKER

Efficient amplification was not the only factor influencing the physical form and acoustic behavior of the early dynamic loudspeaker, however. The archetypal device we have inherited is a manifestation of the promise of the speaker as neutral conduit for the musical expression of an absent other: a music reified as media, abstracted from the social practice of shared performance experience and objectified for commercial consumption. As Jacques Attali [6] and Christopher Small [7] remind us, this process of the professionalization of the musician, concurrent with a detachment from her immediate social context, began in the Middle Ages and continued along a trajectory that is manifest today not only in the media objects of the sound recording but also in a range of cultural conventions, from the etiquette of concert listening and the construction of a celebrity culture that values

musicians somewhat proportionally to their apparent remove from everyday life. Jonathan Sterne relates this theme more specifically to recording technology, demonstrating that the apparatus of sound recording would not even have been conceived were it not for the cultural forces that made these developments desirable [8]. After all, the loudspeaker has always been a device that amplifies not only the *sound* an artist makes but also their cultural status as "other" to the environment in which they are heard. In this regard, the common practice in hi-fi listening of identifying a "sweet spot" can be read not just as optimizing the acoustics of a point of audition but also as emblematic of the desire to transport the listener to another place entirely: a place as removed as possible from the acoustic reality of the listening environment.

BLUMLEIN SPACE

In Edison's 1887 list of possible uses for the phonograph, music was nowhere near the top (and not at all in the way that we now understand it) [9]. However, as the burgeoning music industry became a *recording* industry, the phonograph's relationship with reproduction technology deepened, and with the widespread adoption of stereo in the 1960s, a new formation of audio-spatial relationships emerged. While both the phonograph/gramophone horn and the monaural dynamic speaker valorized a recorded space at the expense of the listening space, stereo listening added a psychoacoustic feat to this cultural imperative, suppressing not only the acoustic characteristics of the listening space but also the objecthood of the loudspeaker itself.

As he was the pioneer of both binaural stereo microphone technique and the theoretical work upon which the "panning" of monaural sounds between two speakers is based, stereo listening owes a great debt to the work of Alan Blumlein. The space created by Blumlein's techniques resides *between* the loudspeakers, causing our aural perception of the speakers themselves to recede, ideally to the point of disappearance, much as the room had done before them. Unlike the intimacy of monaural radio listening, where sound gravitates toward the device that produces it, stereo sound is constituted in space and can only be apprehended *as* stereo once it has left the speaker and entered the space in which it is heard. However, rather than this being a means by which to engage with the acoustic properties of the listening environment, stereo listening requires an even more exacting suppression of room character to function at its optimum.

The cinema was the other significant driver in the development of both loudspeaker technology and an idiomatic mode of loudspeaker listening [10], but while the cinema relied on techniques and technologies similar to those employed in domestic music reproduction, the listener's relationship to space has not been directly equivalent. Although sound was synchronized with film as early as 1900, it did not achieve mainstream adoption until the 1930s, and while experimental spatial audio was explored almost immediately, widespread use of it did not occur until much later. When it did, by contrast with domestic stereo systems that relied on the liminal, Blumleinian space between two loudspeakers, cinema sound emerged from monaural systems into a variety of stereo formats, many of which retained a center channel for dialogue. Stephen Handzo explains that originally this was to benefit from both the diffuse radiation characteristics associated with paper driver cones preferred for music and the more directional characteristics associated with metal drivers preferred for conveying dialogue [11]. However, assigning specific roles to speakers also had the effect of negating Blumlein's articulation of space by means of a prosthetic phantom center and so reinstating the objecthood of the center speaker by gravitating the sound it emitted back to its point of origin.

AUGMENTED AUDIO: THE EXAMPLE OF HEADPHONES

Along the continuum of mixed reality displays described by Milgram et al., VR displays—often synonymous with the use of head-mounted displays (HMDs) such as Oculus Rift—lie at one of the end of the scale, defined by the total immersion of a user in a virtual world. In their ability to isolate a user from all external aural stimuli, headphones can achieve a similar sense of immersion within a single sensory register. Headphones, however, are symptomatic of the limitations of attempting neat parallels between AR and VR as defined in the visual domain because they can perform in both capacities. Indeed, despite the marketing campaigns of headphone manufacturers themselves (Bose Corporation, for example, describes their headphones as a "sanctuary or a concert hall" [12]), headphones have in fact gone further than any loudspeaker technology to point the way toward how mediated sound might interact with the lived experience of an acoustic environment in the active pursuance of augmented audio.

Although open-back headphones and "ear buds" allow the ingress of sounds from the outside environment, both open- and closed-back designs eschew the acoustic signature of the acoustic environment altogether because the audio heard through them does not stimulate that environment. Despite the inability of headphones to interact with environmental acoustics, their portability and the fact that they do not interfere with other sensory registers makes them eminently well suited to locative listening experiences. Michael Bull has written extensively about the ontology of headphone listening, focusing principally on the experience of listening to music not composed for a specific location but heard outside of a domestic environment [13]. In his writing, he gives a lucid account of the relationships listeners form between site and sound, despite the inability of personal stereo technology to respond dynamically to the environment in which it is located. Many sound artists, however, have gone on to exploit the affordances of headphones more specifically. In an interview with Kelly Gordon, Janet Cardiff speaks of her fascination with binaural listening and its potential to elicit new relationships between sound, time and space:

> I was fascinated by the layering of the past onto the present. It had a strange quality of creating a new world, blending together the physical and the virtual. I was also very excited by how my recorded body walking and talking created such an intense physical presence for me, as if there were another woman that was part of me but separate [14].

The journey of sound from a loudspeaker transducer to the ear of a listener necessarily incorporates a series of interactions with the surfaces of the space in which it is heard, and the spatial image created by two or more loudspeakers must be constituted in this space. By contrast, headphones constitute the perceived spatial characteristics of the audio they are rendering within the head of the listener and for this reason the fabric of the listening space can have no bearing on the perceived space of the audio. To subscribe to the putative notion of headphones as a category of audio VR, however, would be to miss the unique characteristics that Cardiff identifies in the quote above, whereby in site-specific binaural listening, spatial integrity is preserved while temporal integrity is disrupted. Although the uncanny phenomenon she describes relies on the recording location being identical to the playback location and on the listener using headphones of a design that allow the natural sounds of the environment to commingle with the recorded audio, Cardiff's work nevertheless represents a paradigm shift in the audio-spatial ontology of mediated listening. Since Cardiff's artistic exploration of headphones, various sound artists have integrated location-sensitive technologies that facilitate the dynamic control of content in response to location, with the artist group Circumstance [15] and Udo Noll's phonography streaming app Aporee [16] being notable examples. Cardiff's own innovation, however, lay not in pioneering a new technology but in imagining a different formulation for the use of an existing one.

AUGMENTED REALITY: A TECHNOLOGICAL TURN

As I have established, the form that early loudspeakers took derived from a combination of both technological limitations and the cultural expectations associated first with monaural radio and later with stereo music reproduction. Surround sound in its various forms has expanded and reified some of the latent assumptions behind hi-fi listening by placing an idealized Cartesian listener at the center of an array of directional speakers from which sounds approach the listener within a space ideally devoid of its own acoustic signature [17].

If artists such as Cardiff have explored alternative formulations for these technologies, we might also look at the technology sector itself for examples of ways in which longstanding ontologically assumed restrictions associated with the loudspeaker might now be destabilized. The resurgence of interest in ambisonic microphone technologies, and in particular the emergence of higher order microphone arrays, for example, affords a much more fluid relationship with recordings and their playback context, challenging the notion of recording media as fixed inscriptions of an acoustic event. Rather than capturing a spatially defined "window" of sound from a fixed perspective, an ambisonic recording captures a full sphere of sound from which multiple perspectives—whether mono, stereo or surround—can be derived later, in mixing, or even by an end user. Similarly, the emergence of object-oriented surround mixing, such as Dolby Atmos, moves away from the notion of mixing into a given number of discrete channels in favor of describing a sound's spatial occupancy and movement by means of vector coordinates, meaning that a spatial audio mix could reconstitute itself according to the context in which it is heard. In this way the same mix could be decoded differently whether heard on headphones, on speakers, in a car or on a multichannel cinema system [18]. With game audio having long since challenged the teleological expectations of fixed media recordings, and now near-ubiquitous mobile technologies providing a range of environmental data to inform details of our listening context, a confluence of technologies herald the potential for recordings to be responsive to the environments in which they are heard.

CONCLUSION

Current developments in visual display technology look set to disrupt both our working definition of the screen and our understanding of its ontology [19]. As we come to make use of these new ways of seeing, our received notions of what it is to engage with visual media and, in turn, how these media can respond to the places in which they are used, will be fundamentally challenged. By comparison, the loudspeaker has for the most part remained stoically unresponsive to the cultural and technological changes surrounding it. While challenges to its orthodox form have occurred at the margins of art practice, the underlying principles of both its design and the listening habits associated with it have remained remarkably durable.

If the current surge of interest in the cultural significance of technologically mediated realities is being driven by innovations in display rather than sound technology, these developments nevertheless provide a welcome provocation by which to reconsider the loudspeaker and the culture of listening that surrounds it. Loudspeaker listening can be understood to be a form in which the naturally immersive affordances of sound—affordances that now constitute the goal of much augmented reality technology—have tended to be suppressed in favor of a mode of listening much more closely aligned with the ideals of what is now being described as virtual reality. The successes of that formation of the loudspeaker can be evidenced in the role it has played in hi-fi, concert and cinema listening, but it is with eager anticipation that we may look forward to how the expectations we have of the loudspeaker might evolve to meet the challenges of augmented reality.

References and Notes

1 Paul Milgram et al., "Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum," in *SPIE* Vol. 2351, *Telemanipulator and Telepresence Technologies* (1994) p. 283.

2 Paul Thebérge and Everett Devine, eds., *Living Stereo: Histories and Cultures of Multi-Channel Sound* (New York and London: Bloomsbury Academic, 2015).

3 See <www.meyersound.com/news/1998/psw-6_test>, accessed 17 December 2015.

4 See <www.kiiaudio.com/en/prod-pro.html>, accessed 17 December 2015.

5 A good example of this can be found in Genelec's Smart Active Monitor (SAM) range of speakers: <www.genelec.com/studio-monitors/sam-studio-monitors>, accessed 17 December 2015.

6 Jacques Attali, *Noise: The Political Economy of Music* (Minnesota: Minnesota Univ. Press, 1999).

7 Christopher Small, *Musicking: The Meanings of Performing and Listening* (Middletown, CT: Wesleyan Univ. Press, 1998).

8 Jonathan Sterne, *The Audible Past: Cultural Origins of Sound Reproduction* (Durham and New York: Duke Univ. Press, 2003).

9 In the 10 uses Edison identifies for his phonograph in 1878, he refers to music in item 6, writing, "Music boxes, toys etc.—A doll which may speak, sing, cry or laugh may be promised our children for the Christmas holidays ensuing," from <www.edisonmuckers.org/edisons-top-ten-uses-for-the-phonograph>, accessed 17 December 2015.

10 Emily Thompson, *The Soundscape of Modernity: Architectural Acoustics and the Culture of Listening in America, 1900–1933* (Cambridge and London: MIT Press, 2004) pp. 235–248.

11 Stephen Handzo, "Appendix: A Narrative Glossary of Film Sound Technology," in E. Weis and J. Belton, eds., *Film Sound: Theory and Practice* (New York: Columbia Univ. Press, 1985).

12 David Prior, "The Rise, Fall and Fallout of Concert Halls as a Primary Space for Listening," in Colin Ripley, Marco Polo and Arthur Wigglesworth, eds., *In the Place of Sound: Architecture | Music | Acoustics* (Newcastle: Cambridge Scholars Press, 2007).

13 Michael Bull, *Sounding Out the City: Personal Stereos and the Management of Everyday Life* (Oxford and New York: Berg, 2000); Michael Bull, *Sound Moves: iPod Culture and Urban Experience* (London: Routledge, 2007).

14 Janet Cardiff, interviewed by Kelly Gordon, <www.hirshhorn.si.edu/collection/directions/#detail=/bio/directions-janet-cardiff-words-drawn-in-water/&collection=directions>, accessed 21 December 2015.

15 See <www.wearecircumstance.com>, accessed 17 December 2015.

16 See <www.aporee.org/mfm>, accessed 17 December 2015.

17 The pseudo-generic man is used here advisedly; see Keir Keightley, "'Turn It Down!' She Shrieked: Gender, Domestic Space, and High Fidelity, 1948–59" *Popular Music* 15, No. 2, 149–177 (1996), for a discussion on hi-fi and gender.

18 For details on a current research program related to object-oriented audio, see <www.s3a-spatialaudio.org/wordpress>, accessed 17 December 2015.

19 Oculus Rift, Sony PlayStation VR, HTC/Valve Vine and Sulon Q as well as a number of smartphone-based HMDs have all been released in 2016.

DAVID PRIOR *is an artist and an academic. He is partner in liminal (<www.liminal.org.uk>) and associate professor in Music and Sound Art at Falmouth University, U.K.*

─ ANNOUNCEMENT ─

MIT Press Journals Podcasts

The MIT Press Journals Podcast series allows *Leonardo* authors and contributors to continue the conversation initiated in the pages of the journal in dialogue with each other. Tune in to hear from authors, editors and artists working at the intersection of art, science and technology. The MIT Press Journals Podcast webpage currently hosts a discussion between Jean-Marc Chomaz, Mikael Fernström, Sean Taylor, Annick Bureaud and Roger Malina on their work in the arts and sciences showcased in the article collection, *Water Is in the Air: Physics, Politics, and Poetics of Water in the Arts*; a conversation inspired by Leonardo e-Book Series release *Art and Atoms*, featuring Tami Spector, Philip Ball, Kathryn de Ridder-Vignone, Julian Voss-Andreae and Roger Malina on the connections between science—especially chemistry—and art; and more lively, engaging discussions of interest to the art/science community. To learn more: <www.mitpressjournals.org/page/podcasts>.