Managing Complexity in Multi-Device Environments

Oliver Bown, UNSW
Frederic Robinson, UNSW
Kurt Mikolajczyk, UTS
Sam Ferguson, UTS

Pre-print copy.
Published in the Routledge Handbook of Sound Design, Michael Filimowitz (ed).

Introduction

Whilst spatial audio concepts have been with us for decades, most spatial audio work, historically, has been concerned with the design and fit-out of fixed listening environments such as cinemas, performance venues, living rooms, music studios and cars. Increasingly, sound design is situated in more dynamic and unpredictable contexts, where it is integrated into a wider range of experiences and pursues more adventurous design goals. A common example is in gallery and museum spaces where sound may be attached to multiple exhibits which audience members navigate interactively, and which may be more or less coordinated. A myriad of factors coalesce in such design scenarios, drawing on knowledge of audio cognition, acoustics, and interaction and experience design. This is an area of continued refinement of practices, standards, and cultures of adoption, interspersed with radical innovation, influenced by an interconnected matrix of transformations in technology and culture (outlined in Table 1).

Technology capabilities	Cultures of adoption	Infrastructure
• Network technologies and related technological solutions for synchronising and coordinating audio streams over wired and wireless networks. • Technological solutions for positioning speakers • Technologies for acoustically modelling spaces. • Technologies for compensating for acoustic factors. • Technologies of online media distribution. • Advances in speech recognition and generation, powered by AI and embedded in devices including robotics.	• The development and adoption of spatial audio formats and standards. • The adoption and standardisation of the Internet of Things (IoT). • The growth of interaction and user experience design and its inclusion of sound. • The increased use of spatial audio in live music and art installation contexts. • The increased use of rendered audio in public spaces, galleries, offices and so on. • Concepts of transmedia. • The improved integration of spatial and multi-speaker possibilities into computer and audio hardware and software. • Interactive and generative audio and music authoring and deployment tools.	• The widespread adoption of smartphones and other personal devices. • The increased use of video and audio-conferencing technology. • Smart speakers and voice assistants. • The increased use of headgear including noise-cancelling headphones and VR/AR devices. • Accumulation of infrastructures that further drive adoption (example: bands performing with their own monitoring setup, having knock-on effects for how venues are set up and the job of live sound engineers).

Table 1: Ongoing developments in technology capabilities, cultures of adoption and infrastructure. Here, cultures of adoption include design developments such as the creation of authoring tools (these are not primarily about technological breakthroughs as about serving user needs through design solutions).

A central technological transformation underway is the steady growth of the Internet of Things (IoT). IoT systems offer new technological solutions for sound designers, but also transform the context in which sound designers work, populating the world with objects that can both make sound by themselves and be coordinated to make sound collectively.

For example, the first author's house contains 11 devices capable of rendering digital audio streams over the internet, another 7 household devices that make digital bleeps to report status (such as microwaves), and a series of other internet connected devices such as IoT plug sockets. None of the networked devices were *made* to be synchronised in multi-device media delivery, but they are entirely capable of it, creating new possibilities. Consequently, home electronics companies such as LG (Kim, Kim, and Kim 2019), and media organisations like the BBC (Francombe et al. 2017), are taking seriously the future potential affordances of media orchestration.

The goal of this chapter is to capture a snapshot of the changing landscape of sonic experience design in an IoT world, and then to present an analysis of this scene in terms of the challenges that both product and sound designers face in creating effective and high-quality sonic experiences. Its central theme is design complexity: how designers make sense of these emerging complex design spaces. The snapshot of the landscape proceeds by presenting emerging application areas and related frontiers in technology development. A subsequent analysis proceeds by introducing the theme of complexity, and then looks at creative design strategies that frame conceptualising sound, finding a working process, and handling hybrid design challenges.

Part I: Survey

Emerging Application Areas

There are a vast number of contexts in which audio is rendered through loudspeakers. People increasingly travel around with noise cancelling headphones listening to audio in highly controlled sonic contexts. Many other carefully designed listening environments exist – music venues, cinemas, theatres, meeting rooms, and lounge rooms – where the acoustic space is controlled to prioritise the rendering of high-definition audio in a shared experience. At the opposite extreme, our world is full of electronic doorbells, alarm clocks and microwave bleeps; simple, point source, functional sounds designed to cut through complex sonic environments.

Recent years have seen these contexts blur, driving techniques for the design of richer sonic experiences in more diverse and complex contexts. New digital technologies continue to drive a persistent embedding of media into our environments. Advertising billboards are replaced with digital screens, buildings draped in bespoke lighting design, and more things bleep or speak. Meanwhile, it has become increasingly feasible for these lights, sounds, and other distributed media to be coordinated together across device multiplicities.

Thus, the answer to the question of where we can find emerging applications of sound design is, increasingly, "everywhere". In the wake of the rising prominence of "ubiquitous computing", correspondent interests now gather in the form of communities of practice in ubiquitous music (ubimus) (Keller, Lazzarini, and Pimenta 2014), the internet of musical things (IoMT) (Turchet, Fischione, et al. 2018; Turchet, Viola, et al. 2018), the internet of sounds (IoS) (Turchet et al. 2023) and sonic interaction design (SID) (Franinovic and Serafin 2013).

This expanded sound design field includes three main areas of development, which we discuss in turn.

Multi-User Experience Environments

"Multi-user" contexts are complex spaces like museums, galleries and showrooms where the careful coordination of sound design had not typically been a priority. There has been increased interest in sound's role in the advancement of media design in such environments. These sound environments are often made up of a multitude of individual sound sources accompanying individual media exhibits. However, as visitors experience these environments as a whole and determine the timeline on which they progress through them, efforts have been made to orchestrate these experiences more holistically in a process referred to as sound scenography. In practice, this means coordinating different streams of audio content like ambient music, UX sound at interactive media stations, audio guide narration, and film soundtracks accompanying surface projections.

Figure 1: The throne room of Schloss Ludwigslust.

Such a scenario can be found in sound scenography work by the second author, working with Swiss audio design studio *Idee und Klang.* To mark the end of restorations of Schloss Ludwigslust in Germany in 2016, *Idee und Klang* were commissioned to create a site-specific audio drama in its historical throne room (Figure 1). A global surround soundscape depicted a baroque orchestra rehearsal which could be heard across the room. Local speakers on benches along the walls reproduced conversations between various historical characters. The overarching sound experience was therefore comprised of local sources and a shared spatial sound environment, and visitors had to move throughout the exhibition space to hear the story in its entirety.

While sound in this context fulfils many of the functions found in other domains like film or video game audio, the spatial – and, by extension, multi-device – nature of the work introduces new challenges and opportunities. Spatial sound is deployed to make environments immersive and believable. As multiple visitors share the same experience, loudspeakers are often preferred over headphones. As audiences are spread out and moving between different sonic experiences, however, designers cannot rely on a unitary listening experience found in traditional surround sound systems. Soundscapes may also compete with adjacent audio sources such as projector fans, or audience chatter, for which broad-band noise can be used for masking. Sounds may also be designed to not interfere with higher-priority sound sources, such as a tour guide. In the context of large-scale surface projections, spatial cues can be used to draw a visitor's attention to specific areas in a visual scene.

Multi-Device / Multi-Function Personal Contexts

A second application area is in home and office contexts where single or multiple users move around a space and engage with sonic media for different purposes including device interaction, communication and entertainment, and increasingly involving voice-based assistants. As part of ongoing research at the BBC's R&D division, Francombe et al. (Francombe et al. 2017) introduced Media Device Orchestration (MDO) as a strategy to optimize media reproduction using commonly available devices, such as laptops, TVs and smartphones. The approach prioritizes the experience of envelopment over precise localization, offering a practical alternative to dedicated surround systems that require more involved setup and calibration. A radio drama involving several narrators, for example, may be distributed across devices in the home to place the listener in the centre of the story. While the adoption of such technology is still at a formative stage, Francome and colleagues' evaluation showed MDO was well received among listeners, demonstrating that media content can be successfully augmented through distributed devices, even if those devices are in unorthodox and unpredictable locations.

Object Design

A third application area is in digital-physical products that increasingly use sound in their interaction design. For example, in the automotive industry, sound is found simultaneously in safety design, interaction design and entertainment. The sound design of mobile phones, tablets and computers is becoming increasingly sophisticated, taking into account the complexities of prioritisation and management of multiple apps making sound.

Sonic Interaction Design (SID) (Franinovic and Serafin 2013) looks at infusing objects with sound design to improve interaction, made possible by low-cost, low-power and small-format digital audio components. SID's focus is on functional augmentation, exploiting sound's affordances as an information channel, but is also concerned with how sonic aesthetics can transform user experience.

On the immediate horizon lies the sonic design of social robots, which due to their close relation to smart devices, both regarding technical capabilities and deployment context, have generally received similar sonic treatment as mobile phones, computers, and smart devices. An example is Amazon's domestic robot Astro, a robotic evolution from its smart display Amazon Echo Show.

Going from "smart" to "social" entails new challenges for sound design. Beyond voice and alert sounds, research into sound design for social robotics extends to nonverbal utterances that convey emotion and intent without semantic content. Their physical presence also introduces motor sound, which has been shown to convey characteristics unintentionally (Moore et al. 2017). Robot sound designers have begun to harness this channel using movement sonification (Robinson, Velonaki, and Bown 2021). In this context, designing sound for robots becomes a more holistic endeavour that involves the creation of a comprehensive sonic identity, rather than of a collection of individual responses and alerts. Commercial social robot sound designers have therefore drawn inspiration from domains like film and video games, where sound identities can be created from the ground up without having to take real-world deployment into account (Robinson, Bown, and Velonaki 2022).

Properties of Contemporary Sonic Media Experiences

For much traditional sound design, the work of creating fixed media content in a purpose-built audio studio completes the process, but for these new emerging contexts, sound design and interaction design fuse to invoke new forms of complexity. The complexifying factors can be captured as follows:

- **Multi-user and/or multi-device and/or multi-function**: Multiplicities introduce complexity into the design process, especially when they overlap into a "multiplicity of multiplicities", such as when multiple users are in a space and sound is being used to achieve multiple functions.
- **Flexible and dynamic**: We increasingly need sound design that is adaptive to different speaker configurations including speakers that move.
- **Situated in more complex acoustic environments**: In these environments, listeners are mobile. The design of sound competes with other noises and other design constraints, and may involve coordinated distributed elements.
- **Interactive and generative**: Where sonic outcomes are not predictable and are constrained by the specifics of available dynamic audio techniques.

These complexifying factors define a design space, which is, above all, shaped by the affordances of the technologies available, to which we now turn.

The Technology Landscape

Whilst audio recording and playback media has gone through consistent evolution over the past 100+ years – now digital, networked and potentially generative – how we render sound through speakers had remained relatively unchanged until the 21st century. It is now going through rapid change in three main areas.

Coordinating Sound-Making: Hardware

Loudspeakers take a great diversity of forms and functions. Indeed, although digital audio delivered via DACs to electronic loudspeakers have almost universal applicability, to be complete we should be mindful of such diverse sound-making devices as resonance transducers, piezo speakers, and mechanically controlled acoustic devices. Loudspeaker design variants include diffusion, multi-directional, bone conducting and parabolic speakers. Spatial audio techniques such as beamforming (Michel 2006) and wavefield synthesis (Berkhout, de Vries, and Vogel 1993) further transform how loudspeakers can deliver sound.

Of particular importance to the affordability of spatial and coordinated distributed configurations is how they are connected. Low latency ethernet-based multi-speaker technologies are replacing traditional analogue electronic audio cables, reducing wiring through daisy-chained configurations. For example, the Dante protocol[1] supports sending hundreds of audio channels over ethernet, which can be decoded at nodes in a wide variety of configurations. Wireless networked systems are also common, but at a cost of significant bandwidth limitations, lower network stability, and significant baseline latencies.

The majority of distributed audio system designs are based on the idea of a central digital device streaming multi-channel audio to speakers via one of the above technologies. However, one can also distribute the process of audio generation, for example using networks of low-cost computers. In that case, a central controller may send networked control messages instead of audio streams. This massively reduces bandwidth usage over the network, as well as reducing the burden of computation performed on the central device (significant when scaling to 1000s of channels of audio). In turn this introduces other design challenges such as synchronisation and pushing the audio computation costs onto the remote devices.

An early experimental performance which predated smartphones was Dialtones, by Golan Levin, Scott Gibbons, Greg Shakar and Yasmin Sohrawardy (Levin, Gibbons, and Shokar 2005). The work simply used audience members' mobile phone ring tones, activated by calling them. Amidst such primitive constraints, the work revealed the creative potential of a mass of networked sounding devices to form a creative medium. In a contemporary take, Garth Paine demonstrates how easy it is to now remotely orchestrate a realtime electronic music performance on audience members' smartphones. Web audio frameworks make it simple to remotely invoke and control complex synthesis algorithms directly through a web browser session (Bevilacqua et al. 2021). In both cases, portable sound emitters are co-opted for synchronised performance, much as MDO aims to do in more functional settings. In our own work (authors 1,3 and 4), we have built arrays of up to 1,000 networked synthesisers, powered by Arduino and Raspberry Pi computers, receiving commands over WiFi. We have also built portable and interactive sonic devices that can be remote-programmed to perform synchronised and interactive music (Bown, Young, and Johnson 2013; Bown et al. 2015).

Spatial Formats

---

[1] https://www.audinate.com/meet-dante

Multi-channel audio systems have long been organised by industry standards, largely driven by cinema. The 5.1 standard emerged along with Dolby's technique of optical matrix encoding for adding sound to film. The film Apocalypse Now was the first to use a Dolby 5.1 soundtrack, and this soon became a standard for cinema theatres. The 5.1 standard comprises a stereo pair, a centre, and two surround speakers, plus a subwoofer. These systems can be expanded through additional speakers, such as overhead speakers that add a vertical dimension (5.1.2) or 7.1 where two additional speakers are placed to the left and right of the listener.

A key limitation of these systems is that surround mixes for the variety of loudspeaker distributions (from stereo upwards) need to be individually hand-crafted, and consumers need to set up their systems accurately to experience sound as intended by the artist. While this is feasible in professional studios and cinemas, musicians and consumers are faced with a high barrier to entry, as significant hardware investments are required. Even then, audiences in cinemas sit in different positions relative to the speakers and so do not receive the same experience.

The concept of object-based audio naturally arises as a means to abstract away from specific speaker configurations. Here, the artist does not map sound to a particular loudspeaker, but instead assigns positional metadata to the individual elements of a spatial mix. The assignment of audio content to individual speakers is done at the point of rendering. Endpoints might then include stereo headphones with binaural rendering, soundbars or surround systems.

A key benefit of this approach is that artists can work in the abstract. In theory one can create and monitor spatial mixes solely using a pair of headphones, and listeners can listen to the result on any system capable of rendering spatial audio, again including headphones.

There is another affordance. Because the mix takes place at the point of rendering, it is possible to adapt mixes for different contexts. For example, the voice tracks in a TV show can be lifted above background music for hard of hearing listeners. There are many other creative applications of this dynamic affordance, such as the playful spatial distribution of elements in a musical work. A downside of this potential for variability of configuration and mix is increased uncertainty on the part of the creator about how the end result will sound.

The dominant object-based spatial audio format is Dolby Atmos, now firmly established in film sound and seeing growing adoption in the music industry as well. A Dolby Atmos master can contain up to 128 channels of audio, which can be individual objects with their own positional data but can also be part of beds that represent canonical speaker locations. A sound file could contain a 7.1 bed with three additional objects, or it could be exclusively made up of audio objects, allowing practitioners to blend object-based audio workflows with traditional multi-channel mixing approaches and tools. In the bandwidth-restricted context of streaming, it is unfeasible to deliver 128 channels of audio to end devices. Atmos objects are therefore eventually rendered down to 16 positions around the listener in a process called spatial coding (Breebaart et al. 2019; Andersen et al. 2004).

Thus, whilst object-based audio formats are a natural evolution towards abstraction, with significant affordances, they also present new design challenges: handling potentially very complicated authoring projects; considering file size and bandwidth limitations; and being able to anticipate and control how things will sound under a diversity of conditions. Another introduced challenge is the reduced ease of interoperability between systems (as compared to simple stereo files).

These issues influence adoption, as does the highly interdependent complex of tool development that helps make an established ecosystem of practice. For example, there are currently few audio effects plugins that make full use of the object-based audio paradigm. While a growing number of musicians are eager to explore the medium, the tools available to them are mostly made for stereo contexts. The mastering process is affected as well, as there is little support for processing clusters of objects without first rendering them.

Such shortcomings can easily fade away as the ecosystem of support shifts around spatial. For example, the recently developed SkyDust 3D synthesizer by developer Sound Particles, integrates spatial distribution as part of the synthesis process. With these complexities in mind, the design of creative workflows for object-based audio continues to develop (e.g., (Coleman et al. 2018)).

Media Device Orchestration

As mentioned above, Media device orchestration (MDO)[2] is the far more speculative, experimental idea of creating rich spatial audio-visual experiences across devices potentially as heterogeneous as fridges and robots. Such devices may be connected over low-bandwidth, high-latency, and unstable wireless connections. This challenge aside, the various application areas and technological considerations above point to a potential culmination of rich spatial audio design in diverse spaces using available infrastructure. There are many reasons this may not eventuate – not least, the low cost of electronics means that we can have dedicated spatial audio systems as well as latent, networked devices capable of sound creation – but it is nevertheless an area of heightened interest as the technology landscape emerges.

An important design differentiator here lies in the distinction between sound-making objects *as objects*, and sound-making objects as elements that serve to render a sonic field; between *collections of objects* and *substrates* (Bown and Ferguson 2018). In home media consumption, we generally engage with screens as objects, but with the pixels that make up those screens as invisible renderers of a substrate. From a user experience perspective, configurations of objects and substrates underlie the potential ways that users might engage with diverse sound-making objects. In MDO, there is a greater tendency to think of the various devices as objects rather than as elements of a substrate.

For example, modern device ecologies (Lyle, Korsgaard, and Bødker 2020) (Bødker and Klokmose 2012) have led to the natural emergence of "second screen" (Cesar, Bulterman,

[2] MDO was coined by researchers at the BBC (citation: https://www.bbc.co.uk/rd/publications/media-device-orchestration-immersive-spatial-audio-reproduction) in a project that looked specifically at using the set of existing in-home speakers for 3D audio rendering.

and Jansen 2008) interaction, such as using a smart phone to look up background relevant to a TV show, or to engage in live audience participation via social media. In such emergent interaction scenarios, the “objectness” of devices is primary to any role they may play rendering a distributed media experience.

Interrelated to this, the poor bandwidth and network stability of wireless devices limits real-time streaming and synchronisation potential, but many devices can make up for this with computational power, meaning that they can perform synthesis. These design scenarios also involve complex spatial and interactive interrelations between people and devices. Thus, whilst MDO can be conceived of as a natural extension of spatial audio to utilise ever more diverse and fluid speaker infrastructure, device orchestration may define a significantly different design space from the traditional spatial audio paradigm. We can, for example, think of the speakers in a room as individual voices in an orchestra.

Most critical to MDO’s place in the technology landscape, though, is how it could yet evolve. As with the slow evolution of an ecosystem for spatial audio, comprising technologies, designs and cultures of practice, IoT-based audio approaches could in the long term become widely adopted. As computing power continues to fall in price, size and power consumption, and network performance continues to improve, the IoT approach may have great advantages, particularly for massively multiplicitous systems. We can currently build a 1,000 wireless speaker system at a cost comparable to using wired networks.

PART II: Analysis

Complexity as a Design Concern: Constraints and Affordances

In broad terms, then, although we have always made sound that is situated in complex and dynamic environments, technologies are enabling greater control over the efficacy with which we can design for those environments. They are also creating tantalising new contexts in which sound design occurs, in turn bringing new design challenges.

Here technology, experience design, and sound design become tightly integrated concerns, and from a design standpoint, perhaps the most significant issue that arises is how practitioners manage complexity. Work on creativity support tools (CSTs) (Shneiderman 2007) provides useful theoretical tools from which we draw on and develop two connected ideas about how we manage constraints and affordances. The first is the idea that creative work frequently involves creatively devising one’s own set of design constraints in order to tame the vastness of unstructured possibility (Stokes and Fisher 2005; Stokes 2009; Biskjaer and Halskov 2014). The second is the idea that we set out to configure our creative environments with certain affordances, specifically to support rapid exploratory search and reflexive iteration (Resnick and Silverman 2005; Shneiderman 2007; Simonton 2011; Bergström and Blackwell 2016).

Both are important in approaching the design issues presented by modern sound technologies and application areas. For example, a complex authoring environment demands many decisions to be made about how relations between sounds are conceptualised, and decisions made early on in a process may constrain possibilities later on in the process (Blackwell and Green 2003; Biskjaer and Halskov 2014). Likewise, networked technologies simultaneously impose technological constraints and rich affordances. Selecting a paradigm for working with such technologies is a demanding task which shapes the various affordances and constraints one then has to work with.

In our work on creative practices for multi-device media environments, we have condensed concepts from contributors such as Shniederman, Resnick, Blackwell and Daalsgaard into the simplified idea of "space shaping" (Mikolajczyk et al. forthcoming). Space shaping describes how creative practitioners make a cascading series of decisions about how they will achieve aesthetic goals though an ongoing process of "shaping" the socio-technological setups they work with – setups that combine technologies, people, media elements, physical objects, environments and more.

A requirement of working with complex technologies is that technical decisions constantly need to be made to get creative work done. For example, channel or computational limitations may force a sound designer to mix down elements in a way that denies them flexibility later. Although such decisions can be – and often are – made on the fly, our experience is that successful practitioners think carefully up front about how they configure systems to afford them greater creativity, at the expense of shutting down certain possibilities. It follows that creative practitioners divide time between setup phases, where the bulk of such decisions are made, and creative phases, where the creative work gets done within the newly defined constraints: the shaped space. Certain types of technological change reduce the burden on such upfront decision making. For example, digital audio workstations, long liberated from the constraints of the four-track recording studio, can now handle as many tracks as the typical musicians would ever need. Other technologies impose such decision making. For example, while programming environments offer an incredible open-endedness of possibilities, much of a creative coder's work is spent making decisions about which libraries to use and dealing with software engineering problems.

The work of technologists – i.e., those creating products such as authoring tools or speaker systems – aims to relieve the creative technologist of such setup work. But such creative decision making is not only about configurations of technology. It integrates technologies, aesthetics, conceptual work and the practical and social work of delivering projects.

With this in mind we look at two themes that creative sound designers contend with: how sound design is conceptualised; and how hybrid design challenges are handled.

Conceptualising Sound Design

One of the most important stages in setting up for a complex design task is to develop appropriate conceptualisations of the elements involved. This serves to conceptually simplify a range of possibilities. In this way, frameworks are an effective means to support practitioners in managing complexity and developing design strategies. Here we consider how sound designers conceptualise sound and space.

Conceptualising Sound

Practitioners tasked with designing for complex sound design scenarios may come from a wide range of backgrounds, including linear audio-visual media, music production, and video game sound. As practitioners' previous assumptions and practices may not be directly applicable to this new space, it is helpful to adopt a first principles thinking approach by considering (i) what functions sound can fulfil, (ii) how different types of sound should interact in context.

Literature in the domains of film and game audio provides a range of sound functions. Wingstedt (Wingstedt 2005, 2004) assembled a comprehensive list of design goals for film music, which can readily be extrapolated to sound in general. Prior work by the second author investigated Wingstedt's framework through practitioners in the domain of distributed sound, particularly sound installation (Robinson, Velonaki, and Bown 2023). In this context, interviewees highlighted the role of sound in conveying emotional states (Emotive Class), guiding attention (Guiding Class), and creating associations to external concepts (Informative Class) in particular (Table 2).

Emotive Class	Guiding Class	Informative Class
Describing/revealing feelings	Direction of attention	Clarify ambiguous situations
Stating relationships	Focus on details	Communicate unspoken thoughts
Adding credibility	Masking unwanted sound	Acknowledge or confirm audience interpretation
Mood induction	Hide weak elements	Evocation of cultural settings, social status
Foreboding		Representing objects, actions
		Representing characters, relationships

Table 2 – A subset of functions in sound design, adapted from (Wingstedt 2004).

To consider how sound elements work in context with each other, Murch proposes a conceptual spectrum of sound, shown in Figure 2 (Murch 2005). He places the various components of a film soundscape along a one-dimensional design space ranging from language to music, with sound design placed in between. Notably, he posits that as sounds move further to the right, listener expectations of the spatial distribution of the sound they hear become less rigid allowing designers more flexibility in distributing sound sources. Consider, for example, the spatialisation of on-screen dialog with that of sound design that complements and expands the visual scene. The previously mentioned sound installation practitioners similarly highlighted the ability of distributed sound to create believable immersive environments and additionally note its ability to guide a listener's attention across them.

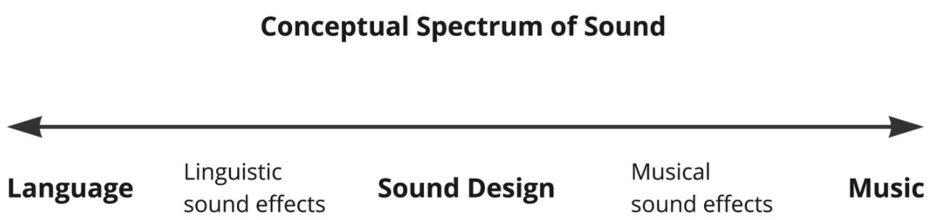

Figure 2 – A conceptual spectrum of sound by Walter Murch.

While these perspectives may be largely tacit for practitioners with established strategies, the challenges presented by new multi-device applications can push such frameworks to the fore. Consider a common contemporary example: A person sits in front of their laptop with their phone placed on their desk. As they receive a FaceTime call, both their laptop and their phone start emitting the same ringtone but out of sync with each other. Both can be used to enter the call. Ringtones fulfil a primary function, to inform the user that there is a call incoming, but also a function that is often implicit, to let the user know where their device is. Based on traditional mobile use cases, a device is considered in isolation and emits alerts regardless of other devices in the environment. Better experience design could result from moving from a device-centric perspective to a user-centric one: it is not devices that receive calls, but users. These users are embedded in multi-device environments which can be orchestrated to most effectively (i) draw attention and (ii) direct the user to an appropriate device. In practice, this could mean muting redundant devices or moving an alert across available devices in a circular manner.

The nature of sound itself has also been conceptualised in ways that have helped classify natural sounds, and consequently understand the interpretation of natural sonic scenes. For example, various formulations have distinguished hitting, resonating and scraping sounds (Gaver 1993) (with equivalents in speech such as plosive and fricative consonants (Molho 1976)) as core building blocks of natural sounds, used, for example, in conceptualising how synthesis techniques can be used to imitate instruments.

Conceptualising Space

The basic affordance of spatial sound technology is that multiple sound objects can be placed anywhere in 3D space. In many situations sounds have a logical location, such as the sound of a voice on screen, but in others there is creative freedom to place and move sounds. One form of simplification is to reduce the dimensionality of 3D space. Indeed, traditional stereo sound is a simplification of spatial sound to a single left-right dimension that proved entirely effective at satisfying the dual goals of placing sounds in a scene and creating spatial effects such as rich reverberation, despite its limitations. Stereo continues to offer an attractive simplicity compared to spatial audio, not least because of the simplicity of the interface required to control it: the one-dimensional pan pot. Whilst 2D controllers are a straightforward evolution from the pan pot, 3D controllers are not.

In many other situations, thinking in terms of abstract 3D space is neither required nor effective. In the Schloss Ludwigslust installation described above, although the many voices playing back over local speakers were designed to be located in specific physical spaces, this was not achieved through 3D spatial audio, but by simply thinking in terms of sound made for each of the individual speakers. This can be considered an “indexical” rather than “speaker agnostic” approach to spatialising sound.

Similarly, in a recent project by the team (authors 1, 3 and 4), a multi-device sound and light installation was being designed in the form of a sphere, which would have flocks of light and sound fly around it. Over time the design evolved under physical constrains from a solid sphere to the surface of a sphere, then to the surface of a hemisphere (Figure 3). At that moment, it became trivial to simplify the conceptual geometry to a two-dimensional, wraparound rectangular surface (i.e., the surface of a cylinder), even though the work was situated in three dimensions. This simplicity made programming the movement of lights and sounds significantly easier, both conceptually and in terms of programming and computation.

Such conceptual simplification reduces work, though it can also reduce the potential for future adaptation.

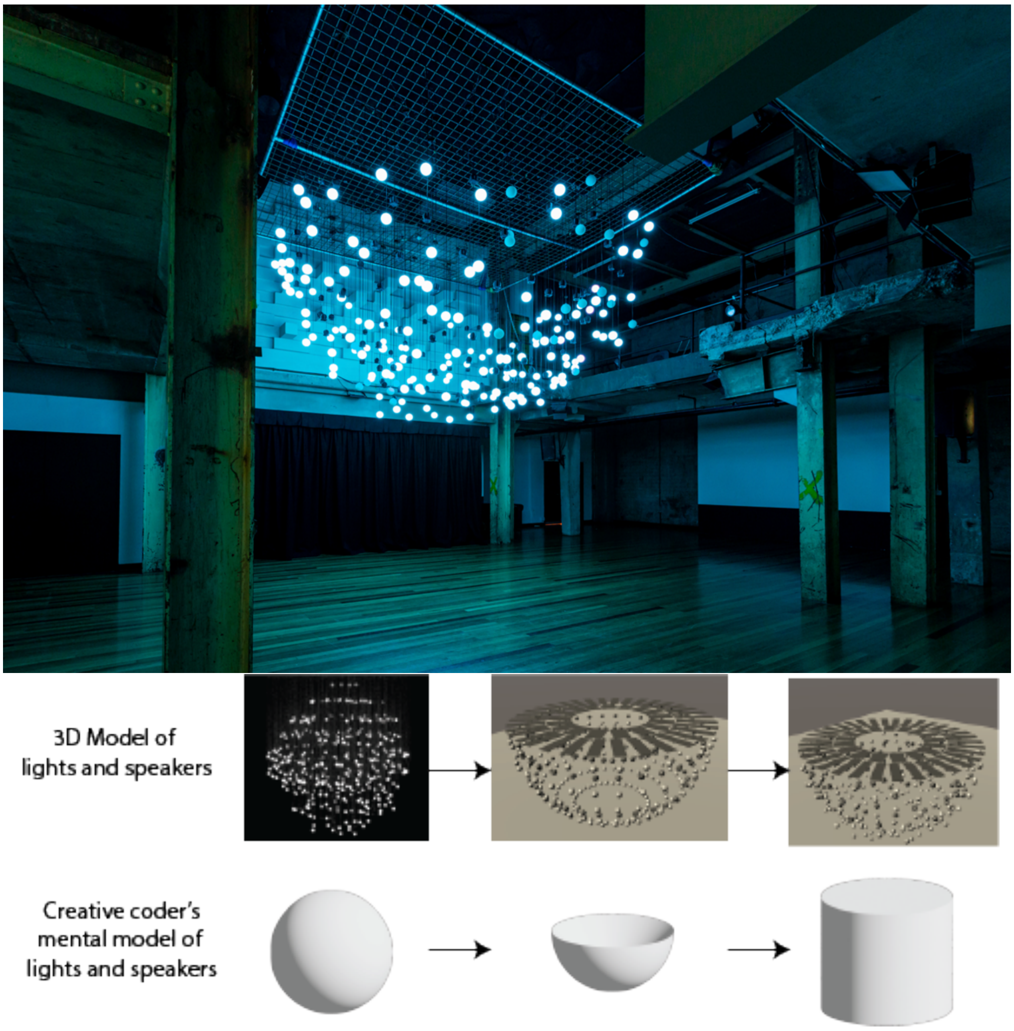

Figure 3: The Mind at Work installation, originally conceived as a sphere. When the spatial design was reconceived as the surface of a hemisphere, the geometry of the sound and lighting design could be reconceived as a simpler rectangular 2D surface.

Handling Hybrid Design Challenges

In realising an output, creators face a series of practical hurdles. These are prominent in more experimental scenarios, as in dynamic music for shopping malls, and more hidden in standardised practices, as in film sound. In our research, such practical hurdles generally fall into three areas: (i) unknowns, (ii) specific technical trade-offs, and (iii) decisions about time commitment to different types of work. Since the latter two issues have been covered elsewhere in this chapter, we focus on unknowns in this section.

There are known knowns, known unknowns and unknown unknowns. A recording engineer knows that their music will be listened to on high-end hi-fis, $5 earbuds, shopping mall sound systems and in cars with growling engines. These are known knowns. They can choose if they want to make a bassline audible on speakers with poor low-frequency response. They can audition tracks on a range of systems easily.

A game music composer works with known unknowns. They know the music may transition from search to battle scenes at arbitrary times, and must ensure the transition is not jarring. They can audition multiple transitions, but depending on the complexity of the situation, may struggle to cover every scenario (Collins 2007). The response becomes heuristic in nature; a gradual fade into sounds with slow attack might be the safest way to cover all instances.

A sound scenographer creating a complex multi-speaker installation can similarly work in a heuristic way to deal with known unknowns. They may not know exactly how things are going to sound in the final space, but they know the general behaviour of a speaker playing a voice track. In this case (unlike the game composer), they will also have the opportunity to fine-tune the sound in the real space: a critical final stage in the process.

In more complex IoT scenarios, unknowns may be more unknown, involving network, processor and software performance issues. In software, profiling performs a sort of "auditioning" of the software's behaviour in a range of scenarios. Wireless networking in IoT scenarios requiring high precision low latency can be invisibly and unpredictably disturbed by unknown unknowns, for example resulting from transmission interference, hardware malfunctions, or software issues.

Heuristics abound in these practices, helping manage knowns and unknowns by acting to conceptualise complexity through abstraction. But the space of constraints and affordances relating to knowns and unknowns can also be reconfigured through the use of both simulation and adaptive techniques. Simulation enables a creator to remove unknowns, to audition in unavailable environments. The sound scenographer can acoustically model spaces and the loudspeakers in those spaces, and the game composer may be able to automate the generation of game state transitions to highlight diverse scenarios. Adaptive techniques automate the solving of problems at the point of delivery, which can introduce new constraints. A common example is the use of limiting and compression in broadcast and stage scenarios: used to avoid unwanted peaks, these effects can also diminish production quality. Smart technologies often serve to support realtime adaptation. In game composition, for example, a virtual musician could be used to make more seamless game state transitions.

As we move to working with more complex, multi-device media configurations, rapidly deployed in physical spaces, we would expect more work to create "digital twin" simulation models of these complex media systems to help remove unknowns and ease creative work, just like Atmos authors can currently simulate different playback systems to monitor their mixes.

Discussion and Conclusion

How do these factors come together in working practices? A sound designer brings training, experience and a changing ecosystem of technologies to their practice, and is also situated in a network of social relations that define how they work. In our model of space shaping, we emphasise the distinction between tasks that can be done quickly enough to allow exploratory search, and those that require more careful planning. The work of designing a complex spatial system is slow, with many unknowns, yet in an ideal scenario one can rapidly experiment with different sonic ideas. Thus in this chapter we have attempted to frame the critical issue of the interplay between slow setup work and rapid experimentation and composition. In this relation, the setup work determines those elements that can be rapidly experimented with. We should think of setup work as encompassing a vast range of time scales, from a sound designer's training (which enables them to be quick at particular practices, say with specific pieces of software), to the immediate work of loading and organising sounds in a DAW for playback (Figure 4). From Heidegger, we can think in terms of which elements become "ready-to-hand" (Heidegger, Macquarrie, and Robinson 1962), that one engages with in an immediate manner.

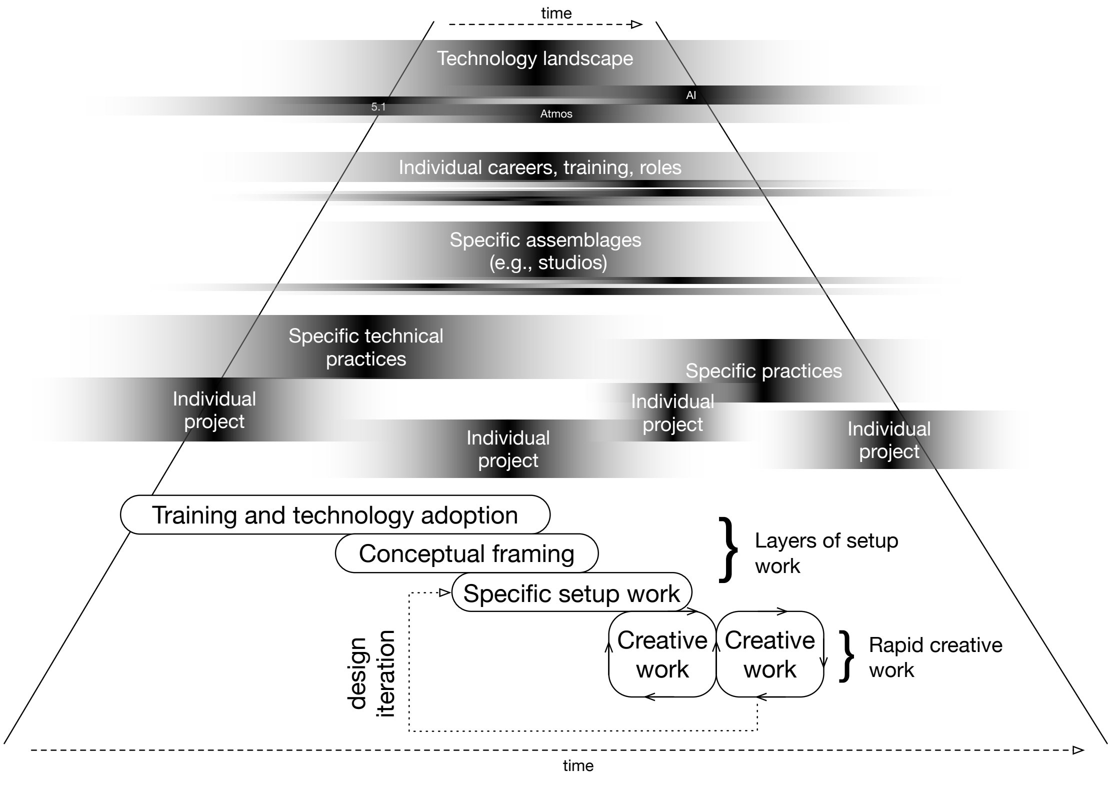

Figure 4: Looking at setup work and creative work as distinct activities within a multi-timescale, ecosystemic view of creative technologies practice. As technologies, individual practices, individual assemblages such as studios and software/hardware configurations come into shape, setup work is done to establish spaces where rapid creative experimentation can be performed. This may involve slower design iterations feeding back into setup work.

Complexity in multi-device environments is what, following Heidegger, can bring elements out of transparent use and "unready-to-hand". Understanding sound design for specific bespoke configurations of devices in their environment requires integrated thinking about sound from multiple perspectives. It means that as a designer conceptualises the knowns and unknowns in a multi-device environment, what should be ready-to-hand can evolve, whether due to challenges or discovering creative affordances.

In a recent study we invited creative musical artists to work with our IoT audio system to create a live performance that involved distributed synthesis (Mikolajczyk, Bown, and Ferguson 2023). The musicians had no prior experience with the system, and their learning of it through a novel interface involved conceptualising ideas for how they would like their music spatially distributed and mapped to light. Those with coding experience also had the potential to write custom distributed synthesis behaviours that went beyond our readymade system. We also offered to implement the musicians' ideas if they could not program.

We attributed three broad strategies to musicians in this scenario: "adopters" learnt how to use the interface we'd made them, thought about how they could use it to augment their existing music, and used exploration to find configurations that worked; "explorers" learnt how to use the system (either via the UI or by programming it) and set out to develop new musical works and concepts that explored the system's creative potential; "stipulators" learnt what the system was capable of, agreed on how it should be used with their music, and requested we develop a specific set of behaviours to realise their idea.

Each of these strategies is witnessed in a very short timescale, compared to the timescale that might be involved in an artist developing a technical practice. They illustrate how an artist shapes a space through a range of choices about their sociotechnical setup and what desired ready-to-hand in the context of multi-device environments We believe that as sound embraces new forms of technical delivery, with new affordances and constraints, sound design is very much in a state of transition. This will no doubt consolidate around standardised technology practices within which sound designers re-establish standardised roles. But even as we move beyond this state of flux, there will be a greater need for sound designers to engage with the working practices of the creative technologist and interaction designer, dealing with networked multi-device systems and understanding user experience in the context of dynamic interactive situations. This chapter offers what we hope is a useful conceptual analysis of the ensuing design space.

References

Andersen, Robert L, Brett G Crockett, Grant A Davidson, Mark F Davis, Louis D Fielder, Stephen C Turner, Mark S Vinton, and Phillip A Williams. 2004. "Introduction to Dolby digital plus, an enhancement to the Dolby digital coding system." Audio Engineering Society Convention 117.

Bergström, Ilias, and Alan F Blackwell. 2016. "The practices of programming." 2016 ieee symposium on visual languages and human-centric computing (vl/hcc).

Berkhout, Augustinus J, Diemer de Vries, and Peter Vogel. 1993. "Acoustic control by wave field synthesis." *The Journal of the Acoustical Society of America* 93 (5): 2764-2778.

Bevilacqua, Frederic, Benjamin Matuszewski, Garth Paine, and Norbert Schnell. 2021. "On designing, composing and performing networked collective interactions." *Organised Sound* 26 (3): 333-339.

Biskjaer, Michael, and Kim Halskov. 2014. "Decisive constraints as a creative resource in interaction design." *Digital Creativity* 25 (1): 27--61. http://www.tandfonline.com/doi/full/10.1080/14626268.2013.855239?src=recsys.

Blackwell, Alan, and Thomas Green. 2003. "Notational systems--the cognitive dimensions of notations framework." *HCI Models, Theories, and Frameworks: Toward an Interdisciplinary Science. Morgan Kaufmann*.

Bødker, Susanne, and Clemens Nylandsted Klokmose. 2012. "Dynamics in artifact ecologies." Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, Copenhagen, Denmark. https://doi.org/10.1145/2399016.2399085.

Bown, Oliver, and Sam Ferguson. 2018. "Understanding media multiplicities." *Entertainment Computing* 25: 62--70.

Bown, Oliver, Lian Loke, Sam Ferguson, and Dagmar Reinhardt. 2015. "Distributed interactive audio devices: Creative strategies and audience responses to novel musical interaction scenarios." Proceedings of the 2015 International Symposium on Electronic Art, Vancouver, Canada.

Bown, Oliver, Miriama Young, and Samuel Johnson. 2013. "A java-based remote live coding system for controlling multiple raspberry pi units." ICMC.

Breebaart, Jeroen, Giulio Cengarle, Lie Lu, Toni Mateos, Heiko Purnhagen, and Nicolas Tsingos. 2019. "Spatial coding of complex object-based program material." *Journal of the Audio Engineering Society* 67 (7/8): 486-497.

Cesar, Pablo, Dick CA Bulterman, and AJ Jansen. 2008. "Usages of the secondary screen in an interactive television environment: Control, enrich, share, and transfer television content." Changing Television Environments: 6th European Conference, EUROITV 2008, Salzburg, Austria, July 3-4, 2008 Proceedings 6.

Coleman, P., A. Franck, J. Francombe, Q. Liu, T. de Campos, R. J. Hughes, D. Menzies, M. F. S. Gálvez, Y. Tang, J. Woodcock, P. J. B. Jackson, F. Melchior, C. Pike, F. M. Fazi, T. J. Cox, and A. Hilton. 2018. "An Audio-Visual System for Object-Based Audio: From Recording to Listening." *IEEE Transactions on Multimedia* 20 (8): 1919-1931. https://doi.org/10.1109/TMM.2018.2794780.

Collins, Karen. 2007. "An introduction to the participatory and non-linear aspects of video games audio." *Essays on sound and vision*: 263-298.

Francombe, Jon, Russell Mason, Philip JB Jackson, Tim Brookes, Richard Hughes, James Woodcock, Andreas Franck, Frank Melchior, and Chris Pike. 2017. "Media device orchestration for immersive spatial audio reproduction." Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences.

Franinovic, Karmen, and Stefania Serafin. 2013. *Sonic interaction design*. Mit Press.

Gaver, William W. 1993. "What in the world do we hear?: An ecological approach to auditory event perception." *Ecological psychology* 5 (1): 1-29.

Heidegger, Martin, John Macquarrie, and Edward Robinson. 1962. "Being and time."

Keller, Damián, Victor Lazzarini, and Marcelo S Pimenta. 2014. *Ubiquitous music*. Springer.

Kim, Ye Jin, Byeong Ha Kim, and Ye Kyung Kim. 2019. Surround audio device and method of providing multi-channel surround audio signal to a plurality of electronic devices including a speaker. Google Patents.
Levin, Golon, Scott Gibbons, and Gregory Shokar. 2005. *Dialtones: A Telesymphony*. Staalplaat.
Lyle, Peter, Henrik Korsgaard, and Susanne Bødker. 2020. "What's in an ecology? A review of artifact, communicative, device and information ecologies." Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society.
Michel, Ulf. 2006. "History of acoustic beamforming." 1st. Berlin Beamforming Conference.
Mikolajczyk, Kurt, Oliver Bown, and Sam Ferguson. 2023. "A Study of Creative Development with an IoT-based Audiovisual System: Creative Strategies and Impacts for System Design." Proceedings of the 15th Conference on Creativity and Cognition, Virtual Event, USA. https://doi.org/10.1145/3591196.3593057.
Mikolajczyk, Kurt, Sam Ferguson, Linda Candy, Augusto Pereira dos Santos, and Oliver Bown. forthcoming. "pace Shaping in the Design Process for Creative Coding: A Case Study in
Media Multiplicities."
Molho, Lee. 1976. "Automatic acoustic-phonetic analysis of fricatives and plosives." ICASSP'76. IEEE International Conference on Acoustics, Speech, and Signal Processing.
Moore, Dylan, Hamish Tennent, Nikolas Martelaro, and Wendy Ju. 2017. "Making Noise Intentional: A Study of Servo Sound Perception." Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, Vienna, Austria. https://doi.org/10.1145/2909824.3020238.
Murch, Walter. 2005. "Dense clarity–clear density." *The transom review* 5 (1): 7-23.
Resnick, Mitchel, and Brian Silverman. 2005. "Some reflections on designing construction kits for kids." Proceedings of the 2005 conference on Interaction design and children.
Robinson, Frederic Anthony, Oliver Bown, and Mari Velonaki. 2022. "Designing Sound for Social Robots: Candidate Design Principles." *International Journal of Social Robotics* 14 (6): 1507-1525. https://doi.org/10.1007/s12369-022-00891-0. https://doi.org/10.1007/s12369-022-00891-0.
Robinson, Frederic Anthony, Mari Velonaki, and Oliver Bown. 2021. "Smooth Operator: Tuning Robot Perception Through Artificial Movement Sound." Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction, Boulder, CO, USA. https://doi.org/10.1145/3434073.3444658.
---. 2023. "Robot Sound in Distributed Audio Environments." In *Sound and Robotics*, edited by Richard Savery. Routledge.
Shneiderman, Ben. 2007. "Creativity support tools: accelerating discovery and innovation." *Communications of the ACM* 50 (12): 20-32.
Simonton, Dean Keith. 2011. "Creativity and discovery as blind variation: Campbell's (1960) BVSR model after the half-century mark." *Review of General Psychology* 15 (2): 158.
Stokes, Patricia D. 2009. "Using constraints to create novelty: A case study." *Psychology of Aesthetics, Creativity, and the Arts* 3 (3): 174.
Stokes, Patricia D, and Danielle Fisher. 2005. "Selection, constraints, and creativity case studies: Max Beckmann and Philip Guston." *Creativity Research Journal* 17 (2-3): 283--291.

Turchet, Luca, Carlo Fischione, Georg Essl, Damián Keller, and Mathieu Barthet. 2018. "Internet of musical things: Vision and challenges." *Ieee access* 6: 61994-62017.
Turchet, Luca, Mathieu Lagrange, Cristina Rottondi, György Fazekas, Nils Peters, Jan Østergaard, Frederic Font, Tom Bäckström, and Carlo Fischione. 2023. "The Internet of Sounds: Convergent Trends, Insights and Future Directions." *IEEE Internet of Things Journal*.
Turchet, Luca, Fabio Viola, György Fazekas, and Mathieu Barthet. 2018. "Towards a semantic architecture for the internet of musical things." 2018 23rd Conference of Open Innovations Association (FRUCT).
Wingstedt, Johnny. 2004. "Narrative functions of film music in a relational perspective." XXVI World Conference ISME 2004:" SOUND WORLDS TO DISCOVER", 11-16 July 2004, Teneriffa, Spain.
---. 2005. "Narrative music: towards an understanding of musical narrative functions in multimedia." Luleå tekniska universitet.