Objects and Structures: Aesthetical inquiry and artistic experimentation into the relationships between sound objects and spatial audio

Ulf A. S. Holbrook

Ph.D. thesis

Department of Musicology
University of Oslo

2022

To Hedda and Iben

Abstract

This thesis is about sound and space, and is an exploration of sounds and spaces using Pierre Schaeffer's sound object theory. It addresses aesthetic and experimental approaches to the exploration of spatial audio and site-specific practices through the intrinsic and extrinsic features of sound objects. These experimental approaches make use of software tools for composition, installation, spatial programming, and sound design, as well as for virtual reality simulation.

The main contribution of the thesis is an exploration of the relationships between sound and space, going beyond the technical issues of the spatialisation paradigm and into issues of place, site, and landscape, as guiding principles for spatial audio practices. The ambisonic soundfield is in this thesis seen as a link between sound objects and spatialisation of sound masses, sharing the same multidimensional space.

The thesis aims to study the various features of sound objects through a multidimensional model where we can access main features as well as sub-features, and sub-sub-features, of sound objects. This thesis is divided into four parts, where the first three parts discuss different aspects of the object–structure relationship, and where the last part is a discussion of possible extensions of Schaeffer's typomorphological system of identification, classification, and description of sound to encompass spatial features.

Acknowledgements

Many times I have thought about this project as a river that meanders its way through an unknown territory making many sinuous curves, loops, bends, and turns on the way to the ocean. This meandering has been a source of great challenges, of great frustrations and of great fulfillment in the end.

First, and foremost, I have to thank my two supervisors, Rolf Inge Godøy and Trond Lossius, for their patient advice as I have meandered my way throughout this project. This work would not be anything had it not been for their support and insights over the course of this research.

To all my colleagues at IMV and RITMO, especially to Anne for support in the final stages. Also thanks to Ancha, Alexander, Nanette, Målfrid, Peter, Ingrid, Victoria, Ellen, Kayla and Heidi. Thanks to Gisela for always finding the books and articles I ask for. To my friends at IMV and RITMO past and present, a deep thanks to: Tejaswinee, Emil, Tore, Áine, Çagri, Ximena, Kai Arne, Marek, Kjell Andreas, Charles, Stéphane, Merve, Bjørnar, Olivier, Catherine, Kyle, Victor and Agata.

A very special thanks to my collaborators over the course of this project: Johannes Lunde Hatfield, Tale Næss Lysestøl and Tori Wrånes.

A big thanks to Tor Halmrast for the many conversations, support and advice through the years. I have most certainly learnt a lot from you. And hopefully now we can continue some experiments together.

Thanks also to Joseph Anderson at DXARTS for opening my eyes and ears to the possibilities and depth of ambisonics through a one-week workshop here in Oslo. To NOTAM for use of the excellent multichannel Studio 3 (*Sirius*).

Thanks to Jøran Rudi for all the discussions, insights and valuable feedback on the text at its final stages.

And thanks to all the authors whose shoulders I stand upon.

Last, but most importantly, to my wife Marthe. My best friend, the love of my life - without you, none of this would have been possible. And to our two children, Hedda and Iben, without whom we would be sleeping much better but also have a much duller life.

Ulf A. S. Holbrook
Nesodden, 2022

This work was partially supported by the Research Council of Norway through its Centres of Excellence scheme, project number 262762 and by NordForsk's Nordic University Hub Nordic Sound and Music Computing Network NordicSMC, project number 86892.

Related publications

- Holbrook, U. A. S. (2019). Sound objects and spatial morphologies. *Organised Sound*, 24(1), 1–10.
- Holbrook, U. A. S. (2019). A question of backgrounds: Sites of listening. In: M. Adkins & S. Cummings (Eds.), *Music Beyond Airports - Appraising ambient music*. University of Huddersfield Press.
- Holbrook, U. A. S. (2018). An approach to stochastic spatialization - a case of Hot Pocket. In: L. Dahl & D. Bowman (Eds.), *Proceedings of the International Conference on New Interfaces for Musical Expression* (pp. 31–32). Virginia Tech.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> It's easier to "bridge" science and art when you don't really think there's a gap between them in the first place, as I don't. The boundaries between subjects are really artificial constructs by humans.
>
> *Eugenia Cheng*[1]

This thesis is about sound in space. It is an exploration of sounds and spaces, of sound objects and spatial audio systems.[2] This thesis examines Pierre Schaeffer's (1910-1997) concepts of the sound object as presented in his 1966 book on music theory, *Treatise on musical objects: An essay across disciplines*,[3] and attempts to provide a suggested extension for how we can use the theories on the sound object in the domains of spatial audio. *Space* is not something "extra" or something added to a sound after creation, rather space is always present in one form or another (Ekeberg, 2013), it is not empty or absolute - space is revealed through the different attributes of the sound matter. This introduction is a presentation and discussion of a conceptual framework that the present work is built upon and

[1]Cheng, 2015.

[2]In this thesis, "spatial audio systems" predominantly refers to *ambisonics*. Ambisonics is a loudspeaker-independent method for the capture, rendering and presentation of spatial scenes. Sounds are *encoded* using spherical harmonics and for the performance situation the sounds are *decoded* in the most fitting way to the available speaker array. Ambisonics and other approaches to multichannel sound are discussed in chapter 3.

[3]All references are from the English translation. The title in the original French is *Traité des objets musicaux: Essai interdisciplines* and will hereafter be referred to as *TOM* to be consistent with the nomenclature in Michel Chion's *Guide des objets sonores* (Chion, 2009).

discusses different models of approaches to sound and space.

Edgard Varèse iconically referred to his practice of composition as *organised sound*, where all possible sounds could be of musical interest, and thus exploding the possible scope of composition. This encompassed all ranges of acoustic phenomena as being equal in "value" to that of the small scope available within the voice and acoustic instruments. Varèse used topological and spatial metaphors to describe his work of "shifting planes, colliding masses, projection, transmutation, repulsion, speeds, angles and zones" (Born, 2013, p. 2). Not only has this discourse influenced a subsequent development of spatialisation approaches, I will argue throughout this text that these metaphors also paved the way for new approaches to understanding sound, to a practice of expanding sound to space. It links music and sound to *a* space, *a* place, and *a* location - between the listener and the sound. Practices of multichannel audio has been at the center of my practice for a over decade through composed music, sound, installations, interaction design, teaching, and theatre productions. The perspectives I present in this text are drawn from my own artistic practice and research, as well as collaborations with others in the fields of music technology and interaction design. This text focuses on the applications and design processes at the interface between spatial attributes and sound objects. Five case studies are presented in chapter 4 that are elaborations on the questions discussed in chapters 2 and 3.

This thesis is an exploration through practice-based methods into the possibilities of how we can listen to the *intrinsic* and *extrinsic* features of sound objects; that is, to be *within* or to "dwell" *inside* a sound, where we can experience the *shape* and the morphological features of the sound, as well as hearing sounds in relation to *outward* and *spatial features*, rather than just having them projected at us from a spatial location. This perspective on sound objects is an attempted development of Schaeffer's theories and draws on his insights that a sound object is a multidimensional unit. This signifies that a sound object is ontologically complex and can contain multiple significations at the same time. This multidimensional model consists of its main features, as well as sub-features and sub-sub-features (Godøy, 2021). These questions will be further discussed in section 1.3 and chapters 2 and 3.

The relationships sketched out above will be covered in more detail in later chapters. The core feature in this development of Schaeffer's theories on the sound object is defining this multidimensional entity as having of three parts: A shape, a (reference to) site, and a model. The shape refers to the intrinsic features of the sound, for example loudness, pitch, or timbre-related features related to Schaeffer's typology of objects (Godøy, 2018), this also includes the external circumstances in which the sound has been captured, as external spatial characteristics has a bearing on the perception of the sound. The site, or a reference to a site, refers to external references in the sound, for example, as to the site it was recorded or heard, which provides us some context for the intrinsic features of the sound; which can range from acoustic features, to historically and culturally siginficant features. For example, an instrument played in a reverberant room will display its own timbral qualities as well as those timbral changes imparted by the space. The final part is a model, how we choose to analyse and then synthesise what we hear. These three parts are not to be considered separate entities but are interdependent and interrelated parts of every sound object.

Backgrounds and motivations for this attempted development in Schaeffer's theories can be found in a multitude of theories on electroacoustic and acousmatic music, notably in spectromorphology. Through spectromorphology, Denis Smalley sought to develop a framework for describing the intrinsic features of sound events and their relationships. He acknowledged that "music is a cultural construct, and an extrinsic foundation in culture is necessary so that the intrinsic can have meaning. The intrinsic and extrinsic are interactive" (Smalley, 1997, p. 110). In this sense, the intrinsic carries with it a reference to the extrinsic, and Smalley uses the term *source-bonding* to represent this intrinsic-extrinsic link, of how listeners would tend to relate sounds to a supposed source or cause (Smalley, 1997). Indeed, this is also emphasised by Brandon LaBelle, in that "sound is intrinsically and unignorably relational: it emanates, propagates, communicates, vibrates and agitates; it leaves a body and enters others; it binds and unhinges, harmonizes and traumatizes; it sends the body moving, the mind dreaming, the air oscillating" (LaBelle, 2015, p. xi). From both these references we can see that the relationship between the intrinsic and the extrinsic are intertwined.

Further, the relationships between the intrinsic and extrinsic can also be considered in categories of material and structure, which has been expressed as

> the 'intrinsic' is the spectrum, its morphology, and the structural organisation concerning the spectral evolution solely as a series of frequencies and articulations, however complex. The 'extrinsic' is the sound's capacity to imply, to refer, or to associate with something other than that empirically present in the spectrum. (Barrett, 2002, p. 314)

Given these references of the relationships between the intrinsic and the extrinsic, I became interested in how this relationship could be further developed in light of theories on the sound object and of spatial practices which extend into the site-specific. This developed into encompassing, analysing, and discussing the spaces of the sounds as much as the sounds themselves, particularly to encompass ideas around site, place, and landscape. This is especially pertinent given that the "sonic image emerges, therefore, as a concept that can integrate different listening approaches and provide an understanding of both the intrinsic and the extrinsic aspects of sonic experience" (Barreiro, 2010, p. 36).

As part of a “spatio-structural theory”, Frank Ekeberg argues that the spatial elements in sound are intertwined and consists of three basic levels corresponding to source material, creative process, and listening experience (Ekeberg, 2013). Within the first level, ‘source material’, spatial elements are divided into intrinsic space, extrinsic space, referential space, and spectral space (Ekeberg, 2013, p. 2). Here, intrinsic space is sound *as* space and the extrinsic refers to sound *in* space. Interestingly, intrinsic space is understood as “an element of the individual sound, intrinsic space concerns the sound as space, and comprises components such as ‘magnitude’, ‘density’ and ‘morphology”’ (p. 2). Ekeberg’s approach has close relationships to the way this is encapsulated in the term “shape” in this thesis. The rationale behind shape is that “any feature may in principle be traced as a shape” (Godøy, 2018, p. 771), that these features are the “traits that make up the composition of the sound” (Holmes, 2012, p. 47). These features could be the amplitude profile, spectral content, dynamics, decay time and so forth, and are related to Schaeffer’s system of typomorphology, discussed in chapter 4.

Sound has been discussed by David Worrall as something which is not projected into space, and rather poetically stated as “space is in the sound. The sound is *of* the space. It is a space of sound, but there is no sound *in* space” (Worrall, 1998, p. 97). A sense of space and how this is conceived is a popular topic among many electroacoustic composers, and Suk-Jun Kim has proposed “place as a sound image” as an aid in framing the potentials of how sound is heard and understood: “spatial cues are invitations to listeners to imagine a sense of place, which may be richer in its details and more suggestive through its inferences than the cues of spatial ambiance by themselves” (Kim, 2010, p. 50). In this thesis, space is, as will be returned to in chapter 3, a boundless three dimensional extent in which objects and events occur, while *place* refers to the physical environment and what is lived. These relationships are further discussed in chapter 3.

Throughout this thesis I am concerned with the relationships between objects and structures. As mentioned previously, *objects* contain the *intrinsic*, internal, features of sounds, their shape, and our interactions with them, as well as the *extrinsic*, external, features of these sounds and their interaction with their surroundings through questions such as landscape and site. *Structures*, are the relational properties of space and describes what relational properties one is focusing on (Marcolli, 2020). The object/structure pair represents an understanding where “every object is perceived as an object only in a context, a structure, which includes it” (Chion, 2009, p. 58), but also every “structure is perceived only as a structure of objects which compose it” and

> every object of perception is at the same time an OBJECT in so far as it is perceived as a unit locatable in a context, and a STRUCTURE in

> so far as it is itself composed of several objects. (Chion, 2009, p. 58)

The object/structure relationship discussed in this thesis is presented as a proposal for applying the ideas on the Schaefferian sound object by mapping out the features of its typology and morphology[4] to that of spatial audio and spatialisation systems and practices. The object/structure pair belongs to the "two infinities of perception", where we can zoom *in* on the infinitely small when "the object is analysed as a structure of constituent objects" and we can zoom *out* to the infinitely great "when we place the object in the structure which contains it, and which can in turn be considered as an object in a context" (Chion, 2009, p. 58). Furthermore, this mapping is proposed to not only belong to a series of localised control structures but rather extend beyond the studio and concert hall to that of landscapes and *sites*. Schaeffer's achievements in his programme of musical research was the insistence on listening as a primary focus of musical research, and out of this practice grew a radical new approach to sound-based music. A fundamental component of this musical research was the sound object, and this thesis seeks to show that the theory of sound objects is still relevant today.

The sound object was conceptualized as a fragment of sound studied to access its various features, and it is used as an interface to access the shape of the sound, a location of/for the sound, and a model we use to analyse, synthesise, and represent the sound. It is a basic but ontologically complex unit of perception, and can contain multiple features and significations at the same time, which means that a sound object can display multiple properties at the same time. This is covered in chapter 2.

There are many lineages which can be traced down the path of "sound and space" apart from the practice of acousmatic music. We could, and should, discuss the traditions in computer music, live electronics, sound art, performance art, soundscape practice, site-specific art practice, improvisation and many others separately. Owing to certain restrictions around space (no pun intended) and, despite a problematic generalisation, the practices of soundscape composition, sound art, electroacoustic music, computer music, and acousmatic music will be collapsed into one category of "acousmatic music". Primarily because of this thesis' focus on music mediated by loudspeakers, however this thesis seeks to discuss practices which extend outside the traditional genre of acousmatic music. The differences between electroacoustic and acousmatic music was formulated by François Bayle as "'electroacoustic music' is a generic term that describes a technical means. It does not usually refer to a style or philosophy. The term 'acousmatic' is our attempt to delimit a particular type of electroacoustic music and a school of com-

[4] The typomorphology proposed by Schaeffer is a system of classification of sounds and is a stage in the compositional process for the exploration and description of sound features (Schaeffer, 2017, pp. 309–376). The typomorphology is one of the topics of chapter 4.

posers working within this philosophy" (Desantos et al., 1997, p. 17). In acousmatic music you might recognize the sound sources but they are out of their usual context.

Sound art is discussed differently, as it is seen to engange more directly with location than composition does. However puzzling this might seem, it is important to look past the lineages and traditions afforded by the different fields of study and rather seek to transgress formalisms and see how space, music, sound, and the listening subject all exist together. In his discussion of the subject, Alan Licht refers to Edgard Varèse's *Poème électronique* and its presentation at the World's Fair in 1958 as an example of sound art and as the first significant installation (Licht, 2009), despite the piece being presented as a concert, of a fixed duration, running in intervals for groups of visitors. Max Neuhaus pointed out that we already have names for the things we lump together as sound art, and the naming of such a broad field of practice is problematic, because

> the medium is not the message. If there is a valid reason for classifying and naming things in culture, certainly it is for the refinement of distinctions. Aesthetic experience lies in the area of fine distinctions, not the destruction of distinctions for promotion of activities with their least common denominator, in this case sound. (Neuhaus, 2000, np)

For Licht, the predominant difference between sound art and composition is one of duration, a piece of sound art has no specified timeline and yet Alvin Lucier's *I am sitting in a room* is often used as an example of sound art. Sound installations and interventions can address site and location in ways that music released on a physical medium cannot, but to say that "sound art is mainly of value in crediting site or object-specific works that are not intended as music per se" (Licht, 2009, p. 9) is an odd conclusion to draw from the discussion. Despite this categorical approach, we could perhaps say that sound art is a practice which straddles the divide between the concert hall and the gallery, sound art has a broader focus on context and concepts, and the practice relates itself to the histories and discourses within the visual arts rather than just those of music.

However difficult it is to find a definition of sound art that could describe all the disparate practices which fall under this rubric, we could, as Jonathan Sterne has described while trying to find a definition of sound studies as with "any definition of an academic field, this is a working definition, imperfect and incomplete(able)" (Sterne, 2012, p. 13). But the important lessons learnt from these definitions of sound studies is that it "recognizes sound as a problem that cuts across academic disciplines, methods and objects, ... reflexively attends to its core concepts and objects, ... is conscious of its own historicity" (Sterne, 2012, p. 5).

Spatial audio systems are models and methods for representing and reproducing sounds over loudspeaker systems. Spatial audio systems are transparent in that they do not produce sounds by themselves, rather these systems are means to represent, project, create, recreate, and structure sonic scenes and environments, be they real or imaginary; and as such, they are not transparent at all. The practice of sound diffusion through the *acousmonium* tradition is an example. Originating at the Groupe de Recherches Musicales (GRM) in 1974, the acousmonium is an approach to spatialisation which has as its design logic a *loudspeaker orchestra* consisting of different types of speakers on a stage in front of the audience controlled from a mixer, where the spectral colourations and differences in dynamics allows for expressive uses of speakers as a collection of "instruments", each with their own register (Bayle, 2007). This approach afforded music composed in stereo to be spatialized so as to enhance the spatial motions, contrasts and articulations contained in the work and can be seen as an extension of the compositional intention (Barrett, 2016).

In this thesis, space is analysed and discussed not just in terms of *spatialisation* but as elements in an artistic discourse which examines the contextual relationships between sound and space, place, site, and location. This discussion is centered on an approach which draws on elements in phenomenology, semiotics, signal processing, sound design, audio engineering, and psychoacoustics.

1.1 Motivation and context

Initially, the conceptual origins of my thesis was formed during the final semester as a Master's student in music technology at the University of Limerick, Ireland, in 2014, and grew from an interest in, and practice of, algorithmic composition. The initial aim for my thesis was to create a toolkit where an algorithmic system of attractors, swarms, and evolutionary approaches was applied to spatialiation. However, as work has progressed, it has become clear that *a framework* for thinking and evaluating sound in space, spatial experiments, and spatial thinking has been more important than the development of "novel approaches to...", which seems to be a popular rhetoric for countless conferences and publication outlets. The experiments contained in this project are built on existing tools and software, and are generally produced on a project-to-project basis that will be reflected upon in practical evaluation of each case. The attitude that any one spatial audio rendering system can fulfill any and all needs is akin to the idea that the technology alone is a solution to solve all problems in the creative process.

In all cases, the experiments and processes involved in all the different case studies has shaped the research presented in this thesis through their practice-based methods. The experiences of working "directly" with the landscape at Lista

Fyr in 2019 (section 4.6) in large parts shaped the focus and awareness of the site-specific discussed in chapter 3. Likewise, the sound design experiments discussed in section 4.4 have shaped my thoughts on the differences between point-sources and sound images, and the role of spatial authoring.

The Norwegian composer Knut Wiggen often criticized electronic music composers for having a technical focus and losing sight of artistic intentions and productions by opting for "novelty" (Rudi, 2018). In discussions on any audio system, it is important to be mindful of the ground we are treading, the music is, if not technologically *created*, it is technologically *mediated*. Spatial audio describes a set of tools and methods for representation and control of sound material through spatialisation/diffusion over multiple loudspeakers. Specifically we have to be mindful about the differences between *techniques* and *technologies* (Baalman, 2010). Techniques are descriptive of a compositional process, while technologies are descriptive of panning, speaker arrays, encoding/decoding functions, and so forth.

I have increasingly, after attending concerts of multichannel audio, over the last 10 to 15 years, found that the pan-pot is for many composers today what the pitch-bend wheel was for electropop outfits in the 1980s. Many composers seem to consider panning to be some form of novelty where individual sounds are moved around the room, rather than considering the possible implications of spatial composition and the possible constructions of complex spatial scenes. The implications of spatial composition extends beyond the *point-in-space* paradigm where individual objects are panned, almost naively, around a space without much thought to the *sound image* that the objects form together. The sound image is a metaphor for the possible mental images the heard sound evokes in the mind of the listener. This is discussed further in section 1.3.

A primary motivation behind this thesis has been a curiosity and a desire to explore the possibilities and limitations in current practices of spatial audio applications and sound composition. Frequently, when approaching *spatialisation*, it is discussed as the last step in an audio effect chain (and in a chain of reasoning) that often can be taken as a practice of *adding* space to sound, indicating that in some way spatial dimensions are not already present in the material. Despite multichannel diffusion of music being an integral part of the early developments of electroacoustic music, many composers still work in the stereo format. Among the issues many composers have cited in a survey from 2011, is that many venues would require an ad-hoc setup, which could make equidistant loudspeaker setups difficult and also that ideal placements of speakers would not be possible due to constraints with stage or lighting design (Peters et al., 2011). These problems could be among the reasons that are "holding spatial music back" (Lyon, 2014, p. 851). An important feature for loudspeaker-independent systems such as ambisonics is the flexibility of rendering the sound materials to best fit the venue and the available loudspeaker setup. A guiding idea in this project is to consider sound and space as being fundamentally linked, where the spatialisation of sound is not an after-thought when the compositional process is completed but is an integral part of the creative process.

Spaces and sound, seen from spatialisation and spatial audio systems, is about uses of space as an element of an artistic discourse and of the space itself as a control and musical parameter. It is a practice of *articulating* space, and the multitude of ways this can be achieved; for example, through an intentional design of sound scenes, spatial scene composition, and notions of site-specificity/place-making (this will be discussed further under the heading of *sound landscape* in chapter 3). Schaeffer and colleagues wanted us to make music that was based on concrete, real-world sounds. Their work used the technological means available to them at the time, and the results and conclusions drawn from their experiments are historically coded by this equipment. The limitations in terms of equipment at the time meant they could only record in mono (later in stereo) but although they explored space in composition with multiple mono tape machines, such as in *Symphonie pour un homme seul* (1949/50), the available technologies made the reproduction and exploration of spatial features difficult. That being said, the ideas presented in the programme of musical research are fully relevant today, specifically as we are steeped in an age of technologically motivated and driven music-making: the relevance of Schaeffer's musical research is on the perceived outcome of the processes not the processes themselves. When stating that the medium is no longer the message, but that specific tools have become the message, Kim Cascone emphasise the fact that the unique fingerprint gained from any system is the artifacts of that system's construction (Cascone, 2000) and this will ultimately be part of the process.

Why should we only be content to model spaces or existing rooms when we have at our disposal the means to create new sonic "worlds" that are derived from concrete, real-world spaces, and experiences, and yet transgress them? The *Gmebaphone* (Clozier, 2001) is constructed similarly to a loudspeaker orchestra, but used a purpose-built frequency splitting device and spatializes sound based on frequency and is seen as an orchestration generator (Emmerson, 2007). Mathematical models provides answers on recording and reproduction of spatial scenes and can calculate the "correct" representation of a spatial model, yet these models are not designed to capture or address the complexity of musical material and perceptual experiences (Carpentier et al., 2016). Why, then, is the focus on space and not just on sound? Because without space as *context*, the sounds have no meaning. Indeed, "the context of a sound object is the whole structure in which it is identified as a unit and from which it is extracted to be examined individually" (Chion, 2009, p. 63). Furthermore, without a clear understanding of the differences between spaces, places, and sites (studios, concert halls, galleries, the outdoors), this “space” has no meaning. If we remove the sounds of the surroundings, then it “ceases to be a living, breathing part of nature, losing much of its depth and becoming just another picture” (Malham, 2001b, p. 32).

As we progress to discuss sound objects in the next chapter, we must never loose sight of the spatial correlate that always follows the individual sounds.

1.1.1 Research questions

Based on the elaborations in the preceding sections, I will formulate the following research questions for this thesis:

- How can we extend the theories of the sound object to include spatial thinking and practice that engages with ideas of site and landscape?

Discussions around this question are made through approaches of *practice-as-research* and *analysis–synthesis*, from where I will formulate the following sub-questions:

- To what extent can we build upon or extend Pierre Schaeffer’s typomorphology to discuss and describe spatial attributes?

- To what extent can we use the discourse bound into maps, mapping, site, landscape, surface, shape, and object to continue to frame a discourse around the practices of spatial composition, from artistic, aesthetic, philosophical, technical, and practice-based perspectives?

My assumptions through out this text is that space is intrinsically linked to sound and as an extension the sound is linked to a site/location. Our perceptions of sound belong to a complex mesh of interactions between people, histories, technologies and ideas, and I hope to shed light on these relationships.

1.2 Methodology

If we go back to the earlier discussion on sound studies, we saw that sound is recognized as a problem which “cuts across disciplines” and that “all disciplines begin as interdisciplines” (Sterne, 2012, p. 5). In the context of the musical, artistic, technological, and philosophical research project I am presenting here, I will delineate two methodological approaches that form the central strategies I have employed throughout this project: *analysis–synthesis* and *practice as research*. These strategies build upon each other and are historically significant in the development of “experimental”, computer, and acousmatic music.

1.2.1 Analysis – synthesis

As an approach, analysis–synthesis[5] can be defined as the systematic exploration of features, it is a method to understand the world by breaking it into smaller parts and looking at the possible interactions between the parts and their surroundings. *Analysis* refers to the decomposition of something of varying degrees of complexity into smaller parts, or elements, and this also includes interactions and perspectives. *Synthesis* refers to the operations involved in putting these decomposed elements back together as themselves, as new configurations or as the combinations of interactions (Risset, 1991; Wright et al., 2000). This will be discussed further in section 1.3 under the domains of semiotics and actor-network theory.

Discussing different layers of abstraction is in itself an analysis, as are different layers of interactions. The temporal unfolding of sound events are subjects for analysis and we can apply methods of analysis to sound scenes and field recordings in order to synthesize the temporal, spatial, and acoustical developments within a recording for use in a composition. When we make incremental adjustments of a sound to make it sound "just right", we can also make incremental adjustments to the spatial positioning of sounds to make their spatial locations and relationships coherent within a framework of spatial composition. We use approaches of analysis to hear the differences between sound scenes, as to what the differences are between a forest, a village, and a city soundscape. We will also use analysis to uncover the differences between similar sound environments. Processes of analysis were important for Schafer and the World Soundscape Project in arriving at the categories of *key sounds*, *signals*, and *soundmarks* (Schafer, 1994) as ways to describe the sound environment, as will be discussed in section 2.7.4. We use the same approach when we are listening to the differences between different forests and different cities. When synthesizing a spatial sound scene, we can draw upon these analyses to synthesize a scene that incorporates the insights learnt from analysing the recordings. In terms of sound synthesis, the process does not have to be built upon a preceding process of analysis.

When making something sound a particular way, either through synthesis or through manipulation of a recording, we would usually pass through several steps of adjusting equalization, filters, delays, compression, and so on where the aim is to find the most suited sound for our purpose. This is a process of *approximation* where each iteration brings you closer to the desired result. Drawing on the synthesis of timbre (Risset & Wessel, 1999), analysis and synthesis was defined as the systematic exploration of trial-by-error, or "learning-by-doing", as a means to "explore categories, categorical thresholds, inter-categorical and intra-categorical variations of sonic objects" (Godøy, 2018, p. 776). Analysis is a method for gain-

[5] The classic term *analysis-by-synthesis* (Risset, 1991) is in this thesis referred to as *analysis–synthesis*, to emphasise that analysis and synthesis are, or can be, two separate processes.

ing insights into why a flute sounds the way it does and what makes it sound different from an oboe.

If we consider Robert Rowe's idea of *composition by refinement*, a program reacts to changes in configurations and inputs from the user which leads to a process of continuous refining of ideas up to the development and articulation of musical ideas (Rowe, 1992). This perspective points to a more heterogeneous approach to what analysis–synthesis *could* be, where the structuring process is guided by the ear and by sound. We can begin simply by saying that a series recorded sounds are first of analysed for their intrinsic properties, then, based on compositional intent, the composer can synthesize new sounds drawing on the analysis of the first sounds. This is an endeavour where the articulation of musical ideas contributes to the dual unfolding of analysis and synthesis.

An obvious further example can be found in Risset's experiments with the synthesis of a bell, or bell-like sound, through additive synthesis. Taking the sound of an existing bell, Risset set up a table of frequencies, amplitudes, and durations of the bell's partials (Dodge & Jerse, 1997, pp. 104–105). The process involved analysing the original sound, synthesizing the sound, then reanalysing the resulting sound in comparison to the original analysis. A bell approximation is generated with one envelope shape and a set of frequency components with different decay times, and if the envelope shape is changed from the sharp attack and exponential decay to a longer attack and decay, then it changes from "bell" to "fluid texture" (Risset, 1991).

In these instances synthesis builds on analysis. If we seek to construct scenes in sound design for film or games set on a different planet, then an analysis of places and sites are integral. Through synthesis we can apply a range of effects drawn on the preceeding analysis, through for example spectral processing and time-stretching to create an alien environment with a different sense of time and gravity than our own environment, but still something based on our own environment. This alien environment draws on the analysis of a real-world sound environment, and can be used to synthesize an artificial sound ecology that becomes artistically credible because it transposes a recognisable sonic eco-logic onto this other world with its differing environment and physical conditions. Our approximations of the new alien sound environment is based on an ecological understanding of sound environments that originate from analyses of the real world. These perspectives are also mirrored in work on procedural audio, where "any sound can be generated from first principles, guided by analysis and synthesis. ... Sounds which are impossible to record becomes accessible" (Farnell, 2010, p. 1). We then use the analysis–synthesis pair to create sensations of alternative realities or indeed to create virtual landscapes. This will be further explored in the case study for *City Dwellers II* in section 4.7.

1.2.2 Practice as research

Creative exploration is a pathway to where new insights, knowledge, and understandings come into existence. The program and methodology described by Schaeffer for musical research (Schaeffer, 2012, 2017) had at its centre an insistence on *listening* and repeated musical *experimentation*. These processes were explorative through technology and mechanical manipulation of recorded sound matter, with the end goal of achieving deeper and wider understanding of the sounds that were the basis for musical creation. This approach belongs to the tradition of "recherches musicales" and guided the creation of the (GRM) as well as the later establishment of *Institut de Recherche et Coordination Acoustique/Musique* (IRCAM).

A creative process is not something that is unique to musical research, but is something that, methodologically speaking, permeates all disciplines (Borgdorff, 2012, pp. 37–53). Research that is dependent on artistic production in some sense or another, where the artistic outcome is deemed inseparable from the research itself, is often referred to as *artistic research*. In the "debate" on artistic research, Henk Borgdorff recognizes three definitions in a terminology for discussing artistic research (Borgdorff, 2006). First, *research on the arts* aims at drawing theoretical conclusions on artistic practice from a distance. This type of research is well established within the academic disciplines of musicology, art history, literature studies, and media studies among others. Second, it is an "instrumental" perspective as *research for the arts*. The theoretical analysis of art is not the primary focus, rather this category focuses on development and extensions of tools and technologies for artists to practice within their field. Lastly, *research in the arts* is the type of research where there is no theoretical or practical separation between artistic practice and research. This is arguably the most controversial of the three categories, as there is no separation between concepts, theories, knowledge, and experiences.

Borgdorff's tripartite perspective is drawn from an earlier definition by Christopher Frayling where he defines artistic research as *research into art*, *research for art*, and *research through art* (Frayling, 1993). Specifically it is the last category, that along with Borgdorff's *research for the arts*, are important methodological pathways. *Research through the arts* provides an interface for flexibility where a wider set of research results are possible, as the methodologies can embrace the practice and outcome of artistic projects. In a similar vein, Robin Nelson has argued for *Practice as Research* where practice is a key method of inquiry and where practice "is submitted as substantial evidence of a research enquiry" (Nelson, 2013, pp. 8–9).

As a means of articulating the *movement of research* within art, Michael Schwab has proposed the term *transposition* to show how something changes its identity when moving across different instances. To contextualize this, Schawb uses the example of Marcel Duchamp's *Fountain* (1917), where a object of utility (a urinal) is moved from its original context and placed in a gallery to become art. A transposition does not merely describe operations within artistic research practice, it is also "an operation *with* artistic research - that is, 'artistic research' emergent as transposition of a project, as speculation on how else knowledge can be gained and what notions of knowledge and perhaps even art are suitable to capture a project's achievements" (Schwab, 2018, p. 9). A transposition considers the flow of data from one instance to another and is seen as "the creation of an experimental context in which epistemic things can be explored for the sake of gaining knowledge about them" (Rheinberger, 2018, p. 215). This movement of research across disciplines and within art, from low-level abstractions to artistic practice, is a method for evaluating features of sound and space, like that of analysis–synthesis.

In between these different notions of research is where I situate my work, where research *for* the arts and research *through* art are important perspectives. Within practice as research, there is a "methodological abundance" (Hannula et al., 2014, pp. 20–33) which affords us a wide field of interactions. Likewise, as we shall see in the following chapters, issues of transpositions and of practice as research are important touchstones in the way this project has refrained from using controlled laboratory experiments in favour of more ecologically valid results, with a strong focus on testing theories on spatial practice in real-world contexts: "where psychoacoustic research tends to focus upon noticeable differences in pitch, spectral structure or duration, ecological research attempts to identify the transformations in acoustic structure which inform the listener of some important change in the environment" (Windsor, 2000, p. 13). As such, I am interested in the affordances of practice in research. The case studies in chapter 4 are all explorations of these different perspectives and explore how practice is a guiding principle of analysing and synthesizing sounds and spaces. These aspects are viewed through the lens of experience and practice, but the ensuing discussions leading up to presentations of the projects are contingent on the relations between the multiplicity of possible interpretations and the historically situated negotiations. The artistic output is not considered the primary focus of the research, rather as a contextualizing bridge between theory and practice - as the result of repeated processes of analysis and synthesis.

1.2.3 Some comments on interdisciplinarity

Any project which moves across disciplines is a risky one. The primary problem is that one person cannot have an expert knowledge of all the different fields and this can lead to frequent "cherry-picking" of ideas, concepts, methods, approaches, and, ultimately, conclusions. Schaeffer's *TOM* has the subtitle of "An essay across disciplines" and is at its core about music and its relationship with technology, listening, acoustics, psychology, subject and object, philosophy, and physics. As we saw at the opening of this section, we can regard all disciplines as starting out as inter-disciplines (Sterne, 2012). The technological, artistic, and philosophical issues discussed in this thesis span a wide range of disciplines and I have no pretension of having an expert knowledge in all these fields.

Any work which deals with multiple conceptual systems will involve lesser or higher degrees of approximations, as similar terminology and concepts can vary between different fields of expertise. We can see that when an attribute has a 1:1 relationship with the terminology assigned to it there is no ambiguity involved (Berg, 2009). However, when this is not the case this is described by the duality of *contrast* - where an attribute is referred to by different terminologies, and *conflict* - where the same terminology is used for different concepts (Shaw & Gaines, 1989). For example in evaluation of attributes related to spatial quality, the attribute *spaciousness* (which describes the size room we perceive to be in) has been replaced by the terms *immersion*, *envelopment* or *ambience* (Berg & Rumsey, 2003).

In musical research this can pose a problem, as we quickly can use multiple methods from multiple fields, without necessarily being aware of hidden ideological and theoretical foundations of the different fields. Given these potential problems, the ecologically valid approach sketched out in the preceding section affords this project a complement by situating it outside the laboratory and at the site where the sounds and events of the real world are experienced.

An important insight from Schaeffer's work across disciplines is the listening subject and the focus on the subjective listening experience. This listening experience is guided by the subject's past experiences, knowledge, and how the listener directs attention and conscious *intention* towards different features of a sound. The differences in focus afford an approach, like that of Husserl's direct consciousness to an object (A. D. Smith, 2003), which incorporates living beings, physical objects, actions, and processes. An interdisciplinary approach to listening and to musical research affords us access to the complexes of interactions in the material.

1.3 Representations

Sounds in acousmatic music can often be perceived as resembling something from the outside world. The real-world context of sounds and their spatial interactions point to a meaning conveyed to us through significations and how these significations are *represented* through sound and space. To convey meaning posits that "no sound event, musical or otherwise, can be isolated from the spatial and temporal conditions of its physical signal propagation. Secondly, sound is also shaped subjectively, depending on the auditory capacity, the attitude and the psychology and culture of the listener" (Augoyard & Torgue, 2014, p. 4). This section will discuss the conceptual framework underpinning this thesis, through the significations and relationships around these objects of perception.

In *The death of the author*, Roland Barthes refers to a text belonging to a "tissue of quotations", or a "fabric of citations" (Barthes, 1977, p. 146), a multi-dimensional space where innumerable citations within a text "blend and clash". This fabric extends far beyond the mere text and exists within the reader's consciousness. Given this perspective the *site* of reception is far beyond the control of the author. This refers to the relationship and *interaction* between the work and the recipient, and the *semiotic* significance in the content of the works. An "author" creates a work with specific intentions, which is interpreted by a reader/listener in the multiple and myriad ways this can happen. There is not one single interpretation but rather a range of possible interactions with the listener's past experiences and knowledge.

Our interaction with the world consists of complex meshes of interactions, and through semiotics we can investigate relationships between symbols and meanings, between words and the understandings they evoke. Not just bounded by studies of language, Jean-Jacques Nattiez provided a foundational discourse of musical semiotics (Nattiez, 1990), where he used a tripartite scheme of *poiesis*, *esthesis*, and the *neutral level*, where "poiesis refers to the creative processes that generate a work and can include authorial intentions; esthesis refers to processes that receivers undertake when interpreting a work . . . [T]he netural level, the trace or physical embodiment of the work that is accessible to the senses" (Demers, 2010, p. 25). Within acousmatic music the mimetic and surrogate relationship between sounds helps to relate an "abstract music" to something outside of the musical work itself.[6] The sounds are signifiers, which point towards some signified content and is related to an image or a narrative by the listeners. Ferdinand de Saussure defined a *linguistic unit* as a double entity, a combination of a *sound image* and a *concept* and the combination of these as a *sign* (de Saussure, 1959). In the original French, the term for a sound image is *image acoustique*, and in other translations this is referred to as a *sound pattern*. Yet, as ambiguity can make the terms difficult to understand, Saussure proposes to replace the terms sound image and concept with that of signified and signifier (de Saussure, 1959, p. 67).

The term *image* and specifically *sound image* has become ubiquitous in the discourse of acousmatic music, to describe the possible narratives or mental images evoked by the music. For example:

[6] "Abstract music" is a term often used to describe music that is *non-representational*, where the sounds (and sources) are unrecognizable.

- On the one hand, sound-images of the voice, or animal and bird cries, have an intrinsic gestural content . . . transformation now becomes the gradual changing of one sound-image into another with its associated metaphorical implications, and a landscape can be seen as a particular kind of timbre-field applying to the space of sound-images. (Wishart, 1996, p. 164)

- Mediatic space creates not so much a direct spatial form, but an image of spaces and places, events, distances, which impinge on, and form part of the spaces within which we act. (Smalley, 2007, p. 39)

Given these statements, we can see the reference to image and sound image clearly indicates that the relationships, contexts, and interactions between sounds creates a perception of mental images in the listener with associations to the outside world. Kim argues that listeners of electroacoustic music imagines a relational and developmental process among the sounds they hear: "The everydayness of the sounds emboldens them to 'picture' what they are hearing as they take cues from the way the sounds sit, move, integrate, disintegrate, or generally relate to one other" (Kim, 2010, p. 46). Saussure emphasised that signs come in systems, where the relevance of each sign is dependent on the positions and relations in the semiotic system. Charles Sanders Peirce highlighted the insight that signs *mediate* meaning between the signifier (the sign), the signified (the idea), and the interpretant (the link between the signifier and signified) (Peirce, 1960). Peirce emphasised that the relationship between a given token and the its object is dependent on the relation where the token is interpreted.

We must be clear that a *musical sign* should not be confused with notation or other graphical approaches to represent music: "Music is a sound index, a morphism, a cross-domain, inference-preserving mapping from thought to sound, without the mediation of symbolic language" (Scaletti, 2018, p. 379). Despite Saussure's theorizing of the linguistic sign as an arbitrary relationship between the signifier and the signified, the meaning in the musical sign is determined by the intrinsic properties of the object itself (Chion, 2009). Schaeffer denoted three terms for considering the sign in musical research: the signal, the index, and the sign (Schaeffer, 2017). If we listen in order to recognize a cause or the event of the sound, then we refer to the sound's *index*, but if our listening intention is towards an understanding of a system of values or language, then we are listening to a *sign*. The signal denotes the *physical signal* is it exists and is perceived within the acoustical context (Chion, 2009, p. 89).

From this we can discern the sound object's relationship to its surroundings, specifically as a *mapping* between *structures* - and as mappings between sign systems. The relationships between objects, or between *aspects* of objects are what

Peirce referred to as the *index*, and this links the signifying element of one object to another (Peirce, 1960). The link can be the event or cause of the sound but also the feature similarities between sounds. When we examine the spatial context where we perceive sound objects, in the performative context of concerts, installations, and the like, then the index of a sound links what we hear to a perceived origin. The sound sequences and constructions signify some form of origin informed by both the sign and the signal, as for example in Bernard Parmegiani's *Sons-Jeux* (1988) (Parmegiani, 2010), where rhythmic train sequences are mirrored in vocal sequences, with and without heavy processing, with silences and long textural sections, this piece creates sound images which have both real and imaginary origins.

An index linking one aspect of an object to another, has also been referred to as a *semiotic morphism*; a mapping of some features of a source to a target, specifically in terms of design objectives. Interfaces, metaphors, and representation have much in common. Joseph Goguen states that:

> A coffee cup is an interface between the coffee and the coffee drinker; questions like thickness, volume, and handle shape are interface design issues. A book can be considered a user interface to its content. Buildings can be seen as providing interfaces to users who want to navigate them, e.g. a directory in the lobby, buttons outside and inside elevators, "EXIT" signs, doorknobs, stairways and even corridors (you make choices with your body, not your mouse). (Goguen, 1999, p. 243)

From this we can draw another parallel and state that the loudspeaker is an interface between the listener and the sound, a microphone is an interface to listen from one specific location at or over a specific time period, the viewfinder of a camera is an interface to see the surrounding landscape from a specific location at a specific angle.

Meaning is created by the ability to relate what we hear to our past experiences and how we situate this along a horizon (Nattiez, 1990), whether something metaphorical is rendered literal and can be better related to a totality of our experiences. This insistence is also reflected in attitudes to the signified meaning in music, where

> musical sounds are above all signals within a semiotic system. Their use is a function of very special devices for synthesis and analysis of physical sounds. ...As physical events, musical sounds are always natural, be they produced by a live violin performance, a computer driven synthesizer via loudspeakers or by the tape patchwork of musique concrète. (Mazzola, 1997, p. 2)

We could easily add a prefix of “spatial” to what we are discussing, as in *spatial audio, spatial mapping, spatial hearing*, and so on. However, this would miss the mark. It is important to not lose sight of the fact that when we experience something in one setting, a different setting can impose great changes to this. A change of representation - for example a sound heard over headphones and then over a large array of loudspeakers - is a morphism, a mapping between two sign systems and by just denoting one as “spatial” and one as not, would not in any way describe the essence of the representations. These mappings between sign systems have been referred to as *cross-domain mappings*, and this leads us into the role of metaphors and analogies.

1.3.1 Metaphors and analogies

Metaphors are interfaces to understand and experience something in the terms of something else. In thought, action, and language, metaphors are pervasive (Lakoff & Johnson, 2008). Loudspeakers have been referred to as a “window into space”, a loudspeaker is physically no such thing, it is an electroacoustic transducer, which converts digitally stored samples to fluctuating voltages to create the sound we hear. However, a metaphoric image of the loudspeaker is of a thing that facilitates the mediation of potentially unknown worlds.

Metaphors must be paired with a target to function and in the theory of structure-mapping (Gentner, 1983; Gentner & Bowdle, 2008) this approach combines metaphor with analogy and similarity. Here, an analogy is a mapping of knowledge from one domain (the base) into another (the target) which conveys that if a system of relations holds among the base objects, it also holds among the target objects. The mapping system holds higher-order relations, rather than isolated predicates (this is explored further in chapter 4).

Thomas Nagel posed the question “What is it like to be a bat?” in an attempt to address the issue that consciousness is of a fundamentally subjective character that cannot be reduced to physical components alone (Nagel, 1974). Imagining what it is like to be a bat is not the same as being a bat, and in attempting to understand how something *is* we break it down into smaller pieces. Where Nagel sees that “any reductionist program has to be based on an analysis of what is to be reduced” (Nagel, 1974, p. 437), it is equally important to examine what we reduce something *to*. Nagel wanted to know what it is like for a bat to be a bat, yet when trying to imagine this one is limited by “the resources of my own mind, and those resources are inadequate for the task” (p. 439). Indeed, even if some external evidence can provide a clue for how it perceives, this experiential perception is still withdrawn (Bogost, 2012). In Nagel’s text, this is about the mind-body problem and “the character of the experience of something is not identical to the *characterization* of that experience by something else” (Bogost, 2012, p. 63).

Musically, we can take the example of Barry Truax's *Riverrun* (1986) (Truax, 1988). The metaphoric construction of the piece is of a river, always moving and flowing, yet permanent. The piece is built upon two 170ms phonemes (Roads, 2004) as its source material, yet, this information is not necessary as it is the title and the subsequent understanding of the sound materials that are interesting in the interpretation and experience of the piece as a *river*. *Riverrun* was created using granular synthesis, a technique where small grains of sound between 30 and 100ms are grouped together in different densities. The constant motion from small individual grains of the opening build in volume from droplets to large masses of sonic densities. The sound is constantly in motion, as of that of a river.

In case study 3, in section 4.6, the installation *Superimposed Landscape (Lista)* is discussed. The conceptual premise of the installation is metaphorical, it is a sound installation which samples from the surrounding sound environment in real-time at irregular intervals, loops these samples back into the gallery space and over time layers different samples taken at different times from the surrounding landscape. This process is reflected in the title of the installation, where it seeks to *superimpose* the sound of the landscape onto itself.

To express sounds and sound experiences we use metaphors to describe their features, whether the sound is synthesized or a field recording. The uses of metaphors can help composers and listeners to describe something vague as more tangible. For example, a kick drum can sound "boxy" at 500Hz, cymbals have "ringing overtones" at 1-6kHz, pianos can have "presence" at 2.5-5kHz and strings exhibit "shrillness" at 8-12kHz.[7] Likewise we can use a series of visual metaphors to describe sounds such as "rough", "smooth", "texture", "contour", and so on. The use of metaphors as a language to describe perceptions of sound can be a means of explaining a mental image of a sound (Porcello, 2004) and even by adopting metaphors from other fields as a means of sensory evaluation, such as using terminology from the wine industry to describe features of concert hall acoustics (Lokki, 2014).

The use of language such as "*in* the sound" and "*into* space" are container metaphors which we use to describe sound, space, and loudspeakers: "each of us is a container, with a bounding surface and an in-out orientation (...) moving *out of* one room and *into* another (...) events and actions are conceptualized metaphorically as objects, activities as substances, states as containers" (Lakoff & Johnson, 2008, pp. 29–32). This is for example illustrated in Anders Vinjar's *Le camere invisibili* (2016)[8] which, inspired by Italo Calvino's 1972 book *Invisible cities*, seeks to create music which moves around in rooms of many forms, it elab-

[7]These subjective audio quality categories are based on the "Audio frequency chart" from Sound on Sound: https://www.soundonsound.com/sound-advice/sos-audio-frequency-chart.

[8]https://www.avinjar.no/works.php

orates large and small rooms, from real to impossible, non-existent to strange and common rooms. Centrally, the piece explores what happens when audio meets boundaries where space breaks down and how these spatial elaborations create rooms within rooms.

We can nest container metaphors further by describing what the room we occupy "sounds" like, be it "boomy", "dead", "reverberant" and so on, where these metaphors refer to the decay times of the room and how this influences the perception of the sound played through *this* room and how it influences our presence *in* the sound. Some uses of metaphors in the audio industry are problematic, which include the language used for clock sync between hardware devices, described as master/slave - based on language derived from colonialism.

1.3.2 The orientational metaphor

When we discuss spatial features, we deal with both *actual* spatial orientations as well as *orientational metaphors* (Lakoff & Johnson, 2008). Our sound experiences are often divided into direction and orientation, left/right between traditional stereo channels, coupled with front/back for surround sound systems, and further still with an added up/down for three dimensional reproductions. These orientations can be exploited for use in compositions given the functioning of our human auditory system, and importantly, "most of our fundamental concepts are organized in terms of one or more spatialization metaphors" (Lakoff & Johnson, 2008, p. 17). This highlights the importance of spatial awareness not just in terms of human sound localization and navigation but also in how we humans describe our experiences through the use of these orientational metaphors in everyday language. An awareness of how these metaphors function is integral for the unfolding of structures and relationships in the composed sound image.

To generalize this last section, we can say that the uses of metaphors for descriptions of sounds is a way of describing *experience*, the choice of metaphors and the way we use them are not arbitrary - they are grounded in cultural and physical experience of the world we live in. We use these experiences to understand something in terms of something else and we orient these experiences along a set of axes.

1.3.3 Axial thinking

Through the preceding sections, the object/structure relationship has been represented as a differentiation between the *intrinsic* and *extrinsic* features of an object and the *contextual* features of a structure, to focus on the potential multidimensionality of features. An *axis* represents a continuum for how we can evaluate the feature dimensions of different objects and the relationships between them.

Rather than considering strict categorizations of features, we could think about the opportunities afforded as a field of possibilities existing along a series of *axes*. An axis is a continuum through space rather than a static point; for example "a continuum of possibilities between two extreme positions: so the axis between black and white is a scale of greys" (Eno, 1996, p. 298).

Axes have been defined by different terms; for example as the "perceptual field" (Chion, 2009, pp. 64–67). Through this field, sound objects can be assembled into significant structures. This points to the relational nature of this perspective. Axes define different features of a multidimensional space and the shape of the object which occupies this space. The organisation of shapes from the features of the objects themselves helps us to define and delimit some space in a minimum and a maximum value per axis. We can consider axes as existing in multiple directions in an N-dimensional space, but with certain defined limitations of acceptable values.

Here we can adopt a perspective from morphodynamic theory, a dynamical system of forms, patterns, and structures (Godøy, 1997; Petitot, 1999, 2011), where this is divided into a *control sphere* and a *morphology sphere*. Each "sphere" represents a structure that can range from linear to nonlinear features. The control sphere includes the input, the control parameters and processes, while the morphology sphere defines the perceived output. These processes can be related to the transformations of sound materials from a recognizable source, through a restructuring of the sound's features and to the spatial presentation. This is discussed again in section 2.5 from the concept of *blackboxing*.

From at the opening of this chapter we saw that Neuhaus stated "aesthetic experience lies in the area of fine distinction" (Neuhaus, 2000, np), and the differentiation afforded by axes are interfaces to exploring subtle values and changes in the features of one object, between groups of objects, and the structural context which includes the objects. Roads considers a repertoire of "spatial oppositions" as an attempt at charting spatial operations for composition of spatial music (Roads, 2015, p. 279). The list sets a system of either/or to how we can articulate sounds in space, rather than thinking about the continuous and possible values between categories. He defines an opposition between "foreground" and "background", "slow motion" and "fast motion", "fixed position in space" and "moving position in space" whereas the interplay and the juxtapositions between the operations are more important than the oppositions between them. Foreground and background are not opposites, they are dependent on each other, as shall be discussed further in section 3.6.1. A composer could employ both operations on a single sound object to create tensions and motions through space by harnessing a rich interplay between delay, spectral colouration and dynamics.

In Figure 1.1, the (Cartography)[3] (MacEachren, 1994) presents key distinctions of evaluation in cartography, where the different criteria for evaluation are placed at opposite ends of several axes so when read together, it generates a point in 3D space. This has been adopted as the Sonification Cube in (Scaletti, 2018).

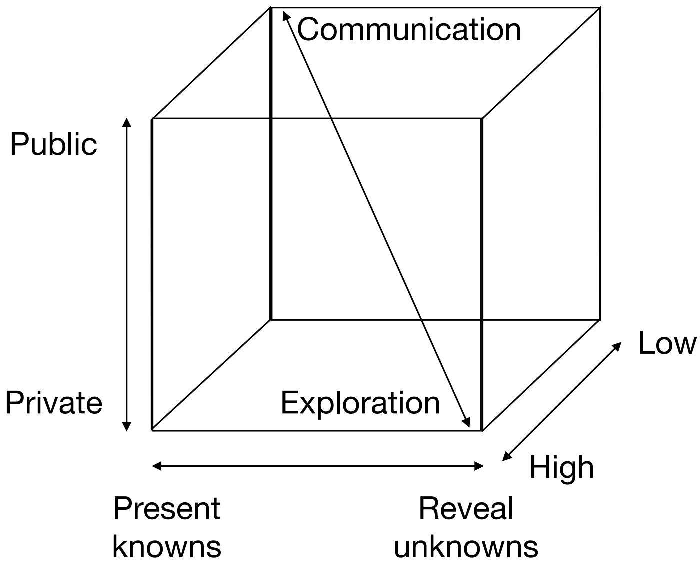


Figure 1.1: The (Cartography)[3], a three-dimensional representation of knowledge creation in cartography along four different axes.

All the axes which make up Figure 1.1 are drawn from a framework for geographic visualisation (DiBiase, 1990), where each axis represents two values, and our navigation within this three-dimensional space is a point which is the intersection and interaction of all the axes:

- The axis of *private* ⟷ *public* denotes if something is created for ones own need to something which is made available for a larger audience.
- The *unknowns* ⟷ *knowns* axis describes if the map is a general way of searching for something to where the user is attempting to access existing information.
- on the *high* ⟷ *low* axis the degree of human interaction is measured, where a high degree indicates that a human interaction can change the map in substantial ways to where the user has a limited ability to change or influence the representation.
- The axis of *exploration* ⟷ *communication* describes to what degree the spatial navigation is an exploration of a map to communication of results.

Drawing on this figure, we can say that working with spatial audio systems can offer possibilities of navigating a multidimensional space of creative and expressive possibilities. These axial relationships are reflected in the multidimensional model where we can access main features, as well as sub-features and sub-sub-features, reflected in the typomorphological framework. Axes should be considered as an approach to multidimensional modelling where, within a given space, we can have an (in theory) infinite number of axes. The location of one axis will be seen in relationship with points on the other axes. The cube-model has also been adopted elsewhere, Grey and Gordon propose a three-dimensional model for spatial resolution of similarity matrices for investigations into the perceptions of timbre (Grey & Gordon, 1978). Likewise, Wessel created two-dimensional representations of timbre spaces, where specific coordinates could be used to hear the timbres at the specified spatial location (Wessel, 1979).

If we consider axes to be universal, they can be applied to more or less anything, whether it is pitch, density, texture, presence, disorder, distribution, proximity, spaciousness, and so on. Within such a multidimensional space, we approach what Bjørn Fongaard referred to as *N-tonality*, the infinitely available frequencies and combinations of frequencies within the human audible range (Rudi, 2019).

We need to define some form of *space* for the axes, rather than using a representation of three-dimensional space as in Figure 1.1. We could look at the relations between axes and between clusters of axes and organise them in a network of interrelated parts which could enable us to consider the effect each individual axis have can have on the resulting sounding matter. Looking at a multidimensional feature space, we then become concerned with the complex meshes of interactions between the different parts and "by placing all its components in continuous variation, music itself becomes a superlinear system, a rhizome instead of a tree, and enters the service of a virtual cosmic continuum of which even holes, silences, ruptures, and breaks are a part" (Deleuze & Guattari, 1988, p. 11). We can then extend the notion of axes in objects and structures to the concept of *actors*.

To explore this, I will borrow some ideas from the *Actor-Network Theory* (Latour, 1996) to draw up a framework for discussing the relations between the objects, structures, and a multidimensional space of axes. In actor-network theory anything can be considered an actor. This means that human and non-human actors are considered as equal, and no distinction is made between the natural, the social, or the technological. An actor is an association of heterogeneous elements constituting a network, presented as a flat ontology (Latour, 2005). If we consider human and non-human actors from this perspective, it affords us a symmetry around the meshes of interactions between them. We can use this network of actors to understand the complex relationships between a sound and its surrounding context. Even though a sound can be de-coupled from the stream of sounds in the environment we can still re-couple the sound's relationship with a surrounding through analysis–synthesis. From here we can see an almost infinite entanglement of interactions, which affords us to juxtapose different environments and study the differences and the effects between them.

To explore this notion further, we can think of this network of actors as the components of a structure. Each structure consists of interrelated points of contact between objects, and their constituent parts of a *shape*, a *site/location* and a *model*, as illustrated in Figure 1.2. This structural overview presents an inter-

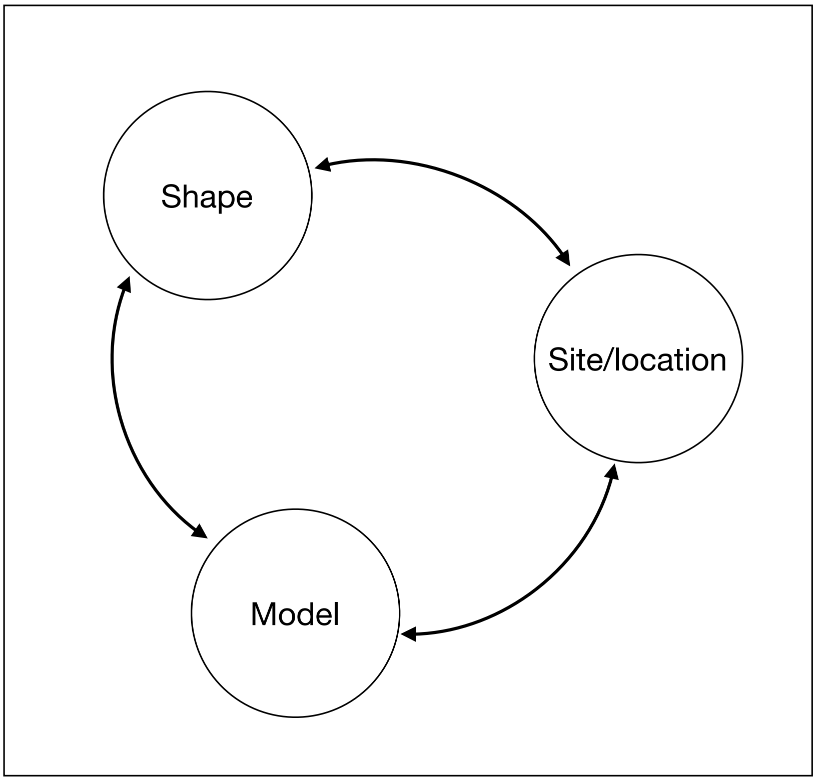


Figure 1.2: A schematic overview of the structure of an object, divided into the three parts of shape, site/location, and model. An object represents both the *intrinsic* and *extrinsic* features of a sound and carries with it a wider context.

dependent interaction that make up the cognition of sound objects, where each object contains a shape (this can be the overall envelope, the mass or the spectral features), a contextual relationship to a site or location, and a model which we use for analysis–synthesis. This model can be a map, a simulation, a theoretical construct, or a scale model, the important feature is that the model can function as a representation on which we base the analysis–synthesis. This structure affords a mapping of sound objects across the multiple axes between the *intrinsic* and the *extrinsic*. Within each structure we can define a series of axes that changes over time, that moves through space and which change the relationship among objects. This means that not only the positioning along an axis changes over time, as with a temporal envelope, but also that the parameter space itself is dynamic and malleable.

In the actor-network theory, an interaction is always situated between other actors and interactions. However, an *interaction* assumes prior existence of entities, and Karen Barad proposes an *intra-action* "to signify *the mutual constitutions of objects and agencies of observation within phenomena*" (Barad, 2007, p. 197) which questions the existence of the actors prior to their interaction. Rather than considering the determinism of preexisting actors, the variable relationships between objects discussed here posits that the objects are defined by their interactions with other objects.

The axes between the intrinsic and the extrinsic create the interactions between the sound object and the spatial reproduction that exist between what Raaijmakers described as "from the smallest sound to liquid form" (Raaijmakers, 2000, p. 81). Possible transmutations and morphogenesis of objects can range from "dull matter, hard resonant matter, flowing liquid, bubbling liquid or steam clouds" (Roads, 2015, p. 312). Objects have multiple significations and we shift our intentional focus to attend to different features of the object, as we also shift our focus between the object as we hear it spatially and how it is situated in three dimensions:

> The essential aim of spatialization, which is often confused with some strange myth of "spatial music", is to improve the definition of objects through their distribution in space, since it so happens that the ear distinguishes two simultaneous sounds better if one comes from the right and the other from the left. We are not dealing here with a luxury added on to our hearing but something to facilitate it. Before even mentioning space and sound architecture, we should talk about the identification of objects and their coexistence. Where they are is of little consequence; it is what this enables that is important: an incomparably clearer, richer, more subtle perception of their contents. In the same way, binocular vision gives the third dimension and by putting things in perspective with each other allows us to judge their properties and relationships better. (Schaeffer, 2017, p. 325)

From this we can see certain features. A piece of composed music (or installation) consists of a large network of parts, the sounds, and their spatial positions - the sound architecture. Each point in the network is a sound object which is informed by its perceived site of origin and its location in space, and a model which informs how this is represented. These sound objects are elements in a complex network of interactions, which extend from the intrinsic and extrinsic features of an individual

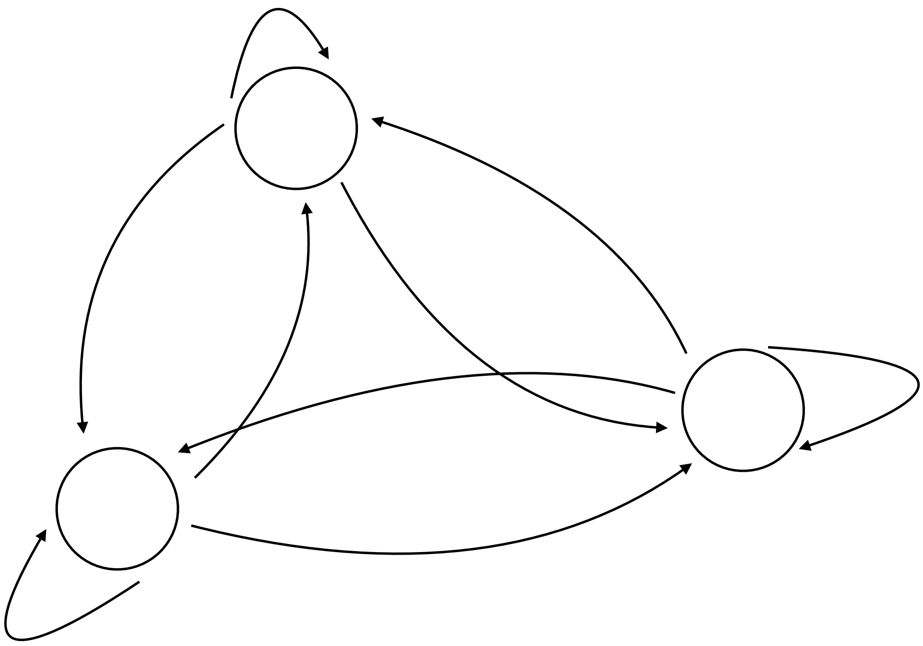


Figure 1.3: The structural relationships between objects, with the interactions that link them.

object to their structural features and their distribution in space (Figure 1.3). This is the topic of chapter 3.

1.3.4 Abstract/concrete

The notion of the abstract/concrete refers back to the experiments in the early days of *musique concrète*, as a fundamental change in how music was made. Musical ideas were notated symbolically and left to known instruments for realisation (Schaeffer, 2017). As seen in Table 1.1, this represented an *abstract* approach to musical creation where the musical concepts came first, then realised symbolically, before being realised through performance. For Schaffer and colleagues the "aim was to gather concrete sound wherever it came from, and to abstract the musical values it potentially contained" (Schaeffer, 2017, p. 7). In this way they sought to create a music that was "carved out of living sound" (Schaeffer, 2017, p. 7). Abstract/concrete is also referred to as the "two faces of perception" (Chion, 2009, p. 37), and the "two isotopes of the real" (Schaeffer, 2017, p. 8), it is a space for evaluation, where we can draw out multidimensional features of a music-making process based on the structures we investigate and the axes we employ.

The space between the abstract/concrete is a space for musical *experimentation*, which can embody all possible sounds. Despite the conflicts between concrete and electronic music in the 1950s and 1960s, Schaeffer still recognized Luciano Berio's *Omaggio a Joyce* (1958) and Karlheinz Stockhausen's *Gesang der Jünglinge* (1955-1956) as examples of electroacoustic experimental music (Schaeffer, 2017, p. 9).

In chapter 2 there will be a discussion on models of, and approaches to, listening, but for now, the concrete side of sound perception refers to the casual references in the sound, the references to the events which cause them and the "raw sound objects" (Chion, 2009, p. 21); while the abstract entails a selection of objects based on selective perception and signifies features of an object that is "stripped down to qualities which describe perception, or constitute a language, express a meaning" (Chion, 2009, p. 21).

Table 1.1: Diagram of the abstract and the concrete, where "abstract" music starts from a concept, realized through symbolic notation and at the last stage performed. While "concrete" music always starts by investigating the available sound material before abstracting musical values from it. (Schaeffer, 2012, p. 25)

ORDINARY MUSIC (*so-called abstract*)		NEW MUSIC (*so-called concrete*)	
Phase 1	Conception (mental)	Phase 3	Composition (material)
Phase 2	Expression (notated)	Phase 2	Drafts (experimentation)
Phase 3	Performance (instrumental)	Phase 1	Materials (making)
(*from the abstract to the concrete*)		(*from the concrete to the abstract*)	

Considering these "two faces of perception" we can look at some models for sound synthesis. In the three-part structure in Figure 1.2, the *model* refers to the approaches we use for analysis–synthesis, which determines how we explore a feature space. This we can briefly consider from a discussion on "standard" and "non-standard" approaches to synthesis (Döbereiner, 2011). Standard approaches seek to imitate, reproduce or vary an existing sound, as an example we have Jean-Claude Risset's experiments with bell and brass-synthesis, where additive synthesis was used to model the sound of existing musical instruments (Dodge & Jerse, 1997, pp. 103–107). Additive synthesis is a sound synthesis technique based on the summation of elementary waveforms to create complex spectra (Roads, 1996, pp. 134–156), and can be used for a wide range of approaches. Non-standard approaches, however, are not concerned with trying to imitate a previously known sound, but seek to find a musical expression through manipulation of material.

Among many possible approaches, Curtis Roads highlights four approaches of “nonstandard” synthesis methods. They are referred to as nonstandard in the sense that they are employed for the creation new electronic sounds. These are motivated by compositional aesthetics and by the creative imagination in searching for new sounds. Roads lists waveform segment techniques, graphic sound synthesis, noise modulation, and stochastic waveform synthesis (Roads, 1996, p. 319) among these nonstandard approaches. What these approaches have in common are noise components that stand in contrast to the “smooth and familiar textures generated by more standard approaches” (Roads, 1996, p. 345).

The differences between “standard” and “nonstandard” are used for differentiation, to define difference between the models employed for creation. Nonstandard approaches are based on compositional ideas surrounding sound production and ideas on musical organisation (Döbereiner, 2011) and more specifically nonstandard approaches seek “the composition of timbre, instead of with timbre” (Brün cited in (Döbereiner, 2011)), as for example in Morton Subotnick’s *Silver apples of the moon* (Subotnick, 1967). Standard approaches to synthesis carry a strong technological determinism, which is guided by the effectiveness of the model used for approximating a known sound.

> A sound-synthesis method is a formalism, and this formalism can be conceived of as a model. . . . The assumption is that we are on the one hand neutrally observing the facts, and on the other hand, actively producing a model. It is a confrontation between a real thing and an artificial reproduction, it is an opposition between reality and thought, and it essentially boils down to nothing more than the opposition of “nature” and “culture”. (Döbereiner, 2011, p. 33)

In the music for the 1956 film *Forbidden Planet*, Louis and Bebe Barron produced the first film soundtrack created entirely out of electronic sounds, which came out of an experimentation with magnetic tape and circuit design (Manning, 1993). On the sleeve notes for the release of the soundtrack the Barrons wrote: “In scoring Forbidden Planet – as in all of our work – we created individual cybernetics circuits for particular themes and leitmotifs, rather than using standard sound generators” (Holmes, 2012).

In standard approaches, there is a description of a sound in terms of an acoustic model and a machine is instructed to simulate the described model (Döbereiner, 2011). The end-goal is always the approximation of a given, existing sound, which is to be reproduced as accurately as possible, where “[T]he most straightforward way to obtain interesting sounds is to draw on nature in some way. Both spectral-modeling and physical-modeling synthesis techniques support incorporation of and/or modeling of natural sounds" (J. Smith, 1991, p. 9). This stands in stark contrast to the concrete methods, Curtis Roads as observed that "the concrète method is empirical; it starts from a harvest of *found sounds* (i.e., recorded sound portraits) and moves towards a musical macrostructure using available tools" (Roads, 2015, p. 80).

Within the field of spatial audio, Wilson and Harrison have made distinctions between specific and non-specific approaches to spatialisation, which are loosely understood as methods to synthesize spaces which do not necessarily refer a real-world context and to use space as an artistic element (S. Wilson & Harrison, 2010). This perspective signals an experimental approach into spatial composition and the creation of sound environments, rather than being content with the approximation and re-construction of a given environment.

1.3.5 Models of sound and space

Through an evaluation of spatial practice, Trochimczyk finds that different spatial designs can be related to different symbolic meanings inherent in these forms, and these are again related to meanings found in our culture (Trochimczyk, 2001). She describes a basic typology of musical-spatial design with a growth from zero to three dimensions, from point through, among others, triangle, hexagon to sphere. Two basic geometric designs stand out: the circle, which represents sound around the audience; and the net, which extends sound throughout the audience. The two geometric designs carries "a range of semiotic and symbolic associations: the circle (and, by extension, the sphere) and the net (or web, of a greater spatial complexity and multidimensionality)" (Trochimczyk, 2001, p. 42).

These evaluations of spatial designs align with the three-part structure of an object presented earlier in section 1.3.3, which divides an object into a shape, a site/location and a model. The *model* we use to analyse and synthesize the sound object *can* be aligned with the processes described as standard and non-standard approaches to synthesis, and are reflective of how the features of one object informs the features of the same or other objects. That is, spatial location, relationships to other objects in the same space, and the temporal evolution of the objects. A net consists "of a potentially unlimited number of interconnected nodes" (Trochimczyk, 2001, p. 50) and we can distribute sound objects in space by spectral similarities, envelope profiles, or textural dimensions.

Creating a framework for the evaluation of digital music instruments, Sile O'Modhrain identifies that *models* are "representations of systems or artifacts that provide a means of reflecting upon the design and behaviour of a system" and a "framework is often used in the context of human-computer interaction to describe a conceptual scaffold that can help to elucidate relationships between design approaches within a given design space" (O'Modhrain, 2011, pp. 29–30). The scaffold "is usually a conceptual scaffold whose purpose is to set out the relationship between different approaches to the development of human-computer interfaces, whether that be in terms of theoretical foundation, design or evaluation of design" (O'Modhrain, 2011, p. 32).

To approach a model of space and the sounds that are contained within it we can look to the *soundfield* as we understand it from ambisonics. This is a method for representing a full 3D soundfield through the decomposition of spherical harmonics regardless of the supposed "size" of the soundfield. We would normally consider a soundfield to be an exterior space, such as a forest, a beach or a city. Through ambisonics we can also create a full 3D soundfield from the inside of a teapot or a drum projected into the concert hall. These approaches will be explored in several of the case studies in chapter 4. This approach to sound and space gives us the opportunity to model the intrinsic and extrinsic relationships of objects and the complex fabrics of interactions between them.

1.4 A brief background review

The rich and varied histories that the present work is based upon, and embedded in, cannot adequately be reflected here. This section will provide a brief literature review, in supplement to the material discussed in the next chapters, to point the reader to key texts to consult as historical and contextual background material.

The use of natural acoustics of enclosed spaces as a direct and conscious support in the composition of music has a long history, with many examples which predate amplification and speaker technologies. For good discussions on "aural space" from prehistory to the present, as well as discussions on examples from the modern ear see *Spaces speak, are you listening?* (Blesser & Salter, 2009). Likewise, for analyses on architectural acoustics and the influence on music, see *Architectural Acoustics* (Long, 2005). The history of spatial audio technology is elegantly summarised in (Peters, 2011, pp. 1–10). Bodily reactions, acoustic influences and environmental alterations of sounds are the subjects of *Sonic Wonderland* (T. Cox, 2014); the culture of listening in America and the influences of architectural acoustics on *The Soundscape of Modernity* and the social significance on these developments is covered in (E. A. Thompson, 2004).

On the history of electronic and experimental music both Manning and Holmes have been good reference guides, which situate the developments of the various discursive practices of music (Holmes, 2012; Manning, 1993). Also *Audio Culture* is an indispensable collection of texts by key historical figures (C. Cox & Warner, 2004). For sound within the arts, especially through the work of The Futurists and John Cage, *Noise, Water, Meat* is an important contribution (Kahn, 2001). Other perspectives on sound art and discussions on artistic practice are found in (Licht, 2007) and (LaBelle, 2015). Mapping the sounds of the 20th and 21st Century's musical practices through ambient and electronic music, David Toop presents disparate practices as memoirs and journals through the books *Ocean of Sound* (Toop, 1995) and *Haunted Weather* (Toop, 2004). Two important books on the history and practices of electroacoustic music can be found in *The Language of Electroacoustic Music* (Emmerson, 1986) and *Living electronic music* (Emmerson, 2007). A book which has influenced this thesis thoroughly is Jonathan Sterne's *Audible Past* (Sterne, 2003). Peter Weibel's catalog from the ZKM exhibition *Sound art - Sound as a medium of art* (Weibel, 2019) presents a broad and comprehensive overview of sound art and associated practices which are relevant to this thesis.

Practices and processes of the synthesis, composition, and performance in *Computer Music* (Dodge & Jerse, 1997) presents a rich historical and technical resource. Written as a *Computer Music Tutorial* to comprehensively cover all aspects of music production on computers, (Roads, 1996) is still a landmark text. It covers not only the technical implementations of everything within computer music, but also the rich history behind the activities in the field. Likewise, covering the development of electronic musical instruments from the *Telharmonium* to the synths of the 1990s, *Electric Sound* tells the stories of the instruments and music through more than 150 interviews with practitioners in the field (Chadabe, 1997). Presenting a theory of composition through electronic music, *Composing Electronic Music* (Roads, 2015) offers a tour de force of historical anecdotes, facts and opinions for the exploration of electronic music. Likewise, *Musimathics* (Loy, 2006, 2007), *Composing Music With Computers* (Miranda, 2001), *Designing Sound* (Farnell, 2010) and *The Audio Programming Book* (Boulanger & Lazzarini, 2010) have been important educational tools throughout this project.

Lastly, Norbert Schnell's PhD thesis *Playing (with) Sound: Of the Animation of Digitized Sounds and their Reenactment by Playful Scenarios in the Design of Interactive Audio Applications* (Schnell, 2013) has had a great impact on this text in its final stages, both through its theoretical abstraction and complexity along with its focus on real-world applications.

1.5 An outline of the thesis

This introduction has presented the conceptual framework that this thesis is embedded in for discussions on sound and space. This has been a brief introduction to the different concepts that will be explored in depth throughout the next chapters. I have tried to place the practice of spatial, acousmatic sound, incorporating practices of electroacoustic music, computer music, soundscape, and sound art under the same term, as the focus is on sound-practices that solely uses the diffusion of sounds from loudspeakers, in a wider context than the dichotomy of production to presentation.

Specifically, the notions of the abstract and the concrete will be important in the coming chapters, where we will be discussing sound objects and the structures that surround them. In section 1.3.3 we looked at the concept of a network of objects and structures, derived from the actor-network theory, I will return to this frequently throughout the next chapters, in discussions on sound objects (chapter 2 and sound landscapes (chapter 3).

The case studies presented in this thesis are a mix of compositions, sound installations, acoustic modelling for virtual reality, and sound design for theatre and exhibition. These different projects are found in chapter 4 and are sites for the discussions of research through theoretical, aesthetic, and technical experiments. As such these should be viewed as informing the project as a whole. Spatialization and spatial audio systems can be plotted and visualized on paper but there is unfortunately few other ways of hearing the sonic outcome of these systems than by being in the space where they are presented. Along with this thesis, sound examples are made available for listening on an accompanying website.[9] In most cases a stereo decode of an ambisonic file will be made available, along with an encoded file for listening in a multichannel array where this is possible.

Overview

Chapter 2, discusses objects from the perspectives of the *intrinsic* and *extrinsic* features of a sound. Specifically, the discussions on the object lead to the sound object and the theories presented by Pierre Schaeffer. We also look into other models of perception, such as *events* and *streams*, along with discussions on perception of shapes and surfaces. Different models and approaches to listening are covered, and aims to provide an overview of different perspectives.

Chapter 3 examines the context of objects through the notion of the *structure*, the contextual features of sounds in space, through listening and *landscape*. The chapter discusses possible ideas of the sound landscape, through the place-bound and the site-specific. Different approaches to spatialisation are discussed, along with aspects of psychoacoustics.

Chapter 4 looks at Schaeffer's typomorphological framework as described in the *TOM* and proposes possible ways to use this system as a multidimensional model for spatial practice. Based on the two preceding chapters, chapter 4 discusses spatial processes and effects, by examining approaches

[9] Examples can be found at: http://u-l-v.org/index.php/research/.

to ambisonics and amplitude panning. This chapter also presents five different case studies that consist of an approach to acoustics simulation in a virtual reality-project and artistic projects involving sound design for theatre, compositions, sound installations, and sound spatialisation for an exhibition venue.

The conclusion in chapter 5 positions the work in the context of artistic and technological practice and presents the insights gained from the work, as well as proposing further avenues for where this practice can be pursued.

Chapter 2

Objects

> A precise description that is somewhat wrong is better than a description so vague that no one can tell if it's wrong.
>
> *Joseph Goguen*[1]

If we follow the concerns of Edgard Varèse cited at the opening section of this thesis, his use of topological metaphors described a vision for the future of music-making. This vision prefigured the developments and practices of spatialisation and spatial audio in acousmatic music and sound art practices, his vision of colliding masses, transmutations, zones, and other metaphors have become pervasive in the contemporary understandings of acousmatic music. How is this reflected in acousmatic music? Acousmatic music is a music where all pitches, timbres, and complexes of morphologies are possible (Landy, 2007), and surely it is also a music where *all complexes of space* are possible. Acousmatic music affords the listener a wide register directly related to human emotions, perceptions, and memories through the uses of sound material related to nature, technological, and human relationships; it affords a decoupling from the human body and its limbs as the visual, gestural references to sound-making are removed. The embodied gestures we normally associate with music are also parts of our identification of music, but the gestures associated with acousmatic music transcend the human body and opens up a deep signification of space, place, identification, and perceptions. This chapter seeks to discuss these features by examining acousmatic music's *object*.

Section 1.3.3 of the previous chapter presented a three-part representation of

[1] Goguen, 1999.

a sound object's features as consisting of a shape, a site/location, and a model. These objects are not isolated, discrete units, but rather tangible objects that are part of the structure which we use to perceive a context. When discussing objects it would seem natural to talk about the physical, sound-producing object (or source), and not about the resulting sound made by the source which we hear. We use the object to talk about features of a sound, and how these are perceived as being representations of dynamic, spatial, and timbral features as *shapes*. The term object is often used as a wildcard, a variable which we often can load up with whatever substitution we might have for "thing". We experience our surrounding world grouped into sources, or *auditory objects* (Van Valkenburg & Kubovy, 2004), indicating that the perception of our surroundings consists of three-dimensional wholes, as objects, and not necessarily as events or streams. The focus in this chapter is on objects and the associated perspectives. First we will look at issues associated with events and streams before moving on to objects and that of sound objects. Before that, we will look at recent discourses on sound's privilege and the "tyrannical, visual hegemony".

2.1 The "tyranny" of the visual

The sound object has previously been described as a fragment of sound, *suitable for study in itself*, and this needs a clarification. In the sonic arts there is a tendency towards a theorization of sound as something with a unique position. With this criticism comes an implicit attitude that experiencing sound and the sonic is in a more important sense modality than experiencing something in the visual domain. Sight has been a privileged sense in European philosophical discourse and this has led several writers and scholars to talk about the "visual hegemony" and "the tyranny of the visual".

Jonathan Sterne has charted many of these ideas along a 2000-year old Christian theology of listening. This is illustrated in his "audio-visual litany", which idealizes hearing as a pure interiority (Sterne, 2003, pp. 10–19). The audio-visual litany rests on a series of binary oppositions, which we discussed briefly in section 1.3.3, where a limited perspective on our thinking leads to a categorical, either/or, problematic. Sterne's critique traces this theology of listening, to hearing (the voice of God) as the (always) preferred path to bring us closer to God. Sterne's audio-visual litany covers such oppositions as "hearing is spherical, vision is directional; hearing immerses its subject, vision offers a perspective; hearing places us inside an event, seeing gives us a perspective on the event; hearing is about affect, vision is about intellect" (Sterne, 2003, p. 15). From this excerpt of the litany we see the way sound can be described as affective, interior, temporal, closer to spirit and (em)bodily; whereas vision is rational, spatial, distanced and intellectual (Kim-Cohen, 2016a). This primacy of the auditory is described as "sound-in-itself", a term which seeks to define the autonomous and detached existence of sound, as was formulated by John Cage in his credo "let sounds be themselves" (Cage, 2012).

Recently, several writers have presented a type of *sonic realism*, to give sound (and hearing) a primacy and a primary status as an experience among sense modalities. Most of all these writers seek to place sound within a phenomenal realm where it exists "as an anonymous *flux* that precedes and exceeds human contributions to it" (C. Cox, 2017, p. 102). A fundamental understanding of *why* sound must have some primary presence in the world is that we, as humans, lack earlids so "we are forever and inescapably bathed in sound" (C. Cox, 2017, p. 101). This, it would seem, is why sound holds a special status, because the sonic flux "elegantly and forcefully models and manifests the myriad fluxes that constitute the natural world" (p. 103) and "the sonic flux is not just one flow among many; it deserves special status insofar as it so elegantly and forcefully models and manifests the myriad fluxes that constitute the natural world" (C. Cox, 2013, np). Sound is unequivocally given some special status because the ears are holes in the head that lets sound in, but light stops at the eyes: "Inside the head, then, it is noisy but dark" (Ingold, 2005, p. 98). That we do not have earlids should in no way mean that we cannot learn to separate and differentiate sounds, nor that we must give up ourselves to the constant sound immersion and the sonic flux precedes and exceeds human contributions.

Through what he has dubbed "visuocentrism", Casey O'Callaghan states that philosophical and scientific activities around "the perception of colors, objects, motion and causation" (O'Callaghan, 2007, p. 2) has shaped our perspective on the nature of perception and that vision has been the preferred method of theorizing about perception. It then becomes a prizing of sound for its un-graspable features, that "one cannot grasp or sit upon a sound" (O'Callaghan, 2007, p. 5) and that "sounds do not seem in auditory experience like ordinary objects, for we do not hear them to qualify ordinary objects in the way that visible attributes do. The world of sounds confounds visuocentric thought" (O'Callaghan, 2007, p. 8). By trying to answer the question: "what is a sound?", the ensuing discussion is a plunge into an esoteric discussion of "events as I wish to understand them are immanent or concrete individuals located in space and time" (O'Callaghan, 2007, p. 58). This proposes a blurry ontology where sounds have a being and are individuals, said to exist temporally but wholly separate from the world around them as they are not tangible and tactile like physical objects. Colours appear to us to be stuck onto objects, while sounds do not act in this way (Steintrager, 2018); where, philosophers fascinated by sound would "follow it where it leads, encountering a strange world in which bodies are dissolved into flows, objects are the residues of events, and effects are unmoored from their causes to float independently as virtual powers and capacities" (C. Cox, 2017, p. 107).

Through this emphasis on an ontological blurriness proposed about sound and sonic experience, it would seem that the desire is to create an environment where sounds can no longer be the object of critical judgement or engagement but just *be*. By privileging sonic experience the intention is to "relieve sound from its submission to ocular primacy . . . through being opposed to the distance of the visual, is consigned to an isolating interiority that actively disables critical engagement" (Schrimshaw, 2017, p. 10). The sonic is then given primacy on the basis of mythic and theological perspectives.

From researchers of acousmatic music as well, there is a move towards this theorizing of sound. François Delalande has made the statement that "pure music" is experienced where no source or cause is identifiable (Landy, 2007). It is, of course, foolish to assume that any one person's experience and perception of the world could possibly be complete, yet we must aim, as Chion has stated, to "disengage sound thinking and its technical and aesthetic applications from its naturalistic rut" (Chion, 1994, p. 94). Sound is always contextual, Cox and O'Callaghan's perspectives discussed previously above disregard the fact that perception and cognition does not rest on one single sense alone. Vision and hearing functions together for us to make sense of the world around us. We can hear things we cannot see but we also see things we cannot hear, and it is in the interplay between the senses that we make sense of the world. We should demand that we can talk about sound, media, and culture without recourse to mysticism.

Then, to return to the opening of this section. What was then, for Schaeffer, the "true" experience of studying a sound for itself? A sound recorded to a closed groove disc or tape could be listened to without having any visual clues apart from seeing the disc spin or the tape machine play the tape. This afforded listening to a sound without having to consider the source and in this way treat the sound as an object of study. The sound remains a sound, but it is separated from the physical signal and listened to *subjectively*. It is only temporarily removed from context, representation, and significations, and analysed, classified, and studied before being reinserted it into the context of a musical work. This approach is not about the "sound in itself". It is an analytic approach to understand the features of a sound which dismisses the perspectives presented above. These issues will be made clearer in section 2.6.

2.1.1 Ontological differentiations

When presenting an ontology of sounds "beyond representation and signification", Christoph Cox wishes to reclaim some primary sensual qualities for sounds and distill them away from the "hegemony of the visual" (C. Cox, 2011, p. 157), which is through an insistence that "[M]usic, participating in the symbolic and imaginary order, is an assemblage that captures sound and turns it into form" (Kane, 2015, p. 9). This is what Kane refers to as onto-aesthetics, in that the cultural (and social) conditions of perception and artistic practice are of less importance than the reality of sound itself. Bound up into onto-aesthetics is the *exemplification* of an ontology, and "this exemplification of a sonic ontology is most readily found in a predisposition towards immersive aesthetics, focused on the presentation of sound itself, beyond its implication in the representation, as a signifying or pre-symbolic affective flux" (Schrimshaw, 2017, p. 110). This exemplification is, according to Kane, based on a category mistake, where embodiment is confused with exemplification. Exemplification can come in degrees, but objects embody their ontology and embodiment "is all or nothing" (Kane, 2015, p. 12). Exemplification is referential, whereas embodiment is a condition, and no object exemplifies its own ontology better than another object or being (Kane, 2015).

Before continuing the discussion on objects, I want to discuss some of the work surrounding perceptual modalities of events and auditory events, streams, and auditory scene analysis.

2.2 Events and streams

Within the complex listening environment surrounding us in the real world, sounds arrive at our ears from all angles. Despite the sound world being a continuous *stream* of sound, human listeners are able to identify and segregate sources into discrete units of perception. In spatial composition we draw directly on studies into the psychoacoustics of sound localisation and the understanding of everyday sound events as we represent the objects associated with them, or the "*piece-wise* cumulative perceptual images" (Godøy, 2019, p. 169) associated with the identification and segregation of sounds to objects. When listening to music we extract information from a continuous sensory stream, which indicates that there is an ongoing mental simulation where we process sensory information as a re-enactment of what we perceive (M. Wilson & Knoblich, 2005). In studies of everyday significance of sounds there are several perspectives associated with an event-centered listening.

Jens Blauert makes a distinction between a *sound*, such as a sound source, a sound signal and a sound wave, and of something that can be denoted with the adjective *auditory*, specifically an *auditory event*. The difference being that a sound source, signal, or wave describes the physical aspect of the phenomena of hearing and that an auditory event is not necessarily caused or determined by sound events. Rather it describes the subjective perception of a sound situation. For example, ringing in the ears from tinnitus occurs without there being an external sound event. An underlying assumption in human sound perception is that we hear something in a medium where we experience a wave or vibratory phenomena between a frequency of 16Hz and 20kHz. This does not mean that whatever vibrations exist in that medium is what is actually heard (Blauert, 1997). "More generally, the fact that a sound event does not necessarily produce an auditory event, and that not every auditory event is connected to a sound event, must exclude the interpretation that one leads to the other in a casual sense" (Blauert, 1997, p. 3). Importantly, "sound events and auditory events are distinct in terms of time, space and other attributes, they occur only at particular times, at particular places and with particular attributes" (Blauert, 1997, p. 3). So, *spatial hearing* needs the relationships between the locations of auditory events and that of sound events.

The explanation of the *sound event* is that the perception of one or more sounds in our surroundings is due to a change is the sonic environment and that an object produces acoustic disturbances that travel to us. Therefore our perception of sound is the perception of this particular sound-producing event. The event-based ontologies does not and cannot separate a sound from its source or cause, as the "sound is the result of an abrupt change in the environment" (Van Valkenburg & Kubovy, 2004, p. 118). Rosenblum, for example, proposes to drop the term object in favor of the event through the emphasis that "we hear events, not sounds" (Rosenblum, 2004, p. 241) and levies parts of his events argument on Heraclitus and the fact that we cannot ever step into the same river twice, because when the water has passed downstream it is a new river (Rosenblum, 2004). This is of course true, but through our own sensory system we cannot determine if this is a new river or what could be new about it. Rather, the river is, for us, the same as it has been and the *object* that rests in our perception and in our memory is not concerned with the passage of water downstream, rather we are concerned with the spot we usually go for a swim, go fishing, or sit to pass the time. The concept of "river" makes this into a perceived object that signifies a spot we have a relation to.

An event suggests it is not just something temporal, but also something localised in time, whereas an auditory object is a carrier of properties, a thing with a set of specific features for all the perceptual values that make up "the thing" (Schnupp et al., 2012). As such the sound event could be understood along Lefevbreian lines, where sound is conceptualised as a burst of energy and this energy changes a space and our perception of it (Lefebvre, 1991). Our perception of a concert hall or gallery can perceptually be completely transformed by a sound event but it is still the same concert hall or gallery.

The EVENT schema is presented as a fundamental context for electroacoustic music (Kendall, 2008), this rests on the notion that the way we organise and characterise the activities around us is as events. At one level, an "event" is an immediate and seemingly automatic understanding of sensory experience, and at another level, an “event” is a metaphorical construct that we use to describe a diversity of sensory experience. Yet, where a “happening” of sorts can be ascribed to events, which is more related towards a mental than a physical experience, Kendall is preoccupied with descriptions of how the event is something which unfolds continuously *through* time.

This perspective has a narrow view about sound materials, sources, and the perception of the origins of sounds. Schaeffer took great pains to point out that the *actual* source is no longer of any concern as we should focus on the *heard* sound and the features and individual properties we can gain from listening. However, the *apparent* source of a sound becomes a concern when several sounds are presented together and are reinserted into the context of a composition or installation.

The event perspective is explored further, in slightly different contexts, through the soundscape. Schafer makes distinctions between cognition of objects and of events, where a sound analysed in the lab is regarded as an object but the same sound experienced in its original context, on *site*, is an event (Schafer, 1994). This perspective of an event is context-dependent and context-driven, and one of Schafer’s examples is hearing church bells in a village. Being present and listening to the sound of the bells and how they fill the space is considered an event, or listening to an event. If the same sound was recorded and listened to in the lab, then it is an object, presumably by virtue of its *schizophonic* status, of splitting the sound from its original context (Schafer, 1994). However, when listening to the sound event in the village, the *experience* of hearing the bell and the subsequent acoustic behaviour in the village space this acts an *object*, as we separate the information related to the “thing” (the bell) and the rest of the surroundings. When the sound event is over, the sensory memory of the event is left as an object in our consciousness, as an analysis of information from the sensory world. This is discussed further in the following section 2.3.

The artistic play of events in electroacoustic music can be experienced as a mirror on the events of everyday life (Kendall, 2008). The event is conceptualized as a method of understanding groups of sounds (and compositions) and not just single sounds. This is similar to Bregman’s insistence on perceiving footsteps as a stream that is interpreted as one sound and not as the individual sounds which constitute the footsteps. If we hear something, we hear the sound that *something* makes. This something can be an object close to us or something in our environment, which leads to the statement that “sounds are public objects of auditory perception. By ‘object’ I mean only that which is perceived - that which is available for attention, thought and demonstrative feedback” (O’Callaghan, 2007, p. 13). When hearing a sound of a passing car or a breaking plate, you not only hear the sound but also something about where the sound occurs.

For Gaver, the event is tied to everyday listening and is influenced by the ecological approach to visual perception (Gibson, 1986). He poses two fundamental questions: "What do we hear?" and "How do we hear it?" (Gaver, 1993b, p. 5) and this guides ecological acoustics and its descriptions of the acoustic properties of sounds and what information they convey about the events that cause them (Gaver, 1993a). Ecological acoustics is discussed further in section 2.7.6.

One auditory event can contain multiple sound objects, as we attend to the individual objects the auditory event is purely ground - it provides us with a temporal background context needed to hear and understand the foreground object. But if we listen to the entirety of the event, we are not hearing the figure at all. The event suggests that the sound has an occupancy and a localisation in time, but only within the perceptual instant of the human. There is no causal relationship to the event and long lasting sounds, or indeed sounds that are masked or hidden by other external factors. Despite these criticisms, we need to consider the sound event if we want to discuss a temporal experience of objects in spatial relationships to something else than the context of their own existence. In chapter 3 we will return to the event, and the needs to conceptually understand the interplay between the supposed "source" and the resulting sound we hear.

In the influential book on *Auditory Scene Analysis*, a *stream* is defined as a "perceptual unit that represents a single happening" (Bregman, 1994, p. 10), exemplified by the sound of footsteps, which consist of several independent sounds but which together form a single event. The justification for using the term *stream* and not *sound* is that a stream can contain a manifold of mental representations of acoustic events. With the word *stream* "we are free to load it up with whatever theoretical properties seem appropriate" (Bregman, 1994, p. 10). The concept of the stream serves the purpose of being able to cluster related qualities, so that we can *group* sounds of the same source, *segregate* those of other sources and understand the "auditory stream".

From a stream of continuous and distinct experience, the auditory stream is the auditory analogy of visual grouping (Schnupp et al., 2012) based on the Gestalt theories of perceptual grouping. The grouping principles employed by Bregman for the auditory stream involves the grouping of percepts of a single source at a given time, and the principles of grouping are interpreted as rules for scene analysis. Among these principles are *proximity*, where objects appear grouped if close together; *similarity*, where objects appear grouped if they have similar properties; *closure*, where perception can fill in the missing parts of a whole; and the principle of *figure and ground*, which describes how we can shift attention between two streams, between a foreground and background.

Bregman uses the term stream to signify an auditory experience, akin to seeing a visual object and states that "properties have to belong to something" (Bregman, 1994, p. 11), particularly when there is more than one such "something". However, hearing "a single burst of white noise against a background of silence wouldn't be a stream, according to current terminology" (Schnupp et al., 2012, p. 265). Auditory scene analysis is in and of itself not one unified and well-defined discipline or methodology but rather a "collection of questions that have to do with hearing in the presence of multiple sound sources" (Schnupp et al., 2012, p. 223).

But what is a stream? In one regard, we follow convention by naming any cluster of sound(s) a stream. We can focus our attention on individual groups of instruments, but we do not hear individual instruments when hearing a symphony orchestra perform, rather we hear the combined orchestra. The *way* it is orchestrated creates textures we cannot find in individual instruments but rather in our perception of the music as a single, fused *stream*. Through the grouping principles, Bregman seeks to explain how we segregate sounds from an auditory scene into separate auditory streams; for example, how we can focus on the sound of someone speaking in a room full of other people or how we can separate the sound of a person from the noise of a city street. What differentiates an auditory event from an auditory object? This is discussed in the next section.

2.3 Defining objects

Defining *objects* is not straightforward. In the Merriam-Webster dictionary an object is defined as "something material that may be perceived by the senses" and as "something mental or physical toward which thought, feeling, or action is directed".[2] Although in common use, and a source of misunderstanding, the term object is the description of a *physical* thing, like a rock, book or building. However the subsequent definition signals the idea that objects correspond to sensory experiences, which need not be visual. The term "object" can also be used for something *immaterial* like "the object of your affection", "the object of study" or "the object of the game is...". Objects can influence each other and one object has an effect on how you perceive another, but are united by one premise: "Objects cannot be known independently of sensory experience" (Griffiths & Warren, 2004, p. 887). When we delve into questions related to objects, we find that these "are often ontological questions, questions about what exists and what form existing things take" (Bogost, 2008, p. 31). A resistance against *auditory objecthood*, of sound as object, is, above all else, because objects are seen as being physical entities and "the bias against the idea of auditory objecthood is embedded in folk ontology. Language itself may lead us to believe that objects are visible by definition" (Kubovy & Van Valkenburg, 2001, p. 98).

An object is a container and a conceptual structure. From the discussion in

[2] https://www.merriam-webster.com/dictionary/object

chapter 1, I set out to expand the notion of a sound object to consist of three parts: a shape, a site/location, and a model. Each object contains multiple features that signify relationships to a model of analysis/synthesis, to a contextual site, and to a sonic shape. From the phenomenological philosopher Edmund Husserl we see that there is a difference between an *entity* and an *object*. An entity is something that exists externally, such as a room or a table. While an object is something that is apprehended by a subject, when it is "cognized" (A. D. Smith, 2003). An object is always an object for a subject. It is precisely its objects that enables acousmatic music to become a coherent musical expression: any sound can be used in a musical context, but it is the analysis of discrete objects through listening, analysis, and transformations, which provides the context and existence of the music.

To equate an object (sound) with a source is arbitrary, as there are nonlinearities between the *source* of the sound, the *physical signal* and the subjectively *apprehended* sound. Schaeffer referred to this as *anamorphosis*, or warping, and urged the musical experimenter to explore the possible correlations between the physical signal and the listener's subjective perceptions. Kittler has noted that "a reproduction authenticated by the object itself is one of physical precision. It refers to the bodily real, which of necessity escapes all symbolic grids" (Kittler, 1999, p. 12). This is discussed further in the following section 2.6.4.

In vision, an object is a surface that stands out from the rest of the surface layout (Gibson, 1982). This indicates that although there might be multiple objects in the environment, only one of these stand out and becomes a figure (Van Valkenburg & Kubovy, 2004). This *figure* is defined as a perceptual object that is susceptible to figure-ground segregation (Van Valkenburg & Kubovy, 2004, p. 124) and it is through this segregation that we discriminate objects from the foreground to the background. We experience the world around us, in sight and in hearing, as separable, distinguishable, three-dimensional wholes - *objects* (Van Valkenburg & Kubovy, 2004). By focusing on object cognition, and not on the events or streams, we are insisting on and discussing a perceived materiality of sound. In the separation of figure and ground, we focus on either the auditory foreground or the auditory background, or "the separation of the foreground and background" (Van Valkenburg & Kubovy, 2004, p. 122), this is also discussed in section 3.6.1 in the context of spatial audio. In reference to music, we can talk about objects as "discontinuous and piece-wise cumulative perceptual images" (Godøy, 2019, p. 169) in terms of sound objects as shape images.

In neuroscience, the concept of "auditory objects" is important for examining how humans experience sound. In reviewing principles regarding auditory objects, Griffith and Warren define four general principles of object analysis:

> First, object analysis involves the analysis of information that corres-

> ponds to things in the sensory world. Second, object analysis involves the separation of information related to the object and information related to the rest of the sensory world. Third, object analysis involves the abstraction of sensory information so that information about an object can be generalized between particular sensory experiences in any one sensory domain (in the visual domain, for example, we recognize the same face regardless of angle or illumination). Fourth, object analysis involves generalization between senses, such as the face and the voice of the same individual. (Griffiths & Warren, 2004, p. 887)

The first two conditions for defining auditory objects align with the scope in this research project, we listen to the sound not what produced the sound, that is, the event. However, the last two conditions define a framework that I find too strict to use objecthood effectively in the context of research into acousmatic music.

Examining differences between notions of spatial and temporal perception, between objects and events, Bernard Mayo identifies that we operate within a spatio-temporal framework where material bodies - objects - are the main occupants (Mayo, 1961). Objects exist in space and time, and events happen to these objects and sometimes change them. As we discussed earlier, objects do this to each other as well. If you pluck a guitar string that vibrates at 440Hz in two distinctly different rooms, it will still be the same 440Hz tone, the sound event is the same but the conditions around our perception changes and hence the perceived sound is different.

Mayo's complementarity between objects and events does not relate to sound, rather it considers that events are "episodes in the history of particular objects" (Mayo, 1961, p. 340), but the connection between events and objects is still of interest to our discussion. This is illustrated in Table 2.1.

In a criticism of the adoption of the term *object* in Object-Oriented Ontology (OOO) (Harman, 2011b), Ian Bogost rightly observes that the uses of the term outside of computing can cause some confusion, in particular for the specific meanings of object and object-orientation in programming languages. As a definition for *object* in programming, Bjarne Stroustrup has defined that "an object is a region in memory with a type that specifies what kind of information can be placed in it. A named object is called a variable" (Stroustrup, 2014, p. 60). This will be relevant for the discussions on the typomorphology in chapter 4. For Bogost, an object implies a physicality and a materiality of what is under discussion and that there are no limits to what the object can encompass, from symbols, corporeal, and incorporeal entities, to ideas (Bogost, 2012). Rather Bogost would have us use the term *unit*, because of its "ambivalent turn, indifferent to the nature of what it names" (Bogost, 2012, p. 25) and *unit operations* are used to denote se-

Table 2.1: An overview of Mayo's complementary analysis of differences between objects and events (Mayo, 1961, p. 342).

	Object	Event
1(a)	An object has a limited extension and an unlimited duration.	An event has a limited duration and an unlimited extension.
(b)	It cannot occupy the whole of space, but it could occupy the whole of time.	It cannot occupy the whole of time, but it could occupy the whole of space.
(c)	There must be room in space for many objects, which may or may not overlap temporally.	There must be room in time for many events, which may or may not overlap spatially.
2	An object can, at different times, occupy the same space (rest, endurance) or different spaces, normally of the same size (locomotion, endurance).	An event can, at different places, occupy the same time (occurrence, extension) or different times, normally of the same length ("propagation", extension).
3	An object cannot be at different places at the same time, unless its spatial size is greater than the interval between the places.	An event cannot be at different times at the same place, unless its temporal size is greater than the interval between the times.

miotic systems of interaction. However appealing this perspective is, the unit does little to address questions around perceptual organisation. With a slightly different angle Sabina Leonelli argues that interactions in the world produce objects, which are processed as data, in turn ordered as models to represent the world, and in the end is interpreted as knowledge before informing further interactions (Leonelli, 2019).

2.3.1 Things and tools

There is a tension in the intentional experience of objects and their qualities, because objects run deeper than the interactions we have with them. Husserl's method of bracketing out the outside world to direct our intentional focus on the phenomena that appear to us, entails that all "perception, judgment, love, and hate is perception, judgment, love, or hate *of some object*" (Harman, 2011b, p. 173). This indicates that intentionality signals a production of the objects we perceive and not as a passive observation or response. Indeed, objects connect further and

rather than viewing objects as too shallow to be the truth, we can treat

> them as too falsely deep to be the truth. This happens whenever a philosophy tells us that an object is nothing more than how it appears to the observer; or an arbitrary bundling of immediately perceived qualities; or when it tells us that there are only "events", not underlying substances; or that objects are real only insofar as they perceive or affect other things. (Harman, 2011b, p. 172)

Husserl's perspective met a radical reworking in his student Martin Heidegger, with the realization that our conscious awareness of the world around us only makes up a small portion of our lives, but mostly "objects withdraw into a shadowy subterranean realm that supports our conscious activity while seldom erupting into view" (Harman, 2011a, p. 37). Heidegger proposed that our understanding of the world is an engagement with things and specifically with the relations between things. This is part of his famous "tool analysis", or *tool-being*.

"Equipment" is a specific class of "tools" defined as *readiness-to-hand*, and its existence is dependent on the interactions with other things. The totality of an object's existence is owed to the other things it references and how the same object is referenced by these things:

> A totality of equipment is constituted by various ways of the "in-order-to", such as serviceability, conduciveness, usability, manipulability. . . . Equipment – in accordance with its equipmentality – always is in terms of its belonging to other equipment: ink-stand, pen, ink, paper, blotting pad, table, lamp, furniture, windows, doors, room. (Heidegger, 1962, p. 97)

When we look at a river, we can never grasp the entirety of its being or reality and most of its existence slips away into a "veiled underworld", apart from the interactions we can perceive. Heidegger's universe is made up of "tool" and "broken tool", and he "recognizes these two basic modes of being, and *only* these two: entities withdraw into a silent underground while also exposing themselves to presence" (Harman, 2011a, p. 39). We expect things to work and it is only when they stop working that we notice them, the functioning of the objects that surround us pass unnoticed. We then become aware of the relational entanglements of the things surrounding us (Olsen, 2010).

Luciano Berio claimed that whatever we intend to hear as music, is music (Rocchesso & Fontana, 2003), and Schaeffer refers to the accidental phenomenological practice that led to the understanding of the sound object (Schaeffer, 2017), where we temporarily suspend our knowledge of the surrounding world and make subjective judgements about what we hear. This way, the object is not an end in and of itself, but a method of analysis of a (potentially) ontologically complex sound. This is grounded in

> A particular object and the various modes through which I relate to this object: perception, memory, desire, imagination, and so forth. In what way is the object *immanent* in these? Because it constitutes an *intentional unit*, involving *acts of synthesis*. These many experiences are directed toward it, and they organize themselves around it so well that I cannot account for the structure of my consciousness except by perpetually recognizing it as "consciousness of something". To this extent the object is contained in it. (Schaeffer, 2017, p. 207)

What is this *intentionality* that Schaeffer was concerned with? This intentionality is related to the objects that lie "before the mind", these objects of our perception are objects of something. Our perception is always directed *towards* something, to some object of experience. Attention and perception is directed towards an object not only for its intrinsic and extrinsic features, but also for its interaction with other objects. Our intentional focus through listening allow us to change the focus of our attention to specific features of a sound and the relationships of those sounds to our surroundings:

> If I hear the squeaking of a door when I am expecting a visitor, I would understand this sound as indicating the arrival of my visitor, and probably not be so much interested in the sonic features of the squeak. But being sensitive to such squeaking noises, I could direct my attention towards the squeak, making a note that I should lubricate the hinges to get rid of the squeak, or I could even be intrigued by the brass instrument like timbral features of the squeak, its pitch and dynamical envelope, in short, start to focus on the squeak as a sonic object. (Godøy, 2018, p. 763)

The shifting intentional focus described here will be discussed further in section 2.7.1, under models of listening, but for now this situation is a good example of the different active components in the listening approach described by Schaeffer, of listening as a basic capacity, hearing as physiology, attention to certain sounds, and the transition between basic listening to grasping the significations of the sounds we hear.

2.4 Shapes and surfaces

Shapes are not just abstract symbols, rather shapes are the features that make up our cognition of objects (Godøy, 2019). All objects have properties such as mass, location, and velocity, among others. Surfaces are interfaces to understand the links between a medium and substance: a surface is an object that stands out from the background. A shape is understood as the potential interaction between an organism and an environment (Gibson, 1986), and we can then define an object as something that is a separation of the auditory foreground and background, as a unit of attention (Kubovy & Van Valkenburg, 2001; Van Valkenburg & Kubovy, 2004). We can relate this definition to Schaeffer's concept of the sound object, as being basic units of perception as small fragments of sound.

A visual and material object has edges that can be discerned by eye and the differences among shapes are perceived by changes in an array of light reflected off the objects. For example, we can say that the illustration of figure and ground in the Rubin vase/face illustration comprises shapes, where we visually can discriminate the edges separating the two objects.

The shape, or the reduction *to* a shape, has an important precedent in the visual arts. The French painter Paul Cézanne laid the foundations for a radical reshaping of the arts in the 20th century, based on his Post-Impressionist paintings. He developed a style and technique where he stripped down a landscape to its constituent geometric elements of the sphere, the cone, and the cylinder as basic building blocks.[3] Cézanne's work had a deep influence on both Pablo Picasso and Henri Matisse and their subsequent developments of painting and sculpture in the 20th Century (Becker, 2008).

Shape perception in sound is ultimately a spatial perception that necessitates an attention to contexts and relationships "as the move into phenomenology proper is made, it is with the *spatiality* of sound that description may begin" (Ihde, 1976, p. 58). One of Don Ihde's principles in listening relates to how "sound can be perceived spatially in the guise of hearing sonic shapes, surfaces and interiors" (Born, 2013, p. 9). If we go back to Mayo's complementarity of objects and events, he claims that objects cannot occupy the whole of space and that makes them into discrete, perceptible shapes located *spatially* around us and available to our perception.

D'Arcy Thompson studied the physical and mathematical properties of morphologies in nature and he showed that all the shapes we encounter in nature have traces of motion in them. Shapes are articulated through motion and materiality:

> The form, then, of any portion of matter, whether it be living or dead, and the changes of form which are apparent in its movements and in its growth, may in all cases alike be described as due to action of force. In short, the form of an object is a "diagram of forces", in this sense, at least, that from it we can judge or deduce the forces that are acting or acted upon it: in this strict and particular sense, it is a diagram – in the case of a solid, of the forces which have been impressed upon it when its conformation was produced, together with

[3] https://www.britannica.com/biography/Paul-Cezanne/Development-of-his-mature-style

> those which enable it to retain conformation; in the case of a liquid (or of a gas) of the forces which are for the moment acting on it to restrain or balance its own inherent mobility. (D. W. Thompson, 1992, p. 16)

From this we can ascertain that the shapes and forms of the objects we encounter bear some traces of the events that caused them and this can help us infer some form of context onto the objects we perceive. A sound has a profile shape that we use to perceive and understand it. Sounds form individual shapes both temporally and spatially as they radiate in different directions and evolve over time. Sounds are determined by their radiation patterns, and different sound sources radiate differently (see section 3.4 for more). The shape of a sound is determined not only by the radiating pattern from the source but also by the sound's interaction with the surrounding architecture. With the interaction of sound to its surroundings, we should also consider the notion of surfaces. Surfaces influence the acoustic perception of a room or an outdoor space, and are distinguished from an physical object. When a sound radiates from a source, "the simplest obstruction it can meet on its way is a single surface, which will give a reflection" (Halmrast, 2015, p. 256). The direct sound combined with these reflections makes up our perception of the sound image heard within a space. Indeed, "Gibson's theories of direct perception account for the reciprocity of a perceiver and their environment in a unified perceptual system" (Worrall, 1998, p. 98).

Gibson recognised nine different surface properties (Gibson, 1982) in visual perception, where he identifies that a surface is not discrete like an object but "is nested within superordinate surfaces" (p. 161), a surface does not have a location as an object does. Moreover, surfaces have perceivable properties that are "hard to soft, luminous or reflecting, illuminated to shaded, high to low reflectance, uniform to speckled reflectance, smooth to rough texture, opacity to transparency, dull to shiny, and hot to cold" (Gibson, 1982, p. 152). Gibson defined surfaces to make up a third of a tripartite relationship, and this relationship also includes *medium* and *substances* alongside surfaces. Ingold argues that surfaces exist at the interface between the medium and substances, this is where energy is absorbed or reflected, vibrations are passed and where vaporization and diffusion occur (Ingold, 2005). This reflection on the properties of surfaces is similar to what Goguen remarked of a coffee cup being an interface between the coffee and the coffee drinker. A surface is an interface in acoustical perception between reflections in space and the ear, between a musician and an instrument, and between the *visual* perception and what we hear.

In the relationship of shapes and the awareness of spatiality in the *typomorphological* system (discussed in chapter 4), there is an inherent problem with Schaeffer's spatial understanding with the proposal of "time as the space where the object inheres" (Chion, 2016, p. 172). This has also been reflected by Ihde, in that immersion and awareness of spatial boundaries are always temporal (Ihde, 1976). Of course, Schaeffer did not have access to the spatialization technologies we have today but his discussion on spatial perceptions of sounds is missing, rather stating that: “Sound objects, unlike visual objects, exist in duration, not space: their physical medium is essentially an energetic event occurring in time” (Schaeffer, 2017, p. 190). This view on space is ambiguous yet is related to his dismissal of “the myth of spatial music” (Schaeffer, 2017, p. 325).

We can derive spatial information from vision, audition, and haptic modalities; we do not rely on just *one* single sense modality. Rather we are dependent on the representation of the sense modalities as spatial and “understanding spatial configurations in the world involves either translating modality-specific information into a common format or providing interfaces between modalities” (Landau & Jackendoff, 1993, p. 217). Spatial understanding maps onto an understanding of objects and shapes, which impose a basic constraint between spatial representation and objects. For an understanding of shape, and its morphology, we look towards its *morphogenesis* or its pattern formation (Lesne & Bourgine, 2010) and its possible relationships.

Kubovy and Van Valkenburg argue that “auditory localization is in the service of visual localization” (Kubovy & Van Valkenburg, 2001, p. 99), which suggests that we need visual correlates to spatial hearing and these in turn informs our perception of sound images. We have already discussed that sound images serve as mental images of the sounds we hear and the images we hear “can integrate different listening approaches and provide an understanding of both the intrinsic and the extrinsic aspects of sonic experience” (Barreiro, 2010, p. 36).

Sound is immaterial, in that we cannot physically touch it. Yet it is tangible through its a physical presence as vibrations in a physical medium, also as encoded to discs or digital formats. The sound object is not just temporal, it is spatio-temporal. When considering objects, motions, gestures, time, and combinations of these without space, they loose context. Space exists independent of the coordinate set we use to describe spatial positions of objects through software, and the combination of this space and sounds creates a spatio-temporal experience that can immerse a listeners physically. A sonic shape is not limited to the implied size of the sound nor to its dynamic amplitude envelope, but rather a sonic shape is the defining element that makes us experience the spatiality of the sound, its localisation in three-dimensional space along with the size, mass, motion, and dynamic profiles. Perceptions of shapes and surfaces are interfaces to a spatial understanding of sound.

In section 1.3.3, we discussed a structural model that borrowed some aspects from Bruno Latour’s actor-network theory. This structure defined that each object we interact with consists of a shape, a site/location, and a model. We then saw that each object is organised in a network-like structure that draws on the actor-network theory's idea of a flat ontology. What we can draw from this approach, based on shapes and the network of structures we perceive, is that the sound objects we hear, study, analyse, and interact with belong to a wider contextual structure. Based on Wishart, in chapter 3 we will define this structure as a *sound landscape*. It is an attempt at exploiting and expanding the concepts described in *reduced listening* where we direct our intentional focus to different salient features of a sound, the values it carries and the events contained in the object.

2.5 The black box

Through "Opening Pandora's Black Box" Bruno Latour introduces the concept of *blackboxing* into his investigation of scientific knowledge (Latour, 1987). This is a process to uncover how scientific knowledge and technologies, with examples from biology and computer science, are revealed and encapsulated through processes of communication. The black box also represents a model of abstraction, where the complexes happening within a device are hidden and the user only has access to a small, designed interface.

> While the term "black box" is not of Latour's own invention, he deserves much of the credit for importing it into philosophy. A black box is any actant so firmly established that we are able to take its interior for granted. The internal properties of a black box do not count as long as we are concerned only with its input and output. (Harman, 2010, p. 33)

This is also reflected in the axial feature space of morphodynamic theory, with its distinction between a control sphere and morphology sphere, discussed earlier. Not unlike Heidegger's tool analysis, the black box allows us to disregard the large network of connections between actants, as long as they function smoothly.

With *blackboxing*, Latour chose a visual metaphor that is also widely adopted in software engineering and in audio applications, and also in terms of the evaluation musical instruments. The black box model describes a way of evaluating a system of which we do not have any knowledge about its internal functions. We input a signal and evaluate what ever change we see on the output. This model provides a way to review, discuss, and understand changes brought on by a system. For example, we can evaluate the differences between the various layers of a composed, spatial scene when mixing in stereo, to presentation over a multichannel speaker array, and from there make necessary changes to individual layers. This describes a process of analysis–synthesis as discussed in section 1.2.1. The black box model can be used as a means to analyse and synthesise how we work with spatial sound (Holbrook, 2019) where the morphological traits of a room will impose timbral changes to performed multichannel music in ways that are almost impossible to foresee, and in this way change the heard music.

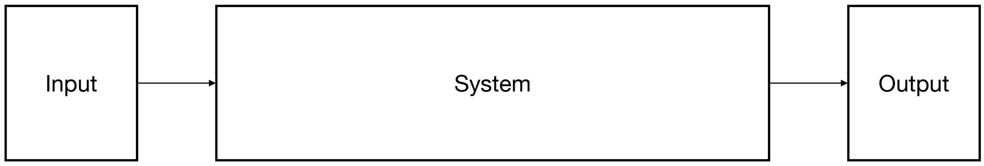


Figure 2.1: Simplified black box model, something is input to a system and we evaluate the system based on the perceived output.

Illustrating the functioning of the black box, the model in Figure 2.1 presents a structure adopted from the evaluation of musical instruments: a sound source and a sound modifier (Howard & Angus, 2009). The source is the input, the instrument, and the system is the sound modifier. For example, we can use this to measure the differences between plucking, bowing, striking, or dropping something on an instrument. This model isolates the instrument (or object) to evaluate it isolated from the surroundings, akin to using an anechoic chamber. This model can be extended further to encompass *two*, or more, systems of modifiers.

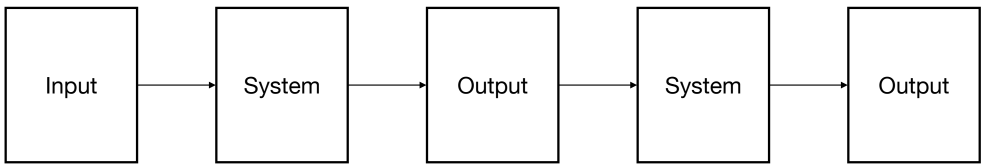


Figure 2.2: Extended black box model, which incorporates two systems of modifiers on a single input.

In the first model, the system that introduces changes on the input was from the instrument itself, seen in isolation. In this the second model, we now have two system modifiers: the instrument's own modifier and the room where the instrument is played. The room imposes its own modification to the sound before it reaches the listener's ears. By only observing the model in Figure 2.1 we are effectively removing the context for how we listen and the important facets of how a sound exists. Predicting the timbral behaviour of an acoustic instrument in a given space is different, and easier given knowledge of the instrument, than music mediated by a series of loudspeakers. However, we must not be limited to the uses of the black box model in directing it only towards issues of musical instruments and biological processes. We can draw on the affordances of this model when evaluating places, spaces, sites, and locations. For example, if we have two recordings of a seagull, one made by the coast and one made in a city; does the seagull sound different in the two recordings? What are the differences? Are they of the same species? Are there backgrounds that influence how the sound of the seagull is understood? What are the spatial characteristics of the sound of the seagulls in the two recordings? What are the characteristics of the backgrounds where the seagulls are recorded? Are there differences in vocalisations due to the backgrounds?

Using the black box as a metaphor explicitly encapsulates *both* exploration (analysis) and boundaries (synthesis) between the known and the unknown. Considering this metaphor will be useful when we next discuss sound objects.

2.6 Sound objects

Schaeffer commented that he and his colleagues had practised phenomenology without realising it (Schaeffer, 2017, p. 206) through the use of the phenomenological *epoché*. Schaeffer and colleagues listened to sounds as loops on closed grooves pressed onto phonographic discs, and after listening to the fragments multiple times, they experienced a shift in attention away from the everyday signification, the event, of the sounds and more towards the sonic features of what they heard. This practice of "bracketing out" a sound, of removing information that is external to it, enabled a study of the features contained in the sound and through this process it was the sound material "itself" that became the object of study.

Edmund Husserl recognized the *epoché* as necessary to gain an understanding of the everyday objects which surround us, "this entire natural world therefore which is continually 'there for us', 'present to our hand', and will ever remain there" (Husserl, 2012, p. 59). This "natural attitude" is always present but in order to examine the world we need to suspend our understanding of it – we need to place the objects we wish to study "in brackets". Husserl substantiated this and said that "the thesis is 'put out of action', bracketed, it passes off into the modified status of 'bracketed thesis', and the judgement simpliciter into 'bracketed judgement"' (Husserl, 2012, p. 58).

Placing something in brackets involves an intentional selection of *what* should be studied. Intentionality and intentional focus is an object-direct focus "to be conscious *of* something" (Husserl, 2012, p. 171). Yet, *what* is placed in brackets? The reduction associated with bracketing out an object entails removing it from its context to see its constituent parts. If we bracket out, or reduce, an object from its surroundings then surely we can go the other way and bracket out the surroundings from the object to listen to the context. It would need a re-evaluation of what we perceive the sound object to be to accomplish this.

To arrive at the sound object, we must suspend our knowledge of the surrounding world in order to make subjective judgements about what we hear. The musical fragment is isolated from its cause and context, so we can listen to the perceived sound - not the physical signal. The object is not an end in and of itself but is a method of analysis to draw out multiple features in an ontologically complex sound, as already discussed. The immanent objectivity in this listening is grounded on what Schaeffer refers to as targeting “a particular object and the various modes through which I relate to this object: perception, memory, desire, imagination, and so forth. In what way is the object immanent in these? Because it constitutes an intentional unit, involving acts of synthesis” (Schaeffer, 2017, p. 207). This comment underlines that a sound object is a multidimensional unit, which signifies that a sound object is ontologically complex and can contain multiple significations at the same time. Through directed, reduced listening we can shift our intentional focus to the various features we wish to study: “The idea of the sound object is critically important because it represented the appreciation of the traits that make up the composition of a sound” (Holmes, 2012, p. 47).

At this point, a distinction should be made between the *sound object* and the *sounding object*. The term sounding object refers to the source of the sound, specifically as it was explored in the Sounding Object project (Rocchesso & Fontana, 2003). If we talk about a sounding object, then we relate this to a specifically located source within the listening environment, how we can examine the processes that produce the sound and how to model these processes; like that of rolling or crashing objects of different materials. The study of the sound object disregards the events of rolling and crashing, rather we are interested in what features we can find in the sound of the rolling and crashing and how it effects the sound we hear. All the features of the sound influences how we hear it, “space is *in* the sound. The sound is *of* the space” (Worrall, 1998, p. 97). The sound is produced by a source, but we focus on what we subjectively hear.

First, objects are classified and described. Then by intentional analysis of their features, the composer can synthesise new sound objects (Landy, 2007). The origins of the original sounds are quickly lost, when sound materials have been transformed or different sound objects have been combined or superimposed onto each other. Indeed, for many composers working with transformed sound material from many different sources “it becomes difficult to remember from where the various sounds originated” (Wishart, 1996, p. 67). It is not only when a sound is removed from its visual context that it is perceived to be acousmatic, new layers are added when sounds are transformed, combined, and re-synthesized. It is when the sound object is contextualised in a performed, composed work that we can start to experience the perceptual effects of the interplay between the sounds and the performance space, which ultimately influences our understanding of the resulting auditory percept.

Musique concrète was theorized and practised as a mode of making music from real-world, concrete sounds. Through a process of reduction we could extract musical values and suitable materials for music composition from the materials of the real world. Schaeffer focussed on the primacy of the ear (Landy, 2007) and sought to create a mode of music-making that transcended the abstract forms of musical notation which dominated classical music. Using the acousmatic curtain as a metaphor to describe a method of separating sound and source, Schaeffer was very clear on what the sound object was not:

1. The sound object is not the sound body,
2. The sound object is not the physical signal,
3. The sound object is not a recorded fragment,
4. The sound object is not a notated symbol on a score,
5. The sound object is not a state of mind. (Chion, 2009, pp. 32–33)

This list does not, however, tell us what the sound object *is*. To access the sound object, we must suspend our knowledge of the surrounding world and what we know about the sound we are listening to, in order to make subjective decisions about what we hear. This is not a permanent "state of mind" nor is it to aim for an essentialist perspective of "sound-in-itself". Rather, we listen analytically and we shift our intentional focus to different features of the sound. The sound object is a sound image in the mind of each individual listener, held together by an intentional focus on a set of features. As we have already discussed, the sound object is not one thing - the image held in our minds changes depending on our intentional focus on different sets of multidimensional features.

Through the phenomenological *epoché*, the listener has access to sound through nothing but its "appearance" and to make no distinctions as to the supposed origins of the sound. However, unfortunate interpretations of the practices around the sound objects has, for example, resulted in statements such as "the sound object is proposed as the ideal and objective form of the signal; the *essence* of any given heard-thing. . . . Acousmatic listening involves a naive, blank reception of the auditory. We asked to let sounds in the door without first asking, 'Who's there?"' (Kim-Cohen, 2009, p. 13). What, unfortunately, Kim-Cohen does not realize is the practice of reduced listening is a temporary, not a permanent, suspension of associations of the heard sounds: "It does not oblige us to repress - even less to deny - our figurative and affective associations. It is only a matter of placing them temporarily outside of the field of naming and of observation" (Chion, 2016, p. 171). In the visual arts also, there was a rejection of the uses of everyday objects where the reduction of the everyday significance of an object to art stated that "the rejection of the readymade by critics and artists such as Greenberg and

Newman was shaped by a fear of the collapse of categories, the fear of identity, of the work of art becoming just another 'arbitrary' object" (Lütticken, 2010, p. 2).

Brian Kane noted on Schaeffer's practices around the sound object that "Schaeffer, unwilling to see his own composing and theorizing as historically conditioned, deludes himself into describing a sonic material that necessarily stands outside history" (Kane, 2014, p. 38). The insistence of the phenomenological reduction poses a naive perspective of the world, by which we can only say something about that which is present to us. By insisting on a phenomenological reading and state of sound perception, it posits an a-historical and a-social perspective on sound that somehow there should be a special treatment for sound from the other sense modalities. The reduction of a sound away from any surrounding or any context, provides an architecture for immersive aesthetics (Schrimshaw, 2017) where any differentiation or representation is actively removed from the reading of an object.

The ontological reduction avoids a historical conditioning, rather it posits that all we perceive belong to a beginning, "through a sleight-of-hand, phenomenology covertly places its ontology prior to experience, and then subsequently 'discloses' the ontological horizon as if it were always already present – as if its ontology made experience possible in the first place" (Kane, 2007, p. 21). There is a desire to keep a sense of purity of experience, which overshadows the potentials in creating a practice where both acoustics and musical exploration could be on equal ground.

2.6.1 The acousmatic

According to antique Greek history, the Acousmatics were the students of Pythagoras, who listened to their master's teaching from behind a curtain (Kane, 2014). The students listened in silence to focus on the content of the lecture and not on the gestures or facial expressions of the master. The adjective *acousmatic* refers to hearing a sound and not seeing the source (Schaeffer, 2017) and this became a central tenet for the development of the idea of the sound object: "Acousmatic listening is the opposite of *direct* listening, which is the 'natural' situation where sound sources are present and visible" (Chion, 2009, p. 11). Experiencing any kind of mediated sound, such as through the radio, records, or telephone is an acousmatic experience, likewise the back of the head is naturally acousmatic and sounds coming to us from outside of our field of vision are equally so. The acousmatic, then, refers to the apprehension of a sound without seeing the source and this also includes the identification of the source without the help of vision. Importantly, it affords us an interface to the discussions we have had so far, that we should focus on the subjective perception of the sounds in correlation with the physical signal and not on the event that caused the sound.

Even though Pythagoras' students listened to his lectures from behind a curtain, we may assume there was never any doubt that they were hearing their master's voice. If we hear a recording of a violin played in the conventional way, there is generally no doubt as to what instrument we are hearing but the lessons learnt though the acousmatic allow us to tune in to different aspects of the violin sound and to focus on certain features of the sound. This will be discussed further in section 2.7.1.

Insisting on *acousmatic listening*, by only focusing on the ears and not on the eyes, does not entail that the auditory has some higher status than the visual. It is important to bear in mind that Schaeffer was drafting the theories of a new type of music using new technologies, where *the practice* of music-making created new avenues for musical research, where the subjective experiences of the heard sound, and how we can draw musical features from theses sounds, were the most important. It is not merely about the supposed *visual* aspect of the source, the violin, the player, or something else, but also about "the social or environmental origins of the sound (from a bird, from a car, from a musical instrument, as language, distress signal, as music, as accident" (Wishart, 1996, p. 67).

For Schaeffer, the loudspeaker became the Pythagorean curtain, likewise for listening situations that utilize multiple loudspeakers to transmit sonic material to a listener, the fact that the back of the head is naturally acousmatic is often overlooked. Of course, we have no source of whatever the sound might originally be but the new source becomes the loudspeaker. This kind of dislocation can be experienced from Denis Smalley's *Empty Vessels* (1997), where the surrounding soundscape from his garden is recorded from the inside of two large garden pots. The angle is not human, and the resonances produced by the pots certainly is not related to the human hearing mechanism. Yet, it is distinctly human, by way of the dislocated sounds, the surrounding man-made sounds, and the composer's intent through mixing.

Using listening as a form of aesthetic exploration through categorization of sound, Tine Surel Lange's series *Works for Listening 1-10* (Surel Lange, 2021) creates sound environments from everyday objects that are abstracted from their original state. The sources included metal and spoon, metal and water, roof work, wires, among others. One of the stated aims by the composer is to make people listen more intently to the experiences in life and to the objects and environments that surround us.

The "voice behind the curtain" is not only relevant for music mediated by loudspeakers, but is also prevalent in sound design. Sounds are recorded, listened to, evaluated, and processed through layers of effect chains in order to represent the sounds as something completely different from their origin. The processes of evaluation and manipulation (analysis–synthesis) in sound design practice bear many relationships to Schaeffer's practice and theorizing around the sound object.

A fundamental functioning of our auditory system is that we can direct our visual attention towards a specific direction to determine the next course of action. In "everyday" listening situations, we can see the sound sources and our experience of the audible source and the visible source are coupled. Hearing sounds behind us, where we do not have eyes, the experience is de-coupled so we need to either turn our heads to see the source or we listen and make assumptions based on the sound we hear.

Our listening is guided, and effected, by something hidden. Here I will borrow a term from Chion's masterful treatment of film sound in *Audio-Vision*, that of the *acousmêtre*. Chion refers to the acousmêtre as an acousmatic character, hidden from view (off-screen) who creates a sense of ambiguity to the scene of a film. It is a character who hides "behind curtains, in rooms or in hideouts" (Chion, 1994, p. 129) and is implicated in the action but is never physically part of it. In film, the acousmêtre is almost always a voice, but the *extended* powers of the acousmêtre can serve as a link for considering the acousmatic nature of a sound as it is presented to us intrinsically, to the contextual presentation extrinsically.

In cinema, the powers of the acousmêtre rests on the character being hidden from the action on screen. In acousmatic music there are multiple voices with no visible body, the *site* is acousmatic, it is hidden from view but the sounds themselves are fundamentally of the site, as they belong to the context from where we view and hear the foreground. The place does not need to be a physical location, but the place as a site of understanding is the central focal point in this mode of listening. Each individual sound becomes a separate protagonist that links the features of a sound object to the features of the context of presentation, the creation of a sound scene and the landscape we listen to.

With sound recording and reproduction, the "source" is now the vibrating cones of the loudspeaker. The acousmatic nature of the sounds and our understanding of them affords us the possibility of defining virtual acoustic landscapes, where we *imagine* the sounds originating. With the virtual we no longer need to consider a "real" landscape but rather can use the potential expressions and explorations of the landscape to place the listener inside a sounding environment. Luc Ferrari referred to his practice of composition as *anecdotal music*, which retained much of (or all) the source-origins in the sound, exemplified clearly through his *Presque Rien No 1* (1968) (Ferrari, 2009), where he compressed several hours of recordings from a small fishing village in Croatia into twenty minutes (Wishart, 1996, p. 136), which transformed the landscape of the beach. Here the acousmatic is an interface towards a spatial, and contextual, understanding of sound, rather than just the isolated recording. Acoustic landscapes and virtual acoustic landscapes will be discussed in chapter 3 with references to the practice of soundscape recording.

2.6.2 *Sillion fermé* and *cloche coupée*

The origins of the programme of musical research by Schaeffer and colleagues in the early days of *musique concrète* sprung from listening to sound fragments engraved onto closed groove phonographic discs (sillion fermé) (Schaeffer, 2012, pp. 10–15). At the time, mixing of sounds was done by starting and stopping playback of these looped fragments, and repeated listening caused the composers to shift their attention to other features of the sounds they heard. The closed groove afforded a listening to individual fragments many times, where "repetition musicalizes the sound fragment by removing the dramatic and anecdotal traces of its original causal context" (Kane, 2014, p. 16). This repeated listening led to the formulation of *reduced listening* where we suspend our knowledge of a sound's origin and causal relationships.

The technology at the time, before tape recorders, offered limited flexibility in terms of composition but afforded the ability to listen to small sound fragments indefinitely. Listening is a cognitive process that attempts to detect differences and interpret information from the environment (Truax, 2001). The repeated listening afforded by the available technology allowed the musical experimenters to shift attention away from the supposed origins of the sounds towards the sonic features that make up the sounds (Godøy, 2019).

The effects of this practice led Chion to his mock-biblical "sermon on the mount":

> Peter spake unto them this parable: "A man went forth to plough a closed groove. At the tenth turn his neighbours and his friends mocked him. But at the thirtieth turn there was more music than in all the fields around. Verily, verily I say unto you, cultivate your Perceptual Field, and the Kingdom of Music shall be yours". (Chion, 2009, p. 5)

The loops of the closed groove discs allowed for a consciousness where "the sound fragment could be described *in itself*" (Chion, 2009, p. 13). If the process of sillion fermé afforded the musical experimenters an interface to *shift* their attention after repeated listening to a sound, then the cloche coupée (cut bell) gave them access to the *shape* of the sound.

Experimentation with the "cut bell" involved removing the attack of the sound, which proved to alter the timbre of it. This made Schaeffer realise that the timbre of a sound is not only linked to the distinct harmonic spectrum that makes up the sound but also describes the temporal *shape* of the sound. Without the attack the sound alters its immediate signification and opens it up for new interpretations: "The 'bell loop', in truth, was at that time no more than a closed groove from a bell taken at a carefully chosen moment of its resonance, and it made a sound like a flute" (Schaeffer, 2017, p. 332). By cutting away the attack and changing its dynamic profile, the shape of the sound was altered. A change in shape, in the sound's *intrinsic* features, influences the *extrinsic* features and ultimately how we can relate this sound to the "horizon" of our experiences.

2.6.3 Technology

It can be easy to romanticize the "genius" and "ingenuity" of these early practices. Kane iterates that although technology is important to the understanding of the sound object "it would be a misunderstanding of Schaeffer's thinking to assume that the sound object is in any way the result of modern sound technology" (Kane, 2014, p. 33). This is an unfortunate perspective, as had it not been for the technological means available at the time, it would have been difficult for Schaeffer and colleagues to arrive on the track to the sound object. Indeed, these practices are is afforded by the technology, "taken as such, technology is understood more richly as the emergent result of people interacting with tools (and other people and so on)" (Green et al., 2019, p. 4). I will extend this further, to say that the site we occupy and represent is part of the emergent interaction between listener, composer, and performance space.

The desire to know the origin of a sound seems to come naturally to most listeners. Identifying a sound and its cause could save you from danger or it could inform you that something you desire is approaching. Listening *into* sound reproduced from loudspeakers could seem to emphasise the desire to uncover the origin of the sound as we have no visible source to relate it to. To structure our thinking and practice, we need to consider how we use technology. As already discussed, the sound under discussion is technologically *created* and technologically *mediated*. In many respects, the work becomes inseparable from the technology that creates it (Essl, 1992) and by ignoring the technology we also ignore aspects of the work itself. Through spectromorphology, Smalley iterates that we must ignore the technology behind the making of the music (Smalley, 1997) due to the fact that "the separation between the act of sound-making and perception, combined with the specialised nature, proliferation and transience of methods and devices, indicate that technological knowledge cannot be part of any method founded on perceptual consensus" (Smalley, 1997, p. 109).

What Smalley advocates is that the technology should always be separated from the perception of the music. The composer and the listener should be "untainted" by technological listening (discussed further in section 2.7.1) to maintain an unclouded judgement. However, it is precisely through the technological mediation that acousmatic music, and the acousmatic situation, exists.

2.6.4 Linear and non-linear relationships

Non-linear relationships exist between the physical signal and the subjective perception of the same sound. These relationships are always present, as we do not only listen to "the sound itself", but also listen from a particular spatial position and it is "not the psychology of the auditor that matters, it is the particular spot where the latter is positioned that does" (Chion, 2016, p. 172). We are concerned with exploring the correlations between the physical signal and the subjective perception of the sound, specifically the spatial relationship existing between the signal, the space, and the listener.

Schaeffer referred to *anamorphosis*, or warping, as the possible non-linear relationships between the physical signal and the sound object, that was characterised by some irregularities that suggest a distortion of physical reality (Schaeffer, 2017). This concerns the mapping of correlations between the subjective images and the acoustic basis in sound. Temporal anamorphosis leads to "time warping" that describes how a "listener's perception affords conclusions that do not concur with physical reality" (Landy, 2007, p. 79). Importantly, this angle on perception was a perspective Schaeffer (and Chion) applied to draw distinctions between the work of the musical experimenter and the physical measurements of the acoustician. The acoustician would seek to explain the black box of the individual (Chion, 2009, p. 16), but the musical experimenter should study the sound world for itself. Observing a sound object is only possible by listening to it again (Chion, 2016) and by each subsequent listening we understand more of the sound's features.

Anamorphosis[4] is a visual distortion that requires the viewer to be in a specific location to see the correct image(s); it is a technique to create pictures within pictures. One example is of Hans Holbein's painting *The Ambassadors* (1533),[5] a much-cited double portrait of two unknown ambassadors with a still life. The painting features a smeared shape across the front of the painting. This shape reveals itself to be a human skull when viewed at a sharp angle from the right, an example of a *memento mori*. Another example is the work of Maurits Cornelis Escher, where the pictures within the pictures are accessible for the viewer from one position, depending on where we focus our gaze. This warping of an image indicates that it is the subjective perception, from a specific angle, that should be considered important, but not the only thing we should attend to. The correlations between the physical signal and the heard sound are important, because what we think we hear is not always what we do in fact hear. This approach allows us to explore the deeper facets between the physical signal propagation and our subjective

[4] A short animation by the Brothers Quay explains the historical origins of anamorphosis in Renaissance painting: https://www.youtube.com/watch?v=cEfwbnMf3jM

[5] https://www.nationalgallery.org.uk/paintings/hans-holbein-the-younger-the-ambassadors

perceptions.

The signal is a carrier of information but it is not the information itself, it is a representation of information (Garnett, 1991) - the physical experience of music is related to the physical vibration that propagates through a medium before it reaches our ears. For example, in the field of sonification and auditory display, the sonification process must be rooted in the data it presents but what is perceived is still sound, from which we can extract information as we would with listening to any sound. The information contained in the sonification should be perceived by the listeners (Grond & Hermann, 2014). The perceptual experience is the psychoacoustic feature attributed to how we make sense of what we hear, and the cognitive features surrounding the listening experience determines the structures we make of what we hear and what it means to us. This perspective reminds us that the black box model discussed in section 2.3 treats the room as an "unknown" and it is through listening to the resulting sound that we can establish the changes imparted to a sound by the space we are listening in. We use the correlations between the signal and the sound as the basis for a continuous analysis and synthesis. For more discussions on correlation, see section 4.3.4.

These non-linear relationships are clear in acousmatic sound mediated by loudspeakers. In traditional instrumental music we can, by vision, see the correlations between the instrumental sound and the gestures of the performer, as we can clearly perceive the relationship between the music and the gestures of the performer. Thompson, Graham and Russo find that the visual impact of popular music recordings enhances a sense of phrasing and anticipation of emotional changes in the perceivers (W. F. Thompson et al., 2005). Later, Thomson, Russo and Quinto find that facial expressions in singers greatly influence the emotional interpretation of the music (W. F. Thompson et al., 2008). Griesinger has found that the aural impression of a concert hall and room acoustics is often dominated by the visual impression (Griesinger, 1997). By vision we can separate sources from the space and make judgements on the music we hear, the acoustics and our perception of it. Indeed, in acousmatic music we have visual contact with loudspeakers as an indication of the location of the sound source, despite the "actual" source remaining hidden.

2.7 Models of listening

The sound object, the acousmatic situation, and anamorphosis point to one thing: *listening*. This section will discuss some approaches to, and models of, listening and presents an overview of different approaches to listening that are relevant to this thesis. Listening involves listening *for* something, that is, we *listen for knowledge*. We listen to attain knowledge:

> On the afternoon of July 11, 2014, Dutch Public Radio 1 broadcast an interview with science journalist Diederik Jekel. He had breaking news: American geologists had discovered a "super ocean" some 300 miles below the earth's surface. The journalist immediately added a qualification. What the Americans had actually found were some stone minerals, originating from the earth's deep layers, that included water molecules. This prompted the talk show host to ask how certain scientists could be of the super ocean's existence. The journalist explained that the American geologists knew they had in fact found water when they sent sound waves deep into the earth. The stones had melted in the earth's heat, and just as a knock on a table sounds different from a knock on a glass of water under the same conditions, the melted material sounded different from the non-melted. The talk show host was quick to conclude: "We know it", she said, "but we have not seen it; it has not been [proven] experimentally".
>
> Apparently, she had trouble believing the geologists' ears. Their findings had not yet been proven, because the phenomenon had not been seen. By suggesting that hearing something is not sufficient to prove its existence, whereas seeing it would actually establish the fact, the interviewer posited a direct link between seeing and true science or ultimate knowledge. She may have learned to do so from scientists themselves, who tend to work in offices packed with printouts and scans around computer screens, producing publications rich in diagrams, graphs, and other images. Indeed, the American geologists in search of water had used seismic data gathered during earthquakes, often referring to infrasound waves: sound waves below the human audible range. At times, these infrasound waves are translated into frequencies that humans can hear, but more often they are made visible, in graphs. (Bijsterveld, 2019, pp. 1–2)

Listening for knowledge is performed by doctors with stethoscopes, marine biologists using hydrophones and by composers, artists, viewers, and listeners. It is an active process to gather information in our interaction with the environment (Tuuri & Eerola, 2012).

Listening is a complex field. From a psychoacoustic perspective, our ears receive information from all angles, which has been reflected off surfaces surrounding us including our own bodies before entering the ear. Hearing is a basic function of the auditory system, which allows us to navigate our surroundings (for descriptions and discussions of the anatomical functions of our hearing systems see (Howard & Angus, 2009; Moore, 2003)). It is our listening intention that indicates how we attune to something specific, to specific features of what we are listening to. Jonathan Sterne, by citing Don Ihde, contends that listening is a directed and learnt activity, with roots in cultural practice, and that “perhaps the biggest error in the audiovisual litany lies in its equation of hearing and listening” (Sterne, 2003, p. 19). We can say that listening is intentional hearing (Carter, 2004), but listening is always dependent on hearing but cannot be reduced to it. Listening experiences display complex spatial attributes, and this section is a bridge to the next chapter where we will discuss the structural features and relationships among objects.

In Schaeffer’s programme of musical research concrete, real-world sounds were prized over synthetically generated sounds, like that of *electronisches musik* which he called “music *a priori*” (Schaeffer, 2017, p. 7). The reason for this was a desire to focus on the sounds of the everyday for their intrinsic musical qualities. It would seem that the primary resistance to *electronisches musik* was because of the serialist approaches of composition (a priori) and the “locked” perspective it afforded. Rather, writing in 1966 Schaeffer states “two unusual modes of sound production, known as musique conctète and electronic music, came into being at about the same time, about fifteen years ago. These developments were in opposition for more than twelve years, before several complementary aspects were revealed” (Schaeffer, 2017, p. 2). Indeed, Schaeffer’s “late repentance” can be found in his electronic composition *Le Trièdre Fertile* (1975-1976) (Schaeffer, 2015).

This section will present and discuss five different perspectives of, and models of, listening based on the work of Schafer and Truax from the World Soundscape Project, Katherine Norman’s concept of “real-world listening”, Schaeffer and Chions models of listening, and Gaver’s ideas on ecological acoustics.

2.7.1 Schaeffer’s four listening modes

Central to the development of Schaeffer’s program of musical research, was the primacy of the ear. Discussed in section 2.6.1, the discovery of sounds decoupled from their source gave rise to an understanding that we can approach music-making by using our ears, and not through symbolic notation. This “new” music was designated as *concrète*, it was based on concrete, real world sounds of the city, nature, objects, instruments, and the like. This was unlike *abstract* music, which was first conceived as an abstract idea and then realised symbolically. In the listening modes, Schaeffer conceived two axes that guides the listening intention: abstract/concrete and objective/subjective.

The two axes refer to different “sets of comparison”, the *subjective* part of the axis concentrates on the person who is perceiving and on the subjective perception of the sound. The *objective* refers to the objects of perception. Within these modes, the *concrete* refers to the sound event or causes behind the sound, and the "raw" sound objects - the recognition of the real-world source and its agent. The *abstract* side of the axis is the shift in attention towards certain qualities of sounds, and towards the comprehension of meaning through signs, or codes, based on the perceived features of the sound.

The four modes of listening are divided into two categories of abstract/concrete (horizontal axis) and objective/subjective (vertical axis):

- The first, *listening* (écouter) is objective/concrete, we treat sound as a *sign* of a source, where we identify the source, the event, the cause of the sound. We direct ourselves towards someone or something that is signalled or described by a sound.
- Secondly, *perceiving* (ouïr) is subjective/concrete, we perceive by ear, we hear passively and are not listening with intent.
- The third part, *hearing* (entendre) is subjective/abstract, we display an intention to listen, and by that choose what interests us to listen to and what features to focus on.
- The fourth and final mode, *comprehending* (comprendre), displays a *listening intention*: we understand the meaning conveyed by the sign, we grasp a meaning of the sound by our shifting intentional focus. This is objective/abstract.

The four modes of listening around these two axes was structured with the aim of understanding the correlations between the physical signal from the sound event and the subjective listening to an abstracted sound. This was designed as a "circuit" diagram.

Table 2.2: The "circuit" of the four listening modes, we move from the *concrete* to the *abstract* - from listening to the event and causes of the sound (1) to understanding the intrinsic features of the sound (4). The subjective/objective indicates that we turn towards the object of perception, to the perceiving subject and back to the object of perception.

Abstract	Concrete	
(4) Comprehending (*comprendre*)	(1) Listening (*écouter*)	Objective
(3) Hearing (*entendre*)	(2) Perceiving (*ouïr*)	Subjective

The diagram in Table 2.2 is not conceived as a chronological sequence but a circuit where "perception moves in every direction and where the four sectors are most often involved simultaneously, interacting with each other" (Chion, 2009,

p. 20). The four modes of listening are not isolated "modes", rather the four modes operate together and are interdependent.

The arrangement of the four listening modes in Table 2.2, where they are clearly located, "will help us to understand not only musical research and the functioning of traditional music, but also the relationship between music and language and the physical signal and the musical object" (Chion, 2009, p. 19). In other words, listening will reveal knowledge about the correlations between the signal, the sound, the spatial attributes and the wider contextual space.

With listening, we can access the various features of the sound object (shape, site, model), through its mediation to us as a sign. When shifting our intentional focus to the features contained in the sound object, we cannot see the source and this is no longer of any relevance to us, as we focus on what we hear and how we hear it. The potential references to an external sound event will still be evident, depending our intentional focus, likewise "smoke is only a sign of fire to the extent that fire is not actually perceived along with the smoke" (Eco, 1979, p. 17). This will lead us on into *reduced listening*. As a correlate of reduced listening, the sound object "is independent of the material and physical cause of the sound" (Chion, 2016, p. 171) as a means to gain an understanding of what we are listening to outside of association and observations.

Smalley adds to Schaeffer's listening model a *5ième écoute*, that of *technological listening* (Smalley, 1997) or "recipe listening" (Landy, 2007). This occurs when a listener perceives or focuses on the technology behind the music, and how the music is generated, processed, recorded, or manipulated rather than the music itself, and this often occurs at times where the sources are hidden. Smalley also warned that the practice of reduced listening is as dangerous as it is useful:

> many composers regard reduced listening as an ultimate mode of perceptual contemplation. But it is as dangerous as it is useful for two reasons. Firstly, once one has discovered an aural interest in the more detailed spectromorphological features, it becomes very difficult to restore the extrinsic threads to their rightful place. Secondly, microscopic perceptual scanning tends to highlight less pertinent, low-level, intrinsic detail such that the composer–listener can easily focus too much on background at the expense of foreground. (Smalley, 1997, p. 111)

Smalley's technological listening is a type in indexical listening that is already encapsulated within the four listening modes, where it is our *intentional* focus that guides our listening and the features we choose to focus on. Importantly, with the idea of the acousmatic in mind, Schaeffer's aim was that we should not describe a sound in terms of its supposed origin but rather in terms of its morphological, sonic qualities.

2.7.2 Reduced listening

Of all approaches to, and models of, listening, reduced listening has often been misunderstood. It is unfortunate, as we all practice reduced listening in one form or another throughout our lives. Joanna Dermers highlights one of the problems as "pure reduced listening is virtually impossible when dealing with recognizable instruments and for many electroacoustic composers is unsettlingly ahistorical, although some advocate it as one component of a larger listening experience that simultaneously acknowledges the external associations of sound" (Demers, 2010, p. 29). Reduced listening is not about repressing associations nor denying ourselves access to the source or its meaning; it is about a *temporary* suspension of our knowledge about the sound in order to hear it for itself, as a sound object. It is also important to remember that reduced listening and the practices surrounding the phenomenological reductions took place in the studio, and Schaeffer readily acknowledged that performance of music would change the reception of it.

It is unfortunate that reduced listening in much literature is often seen as a separate endeavor from the other listening modes, indeed as Smalley indicated previously. It is not a mode of listening that exists on its own, Schaeffer's four listening modes exist side by side in a symbiotic relationship and are interdependent and interrelated (Kane, 2014). By temporarily suspending figurative and affective associations about a heard sound, we gain an understanding about the features of the sound itself, based on our intentional focus. As an extension of the discussion on Kim-Cohen's understanding in section 2.6, there are similar misunderstandings as to the nature and practice of reduced listening. Others have reduced the acousmatic listening situation to "a darkened room, sets of multiple loudspeakers and a mixing console. In this sense, what is staged is the sound object without external interference or reference as architecture built only in sound itself" (LaBelle, 2015, p. 30). What LaBelle is actually referring to here is a studio-like listening situation, and not a concert reception of acousmatic music. This misunderstanding can be forgiven but this image is still a widespread misunderstanding of Schaeffer's music theory, namely that, somehow, we are seeking to achieve some form of sonic contemplation that allows us to hear sounds in a completely dissociated way.

What does reduced listening afford? What does it even mean to "reduce" something such as listening? Reduced listening is a temporary suspension of our understanding of our surroundings, the sound producing objects, and events around us in order to listen to a sound "as itself". This "as itself" is a reference to the features of the sound: its texture, mass, fluctuations, timbral qualities, spatial attributes, and so on. One of the purposes of reduced listening is analysis, where we examine the building blocks of the sound itself. It is not censure, nor is it some form idealized state of insight. It is a temporary suspension of our field of observation, said differently: it is a shift in intentional focus.

Reduced listening is a way to teach oneself listening, concentration, analysis, and an artistic sensibility towards sounding materials. Then, our initial question again: What does reduced listening afford? Sound, or *the* sound, ceases to be placeholder for something else and affords a listening to the sound for it potential sonic and timbral properties. As such, it is an exercise in listening, in concentration, in analysis, and, ultimately, of synthesis. We all perform acts of unconscious reduced listening: If we hear a sound, (gasp) and think “What’s that sound?”, listen, and deduce that “it is just my cat”, then we have for a brief second listened intensely to the sound itself, to the nature of the sound, to its properties, and reached a conclusion as to what the sound is (as we saw at the end of section 2.3.1).

2.7.3 Chion’s three listening modes

Drawing on Schaeffer’s system of listening, and part of his treatment of sound on film, Michel Chion describes a three-part listening model of *casual, semantic*, and *reduced listening* (Chion, 1994, pp. 25–34). Listening takes place at many levels, and again, it is our listening intention that guides how we listen. Casual listening involves listening to gather information about a source or cause; for example, a person’s voice or the sound made by a particular object. It is also a mode of listening to a particular category - that of a man’s voice or a songbird. Listening that refers to a code or language, is called semantic listening, especially studied in linguistic research. This is a listening mode which specifically targets messages, as in spoken language, morse code, and so on. Where Chion perhaps has the most in common with Schaeffer, is reduced listening. Reduced listening focuses on the traits of the sound itself, independent of cause and meaning: “Reduced listening takes the sound ... as itself the object to be observed instead of as a vehicle for something else” (Chion, 1994, p. 29). Yet for what reason should someone practice reduced listening in relation to something we look at? The answer is more obvious than one would think. In film, sound is extremely important and reduced listening “has the enormous advantage of opening up our ears and sharpening our power of listening” (Chion, 1994, p. 31).

It must be emphasised that the differences between Schaeffer’s and Chion’s listening modes is one of use. Schaeffer was explicit in the focus towards music-making and how we can free our process of creation from the conceptual, symbolic nature of music. While Chion here writes about how we perceive sound on screen, these three listening modes are always accompanied by the visual. Through the audio-visual contract, we experience a multimodal integration of the senses, where sonic and visual object coexist through synchronization and object-sound mechanisms (Farnell, 2010), where the viewer considers the sound and image coming from the screen to belong to the one and same world. This is a symbiotic relationship between the audio-spectator and what happens on screen.

2.7.4 Soundscape listening

The World Soundscape Project was established by R. Murray Schafer at Simon Fraser University, Vancouver, Canada in the 1960s. The group of researchers involved in the project worked on the documentation of the *soundscape* from urban and rural environments. Schafer defined the soundscape as

> [T]he sonic environment. Technically, any portion of the sonic environment regarded as a field of study. The term may refer to actual environments, or to abstract constructions such as musical compositions and tape montages, particularly when considered as an environment. (Schafer, 1994, p. 274)

An important concept in Schafer's soundscape is the notion of *schizophonia*, which is an acousmatic reproduction of the sound and "the split between an original sound and its electroacoustical transmission or reproduction" (p. 90). For Schafer this was a lost connection, where people lost contact with nature through the loss of hearing nature in its correct place of appreciation. Yet, the soundscape is not some naive concept that objectifies nature, rather it examines a relationship between sounds in the world and the people who hear them. It directly examines the relationship between an *actual* landscape, the sounds contained in it and the listening to those sounds. For Truax, the term soundscape "refers to how the individual and society as a whole understand the acoustic environment through listening" (Truax, 2001, p. xii), and listening creates a relationship between the individual and the environment.

As such, Schafer's perspective on our lost connection between the sound and its origin is antithetical to Schaeffer's. As noted, the sound object - the decoupling of a sound from its origin - is a path to a new practice of music, whereas for Schafer this leads people to loose contact with the *context* the sounds have to the world and to us.

In distinguishing the main themes in understanding a soundscape, our understanding of the sounds that comprise a soundscape is divided into three: *keynote sounds*, *signals*, and *soundmarks*. *Keynote sounds* refers to the musical use, where the key of a piece is its fundamental tonality (Truax, 2001). Keynote sounds can be overheard but cannot be overlooked, and are sounds that belong to a sonic background against which other sounds are perceived. The keynote sounds of a site are those defined by its geography and climate. *Signals* are foreground sounds that are listened to consciously. *Soundmarks*, derived from landmarks, are unique sounds that belong to or is specially regarded by a community. Soundmarks "make the acoustic life of a community unique" (Schafer, 1994, p. 10).

In terms of *soundscape composition*, the ear and the microphone are the starting points (Westerkamp, 2002). In soundscape discourse, (sound) signal and keynote sounds are contrasted, like the separation of figure and ground in visual perception (Schafer, 1994). However, it is when focusing on individual sounds to classify them as keynote, signal, or soundmark, that it is defined as a *sound event*. In Schafer's view, Schaefferian sound objects are laboratory entities and sound events are sounds experienced in a context - in a place at a specific time. As an example, if a church bell is recorded and analysed in the lab/studio, then it is considered a sound object, but if it is identified and analysed in the community, then it is a sound event (Schafer, 1994, p. 131).

Based on these categories, Barry Truax structures a listening system on how the brain processes information and structures significance from what we hear. This system is divided into three levels of listening:

1. *Listening-in-search*, which is a conscious search of the environment for cues, a search for messages that can be derived from the sounds we hear.
2. *Listening-in-readiness*, a state where we are ready to receive information but our focus is elsewhere.
3. *Background listening*, where we are aware of sounds but not involved in active search or recognition. (Truax, 2001, pp. 19–24)

Stemming from the practice of the World Soundscape Project, this system is based on listening in the natural environment and not necessarily to studio-based practice.

Although not affiliated with the World Soundscape Project, Luc Ferrari explored anecdotal music and created *virtual soundscapes*, where the soundscape composition is composed of recordings made in many different locations and through layering with electronic sounds, instruments, and so on, creates a "phantasmagorical composite" (Roads, 2015). *Anecdotal* refers to sound where the source is clearly recognizable. The drive to create anecdotal music was seen as an act of rebellion against the culture at the GRM (Caux & Ferrari, 2012), where Ferrari sought to keep the structures of *musique concrète* without disregarding the reality content of the music. He looked upon this music as "electroacoustic nature photographs" (Emmerson, 2007).

2.7.5 Real-world listening

Katherine Norman builds a framework of possible ways of listening where she defines "listening is as much a material for the composer as the sounds themselves" (Norman, 1996, p. 2). Norman's perspectives is an interface towards Luigi Russolo's (Russolo, 1913) "orchestration of real-world sounds" in the listener's imagination. This framework directly considers the composer and how different listening modes are used as artistic strategies, specifically for music that seeks to preserve a connection to the recorded source - pieces that are *about* the real world. Furthermore, these real world sounds are important for the creation of meaning, being, and the sonic implications from non-narrative sonic journeys.

Part of the framework are categories such as referential listening, reflective listening, contextual listening, real-world listening, self-intended listening, and composer-intended listening. Referential listening "invokes" contexts and references about a sounds relevance to our situation and connects sounds to objects. Reflective listening is an appraisal of a sound for its acoustic properties, as for example "the song of the sea" and this mode refers to the uses of metaphors to describe what we hear. Norman is particularly concerned with the importance of obtaining the "right" metaphor for an acquisition of knowledge.

2.7.6 Ecological acoustics

It is outside the laboratory that we encounter the world and the sounds that occupy it. Ecological acoustics is the analogy to the ecological approach to optics and visual perception (Gibson, 1986), and the ecological approach involves an analysis of the physical event, identification of higher-order acoustic properties and empirical tests of listener's ability to detect such information (Warren & Verbrugge, 1984).

William Gaver's studies point to how listeners are concerned with the objects and events which cause the sound, from studies he has found that only "[t]ruly acousmatic descriptions of sounds occur when their source events are unidentifiable" (Demers, 2010, p. 36). The listening and compositional strategies that grew out of Schaeffer's experiments sought to turn attention away from the supposed origins of the sound to focus on a sound's features for use in artistic practice. To reiterate from section 2.6.1, acousmatic sound is the perception of sound from where you cannot see the source, not from where the source is impossible to identify. Pythagoras' students listened to their master from behind a curtain but there was most likely no mistaking who was speaking to them.

In contrast, Gaver has constructed a framework that describes sounds in terms of their audible source attributes and, not unlike the Sounding Object-project (Rocchesso & Fontana, 2003), seeks to study the vast richness of everyday listening. The framework seeks to "evoke intuitions about the structured information provided by sound about events at locations in an environment and to point out the relatively direct route from the structuring event to the auditory system" (Gaver, 1993b, p. 26). Although the "acousmatic school", drawn from the work of Schaeffer, sought to disregard the source of a sound, the sound producing event can still inform us in the treatment of the materials. When approaches to ecological acoustics attempt to identify the sound-producing events behind the sound, it seeks a context to the lived experience of our everyday experience of sound.

2.8 Summary

Through the discussions in this chapter, I have built on the presentation of the features of objects from chapter 1, and have focused on the features of objects through discussions on auditory objecthood, auditory events, shapes, and sound objects. Rather than discussing objects from a perspective of the physical, sound producing source, the object is used to discuss features of sounds and how objects are perceived as being representations of dynamic, spatial, and timbral features as shapes.

The intrinsic and extrinsic features of an object are reflected in the models we employ and become clear to us through listening with a shifting intentional focus. As we move into a phenomenological perception of sound, it is through *spatiality* of sound that description may begin (Ihde, 1976). Our understanding of sounds and their features is based on listening and the complex spatial attributes we can uncover through listening.

The reduction associated with "bracketing out" an object entails removing it from its context to see its constituent parts. If we bracket out and reduce an object from its surroundings, then surely we can go the other way and bracket out the surroundings from the object in order to listen to the context. The next chapter will expand these ideas to the *spatial context* where these sounds exist, to the places where we encounter them, and expand these ideas to discussions around perspectives of space/place and site/location that make up the term *sound landscape*.

Chapter 3

Structures

> For many years, I have been moved by the blue at the far edge of what can be seen, that color of horizons, of remote mountain ranges, of anything far away. The color of distance is the color of an emotion, the color of solitude and of desire, the color of there seen from here, the color of where you are not.
>
> *Rebecca Solnit*[1]

Where the previous chapter discussed the various features of a sound object, this chapter looks at the structural relationships among objects, that is, the sound object's relationships to space, interactions with space, and existence in space. In section 1.3, we looked at questions surrounding the sound image, its root in semiotic theory and the possible significations a composer's intention can have on the perception of mental images evoked by music. I referenced Roland Barthes' concept of a "fabric of citations" as a means to discuss the multiplicity of interactions that can exist between symbols and meanings: "The reader is the space on which all the quotations that make up a writing are inscribed without any of them being lost; a text's unity lies not in its origin but in its destination" (Barthes, 1977, p. 148). We are dealing with the possible *contexts* in which the intrinsic-extrinsic relationship exists.

A *structure* is a circular organising principle and "it describes both the perceived structure and the activity of perception" (Chion, 2009, p. 60) as well as the relational properties of space and what relational properties one is focusing on

[1] Solnit, 2006.

(Marcolli, 2020). As we saw in the previous chapter, objects are only perceived as objects *within* a structure, and every structure is perceived as a structure of the objects that comprise it. These tangible structures between the invisible, yet enveloping sound and references to external spatial reality are parts of the guiding principle for this chapter. I have previously stated that "space" is not something extra, rather it is woven into everything, it is "a container for action, something which is involved in action and cannot be divorced from it. As such, space does not and cannot exist apart from the events and activities within which it is implicated" (Tilley, 1994, p. 10).

In this chapter the questions surrounding space and spatial perception borrow a term from Trevor Wishart: the *sound landscape* (Wishart, 1996, pp. 129–161). This term is used to discuss the structural relationships around the acousmatic notions of space, place, site, and location. The discussions will include site-specificity, psychoacoustics, and spatial practice through ambisonics. Referring to Tchaikovsky's Manfred Symphony, Wishart uses the metaphor of a *landscape* to describe the perceived *physical source* of the sounds we hear. Landscape is differentiated from *association*, where we are led to associate the acoustic events of the music to images of forests or fields, evoked by the music and the programme note. The landscape of this symphony "is however musicians playing instruments" (Wishart, 1996, p. 130). The *sound landscape* is a redefinition of this landscape to a virtual acoustic space, from where we *imagine* the sounds coming *from* (Wishart, 1996) through the possibilities afforded by loudspeaker mediation. I will continue to elaborate on the use of the landscape as metaphor in the following sections, focussing on the perspective that this landscape is *an area of activity*.

Seth Kim-Cohen argues for an expansion of sonic practice, which includes both composition and sound art and all the disparate fields that fall around and in between these two, *from* a focus on medium and material to that of conceptualism. The direction of art practices accepts "the appeal of ambient phenomena, like sound and light, is attributable to their evanescence, ineffability and immersiveness" (Kim-Cohen, 2016a, p. 4). Works of sonic art concern themselves often with their own materials and with the perceptual properties of immersion, vibration, and resonances (Waters, 2000). This call for conceptualism is to bring the focus and perspectives of artists out of the narrow vision of self and to look at the wider social and societal implications of sonic art. Writing on "expanded" fields of sculpture in the late 1970s, Rosalind Krauss refers to medium that "the situation of postmodernism, practice is not defined in relation to a given medium - sculpture - but rather in relation to the logical operations on a set of cultural terms, for which any medium might be used" (Krauss, 1979, p. 42). These perspectives will be discussed further later, under discussions on these features and relationships of sound and space.

If we go back to the discussion on the "tyranny of the visual" in section 2.1, the discussion centered on Jonathan Sterne's audio-visual litany (Sterne, 2003, 2011) and its critique of the move to give primacy to hearing as the most important sense modality. In the catalogue for the monumental ZKM exhibition *Sound art - Sound as a medium of art*, the curator Peter Weibel writes about the "quest for spatial sound", and gives an example of IRCAM's *Espace de projection* as being completely soundproof, so that "what one hears is truly the pure sound and nothing else" (Weibel, 2019, p. 64). This image of purity represents a decontextualisation of sound and a shift away from conceptualisation and back to a focus on its own materials. The "quest for spatial sound" likewise, through a focus on sound's purity, indicates how the spatial context is missing and this represents a radical aestheticisation of sound that only focuses on interiority and not exteriority, of affect over intellect and a lack of signification. This "purity" of sound is only possible if one is immersed in it (Sterne, 2003). There are different approaches to the representation and portrayal of space and place. Often, the listener can only experience this through the composer's microphones. However, the composer is completely absent from the scene.

The following three examples all use different methods and activities to present ideas and evocations of place to a listener, through the recorded medium. Kjell Samkopf's *Mårådalen Walk* (1993) is a record featuring "recordings of walking" (Samkopf, 1994). The recordings are of someone (presumably the composer) walking in different terrains of the mountains. This is a sonic portrait of place, through an interaction with the material substances found on site. The interaction with the materials in the mountain is not in the traditional *musique concrète* sense, but rather the recordings are untreated and situates a body in the landscape. *Kits Beach Soundwalk* (1989) (Westerkamp, 2010) is an acoustic narration and exploration of the experience of being on the beach, surrounded by the sounds of the sea, the barnacles, and the city. Hildegard Westerkamp narrates the scene and describes what we are hearing. After introducing the place, the relationships between her position, the sea, and the city, she plays with the technological mediation of the sound and demonstrates methods of sound manipulation to remind the listener that what we are hearing is recorded, transformed, mixed, and presented by a human. Described as a "zoom lens", this process is discussed by Emmerson as using "the real filters of electronic technology to enhance the psychological filters of perception" (Emmerson, 2007, p. 10). This perspective is similar to what we discussed in section 1.3, where the microphone and the camera are both interfaces for accessing that which the ears and the eyes cannot perceive on their own.

In Annea Lockwood's *A Sound Map of the Hudson River* (Lockwood, 1989), an aural journey is traced from the source of the River Hudson to the Atlantic Ocean, documenting the changing sonic textures of the river through changes in the terrain. Where *Mårådalen Walk* is concerned with the composer's interaction with the materials of the mountains and to the textural features of the different surfaces, *A sound map of the Hudson River* is directly concerned with the natural morphology of the river's sound, and how the terrain and the human environment is woven into its sound. The acoustic exploration in *Kits Beach Soundwalk* goes from exploring the macro structures of the city, to the micro structures of barnacles, and on into inner spaces and dreams. All three examples explore different aspects of what we are discussing under the heading of the *sound landscape*, an imagined source of the sounds we are experiencing. Westergard's *Kits Beach Soundwalk* makes the most literal exposition of this by narrating how the concrete and anecdotal experience of the beach can move into abstraction.

This chapter will discuss further these ideas of a sound landscape in the context of sound perception, psychoacoustics, spatialisation technologies, and the morphological features associated with motion and shapes. Within these structuring principles we will consider sound and spatial features as the distance and proximity between sounds, between spatial causation, between multiple possible locations and distances of sounds, between size and mass of sounds, and how they move through space. An angular difference, panning, provides a change in location, depth, and distance of a sound object to the perceived space it inhabits. However, this is just one element in spatial perception. I will attempt to show that we can work directly with sound materials *and* with coupled schemes though a focus on the complex meshes of interaction any practice belongs to. To discuss this further we will first investigate other ideas surrounding notions of space and place, to site-specific art practice and the *sound landscape*. After these theoretical discussions, the focus is turned to more technical topics surrounding spatial sound.

3.1 Space – place

From the discussions leading up to this section, we can see that "space" is fundamental. By drawing on a dictionary definition of *space*,[2], it is considered as a "boundless 3D extent in which objects and events occur". In addition, *place*[3] is defined as "the physical environment/surroundings, building or locality, used for a special purpose". The definition of space is a boundless volume that has no clear edges or boundaries: for example, if we come up to a long fence in the country which divides two properties, both sides of the fence belong to the same space but not to the same *place*. Lefebvre considers place to be where everyday life is situated, and where *space* represents "the flows of capital, money, commodities and information ... place is shaped by the grounding (the 'thingification', if you

[2]https://www.merriam-webster.com/dictionary/space
[3]https://www.merriam-webster.com/dictionary/place

will) of these material flows" (Merrifield, 1993, p. 525). *Place* is what is lived in our daily lives.

Space should, however, not be considered to be static and unconnected to time, rather space is dynamic and in motion (Thrift, 2003), it is a "social morphology: it is to lived experience what form itself is to the living organism, and just as intimately bound up with function and structure" (Lefebvre, 1991, p. 94). *Place*, "is involved with embodiment. It is difficult to think of places outside the body" (Thrift, 2003, p. 103). Where space is the fabric that is part of our daily lives and the dynamics that surround us, then place is directly connected to human activity and presence.

Journeying into language and radio signals, Mais Urstad's *Time-tone passages* (2017-2019) mapped the history of the German broadcaster Deutsche Welle to a 40-channel sound installation (Seiffarth, 2017). Here Urstad created a site-specific installation that used material from the archives of broadcasts in thirty-one languages dating back to 1953. Urstad created a portrait of the international broadcaster as a "Stimme der Freiheit" (voice of freedom), rising out of the post-war turbulence. The sound installation was installed in the Welckerpassage in Bonn where all people moving through the passage could be present in the radio broadcaster's history and be connected to its human activity. Still today Deutsche Welle transmits radio signals to the world but only to a few regions with no access to digital technology (Seiffarth, 2017). This connects *Time-tone passages* to this wider fabric of places and human activty.

The term *space* is used frequently in discussions around *spatialisation* and artistic practices. Space can have different meanings dependent on use, and we can talk about *outer space* or *inner space*, but for artistic practices the use of "space" is a useful but generic term to be loaded up with whatever would suit the current moment. In the introduction of this thesis, I mentioned that space is not something "extra", which is added after creation. Rather, space is always present, as a boundless extent, and therefore practices of spatialisation and spatial audio are concerned with the *production of space* and with the *situatedness in space*. Through practices of spatial audio we are creating spaces within the loudspeaker array, but if we only think about space as a product of panned sources, as something technical or as a mere "effect", then

> we fall into the trap of treating space as space "in itself", as space as such. We come to think in terms of spatiality, and so to fetishize space in a way reminiscent of the old fetishism of commodities, where the trap lay in exchange, and the error was to consider "things" in isolation, as "things in themselves". (Lefebvre, 1991, p. 90)

Lefebvre was not concerned with music or sound, his perspectives on the production and practice of space reflects the capitalist modes of production and the importance of everyday life. For Lefevbre, place is synonymous with what is lived in our daily lives; social practice is place-bound. Space is a social dynamic, constantly in motion and in development and is manifested and grounded through place: “Place is more than just lived everyday life. It is the “moment” when the conceived, the perceived and the lived attain a certain “structured coherence”” (Merrifield, 1993, p. 525). If we then, in extension of this, consider space and *spatialisation* as a representation of space, and as a boundless extent that *becomes* place-bound - through our activities, how we situate place and incorporate these notions to practice - we can use it for specific purposes.

An example of the place-bound can be found in *The Waves* by Espen Sommer Eide, where the musical result is both an album and a building - that is, the resulting exhibition at Marres in Maastricht (Sommer Eide, 2019). This is a piece that uses a collective stream-of-consciousness approach inspired by Virginia Woolf’s novel *The Waves*. Here several different musicians with specially constructed instruments move around the building playing melodies, sequences, and rhythms, which later makes the audience move around the building in the musicians shadows and makes the rooms and spaces tangible. This is a music that highlights the building itself as container for possible worlds, unobserved, unoccupied by any particular subject.

Where Michel de Certaeu states that “space is practiced place” (De Certeau, 1988, p. 117), we could flip this to say that place is practised space (Merrifield, 1993). Place is a synthesis of the flowing social dynamic that it is grounded in place. At the start of this chapter, I referenced Tilley and how space cannot exist apart from events and activities in which it is implicated, and extending this “[p]lace is about situatedness in relation to identity and action. In this sense place is context and there can be no non-contextual definition of context or place” (Tilley, 1994, p. 10).

Bound by place as context, Signe Lidén’s 2019 installation *The Tidal Sense*[4] used a large canvas stretched through the intertidal zone outside Ramberg in Lofoten, which acted like a giant 28-meter long microphone membrane. This membrane is tuned to the rhythm of the tide, at high tide the membrane would be partially covered and at low tide it would be manipulated by the wind and weather. The resulting long-term recordings of the membrane voices something about what it means to be part of this particular inter-tidal situation.

It is also possible to sonically allude to two places at the same time, or perhaps to two contexts of the same place. *Beneath The Forest Floor* (1992) was composed based on sounds recorded in a rainforest in Carmanah Valley on Vancouver Island (Westerkamp, 2010). The piece captures the stillness and the motions of wind, water, birds, and insects through the forest. The sudden presence of a chainsaw

[4] https://signeliden.com/?p=1994

within the sound image draws the allusion to two different places in the same musical instant: first, the forest we have been listening to, its stillness, water, and insects; and second, that of the open landscape after clear-cut logging of the trees. When Westerkamp made the recordings, half of the forest was gone due to logging. Where *Beneath The Forest Floor* uses music technologies for a documentary and artistic approach to make listeners aware of ecological treasures grounded in place, the music technologies of recording and distribution can cause music to become detached from place and embedded in new contexts (Cook, 2013), as with the relocating of commercial world music from Third World cultural property to First World capitalism.

This problem with the social and technological mediation of spatial music becomes a sonic essentialism (in reference to the discussions in section 2.1), where space is considered as something "in itself" and the wider social implications are left out. Forgetting that space, and in its extension, place

> is the terrain where basic social practices - consumption, enjoyment, tradition, self-identification, solidarity, social support and social reproduction, etc - are lived out. As a moment of capitalist space, place is where everyday life is *situated*. And as such, place can be taken as *practiced space*. (Merrifield, 1993, p. 522)

These notions of space and place will be discussed further under the heading of the *sound landscape* in the next section. Section 3.3 on site-specificity is a synthesis of the discussions in the previous and the following sections.

3.2 Sound landscape

A "landscape" is normally a reference to the visible features of a land, or an area of land. Something we normally look *at*. When discussing meaning and the perception of meaning in a musical work, Jean-Jacques Nattiez states that "meaning exists when an object is situated in relation to a horizon ... The meaning of an object of any kind is the constellation of interpretants drawn from the lived experience of the sign's user - the 'producer' or 'receiver' - in a given situation" (Nattiez, 1990, pp. 9–10). We ascribe meaning to an object by seeing it in relation to other objects within a larger structure, not only by evaluating several objects next to each other, but also by attributing a certain depth of perspective. In section 1.3.3 we discussed the structure of a sound object as consisting of three parts - a shape, a site/location, and a model in order to tie individual sound objects into a wider mesh of interactions. The "horizon" is a metaphor for the perspective we employ when evaluating different features of objects in relation to each other, to ourselves, and to the surroundings.

Let us look further at the concept of "landscape". A unifying principle with the idea of landscape is "the active engagement of a human subject with the material object. In other words landscape denotes the external world mediated though subjective human experience" (Cosgrove, 1998, p. 13). Landscape is a way of seeing the world. Simply equating landscape with area or region, it does not encompass the possible multiplicities contained in the term. Furthermore, "the painter's use of landscape implies, precisely, observation by an individual" (Cosgrove, 1998, p. 18). The landscape is an *externalisation* (what makes the world and its objects "out there", discussed further in sections 3.4 and 3.6.2) where we are offered a personal control of our perspective of the external world, the symbolic dimensions of landscape and the cultural significations that come with it. Indeed, as Ingold states: "The landscape, by contrast, consists in the first place of surfaces. So at a first approximation, we could say that the question of the relation between weather and landscape is really one about the relation between medium and surfaces" (Ingold, 2005, p. 101). By applying the "morphological method" to landscape, Sauer sought to describe physical and cultural landscapes to identify patterns across landscapes to determine the connections between culture and nature (Sauer, 1925/2008).

Mentioned in the introduction to this chapter, Wishart introduces ideas around the landscape and the sound landscape with references to the "Pastoral" and Manfred symphonies.[5] These pieces are used as examples in that they create analogical relationships between the acoustic events in concert and the events described in the programme note, by evoking images of landscapes and pastoral fields. However, the landscape these pieces belong to, the physical sources of sound, is the concert hall and the human musicians playing instruments (Wishart, 1996, p. 134). By applying the ideas of a landscape to acousmatic music, then "landscape" is redefined to where we *imagine* the sounds to come from. The idea of a *sound landscape* is an attempt at explaining our experience of source recognition in listening situations where we have no other visual cue than loudspeakers.

Wishart recognizes that the destruction of clues as to the sources or origins of sounds in acousmatic music, can be seen as a specific problem in understanding landscape in acousmatic music (Wishart, 1996, p. 139). A sound is often referred to as *abstract* when it is non-representational - when the reference to a source or identity is completely missing. Our perception of the sound cannot be separated from the apparent acoustic space that the sound objects occupy, and the location and size of the sound informs our perception of the supposed size of the acoustic space that the sound objects inhabit. The sound landscape is concerned with the sound images of acousmatic music and their representation, and I have adopted

[5] Beethoven, Symphony no. 6 in F major ("Pastoral"), Op. 68, (1808); Tchaikovsky, Manfred Symphony in B minor, Op. 58 (1817).

this term as an interface to discuss issues related to space/place and site/location as spatial representation.

Within the *visual* field of the landscape, we define the boundary of the visible land and how we organise the objects we see. Likewise, within the *sound* landscape, we define boundaries for what we hear and how we localise what we hear within the bounds of the space around us. The sound landscape is an interface to the discussions on space-place, and site-specific art practices in the next section.

The perception of a *sound landscape* can be broken down into three components. These are interdependent features and can individually or together be subject to analysis–synthesis:

1. The nature of the perceived sounds.
2. the disposition of sound objects within the space.
3. the recognition of individual sound objects. (Wishart, 1996)

These three components of the sound landscape can be tied back to the three-part sound object referred to earlier. The nature of the perceived sound refers to the *shape* of the sound, the dynamic or temporal envelope of a sound and its spectrum; the disposition of the sound objects refer to the site/location-component and its relationship to a site or location; and the recognition of individual sound objects is related to the model and the methods we use for analysis and synthesis.

The resulting *acoustic image*, whether we experience it first-hand or as a recording, creates a large network of semiotic significations and associations, which is rhizomatic in its "semiotic play and culturally specific associations that can be 'messy' to disentangle" (Filimowicz & Stockholm, 2010, p. 6). If we stand at one vantage point and listen to the sounds surrounding us, we listen to both a foreground and a background and we *see* a foreground and a background. We can see our immediate surroundings and we can see the distant horizon and landscape. We can hear (and see) the surroundings where sounds are located, we can localise aurally and visually the directions of the sounds and we can recognise the different sources of the sounds we hear. When listening to the sound landscape, we are dependent on the complexes of interactions between our listening and existing knowledge of the possible landscapes we know. Then, through a synthesis of active listening, memories, and knowledge we make sense of what we hear.

The sound landscape rests on a duality between the real and the imaginary (or virtual) and the possible juxtapositions between them - from the real landscapes of Luc Ferrari's *Presque Rien*-series (Ferrari, 2009), to the imaginary and surreal landscapes of Bernard Parmegiani's *De Natura Sonorum* (Parmegiani, 2010). For example, according to Bregman, frequency proximity, spectral similarity, and correlations of changes in acoustic properties provide certain clues to the listener as to the correct grouping of sonic features in an auditory scene (Bregman, 1994).

We use these clues to make sense of the sound landscape.

However, the landscape must not simply serve as an abstract idea. When landscape is divorced from place it becomes an abstract idea, and as an extension a metaphor. In section 1.3.1 we discussed how we use metaphors, and specifically orientational metaphors, to make sense of spatial orientations in the world and how it influences language. When we use the landscape as a metaphor it is a *container metaphor* (Lakoff & Johnson, 2008), we project an orientation among objects and their surfaces, where we create boundaries and orientations between them. We can move from container to container, out of one container and into another. On a higher level, these are referred to as *ontological metaphors* as "ways of viewing events, activities, emotions, ideas, etc., as entities and substances" (Lakoff & Johnson, 2008, p. 26). This class of metaphor goes beyond orientation and affords understandings of experiences of objects and surfaces - and ultimately shapes. To contextualize this further, let us consider two musical examples.

It can be messy to untangle possible significations in a musical work, and if we look at Jana Winderen's *Spring Bloom In The Marginal Ice Zone* (Winderen, 2018), the listener is placed mainly under water, in the area around Spitsbergen in the Barents Sea. The listener can experience bearded seals, migrating humpback whales and orcas, spawning cod and other species that are attracted to the area during the spring bloom of plankton. Viewed from a purely musical perspective we hear shifting tonal calls, creaks, pulsating and droning waves, along with complex rhythms covering the entire stereo field of the recording in a wide frequency band.[6] But if we consider the *landscape* of *Spring Bloom In The Marginal Ice Zone*, then the site the piece occupies is not only the Barents Sea around the Spitsbergen, but also the ecological conditions which surrounds the global climate emergency. The uses of hydrophone recordings to capture the marine life places the listener firmly into a space that cannot be accessed without specialist equipment and, as with all field recording-based materials, we listen through the perspective of the recordist. The piece ultimately places the listener at the centre of a geopolitical debate between the continuous excavation and erosion of the natural world for profit on the one side, and the need and desire for conservation of vulnerable ecologically important habitats on the other side. The spring bloom is an important source of food for numerous species and is also important for the production of oxygen on the planet, all the while being an aesthetic experience.

A different perspective was taken by Jean-Claude Risset in *Sud* (1985) (Risset, 1988). Based upon location recordings from the south of France, the piece uses recordings taken from a beach and a forest, places where most listeners would have experiences of its sounding nature. The sounds of sea waves, birds, and insects are juxtaposed, layered, and covered over by textural, granular, inharmonic, and

[6] *Spring bloom in the marginal ice zone* also exists as a multi-channel sound installation.

harmonic manipulations of the sound recordings. The play between the real and the synthetic shifts between foreground and background layers to produce hybrids and chimeras. But importantly, it shifts the listeners attention between the landscape of the beach and the forest to the sound landscape of the recording, between a known and an unknown space.

Embedded within the discussion on landscape and sound landscape are the notions of space, place, site, and location. For this we shall refer back to the object discussion in chapter 2, where an object was considered to have three parts: a shape (the overall envelope of the sound, the spectrum), a site/location (the presumed contextual site of the sound), and a model (used for analysis–synthesis of the sound). The spatial structures in acousmatic music are manifold and the design and positioning of spatial features are important to the form of a composition. When Pierre Schaeffer considered musical structure, it was considered temporally, not spatially. The music he advocated went top-down, meaning that the "sound-based music starts with the sound itself and develops into structures by means of definable sonic characteristics" (Landy, 2007, p. 137). This refers not only the to the organisation and arrangement of the sound images temporally but equally to spatial structures. Acousmatic music and its references to real-world scenarios does not just explicitly reference space in terms of spatialisation but also to the site-contingent existence of sounds - the sound landscape.

Schaeffer's focus on the acousmatic was not a means to give primacy to the ear only. The acousmatic refers to hearing a sound and not seeing the cause, which affords us to study what we hear while not being guided by our visual confirmations. If we hear a violin played in the conventional way, there is no doubt that there is a human performer making the sound, however the recording of the violin allows us to listen deeper and past the sound and its coupling with the gestures of the performer. The use of terms such as *evocative* and *mental* images, point to listeners recognizing distinct images, such as photographs, of the external world. These are in no way replicas of something from the external world: "When sensing a spatial environment, an individual builds a cognitive map of space using a combination of sensory information and experiences accumulated over a lifetime" (Blesser & Salter, 2009, p. 46). A cognitive map of space consists of a combination of geometry and knowledge about the external world, and these cognitive maps differ between people due to background and personal bias towards the different sense modalities.

We can relate this back to the discussion in section 2.3, where we looked at Heidegger's ideas on our engagement with the world and how an object's existence it owed to the other things it references and how the same object is again referenced by these things through their complex interactions. Our spatial, cognitive maps are not different, and

> technology, things, science and nature are not "extra" to society - that is they are not elements of an outside reality that causes something to happen in this authenitcally pure social environment (or vice versa). Neither is society itself an embracing container or structuring a priori to which all individual actions may be anchored (and projected toward). A society is rather a complex fabric of intimate relations that link and associate people and things - in short, a collective in which humans and nonhumans co-habitate and collaborate. (Olsen, 2010, p. 138)

In short, the objects we interact with, the objects referenced in this interaction and our navigation of the world is contingent on the container metaphor of the landscape presented previously, as a way we contextually interact with the world.

As has already been alluded to, acousmatic music affords the listener a deep connection to human emotions and relationships precisely because it uses materials that are of direct physical, extensions of the human body into the surrounding world. Spatial cognition is the result of these overlapping sense modalities, not the individual modalities on their own. When we experience the size of something, we relate it to our human body and not to a standardized measurement. If we walk through a doorway, it is not the measured size of the door we are interested in but how our bodies fit *through* the doorway. This shape-relation is pertinent when we assess the perceived size of a space and the perceived size of the sounds (and their implied sources) when we listen to music.

By extending this *outward*, based on a listening experience of the Orbieu soundscape in the south of France, Denis Smalley presents a tour de force of spatial taxonomies, as, among others, arrays of nested *zoned spaces* (Smalley, 2007). Among these, Smalley defines a *perspectival space* "of the acousmatic image as the relations of position, movement and scale among spectromorphologies, viewed from the listener's vantage point" (Smalley, 2007, p. 48), with a reference to visual art and the perception of positions and distances of three-dimensional shapes on a two-dimensional surface. The effectiveness of spatial conceptions and orientations depend on the "vantage point" of the listener, and the sound image refers to "seeing the invisible", of how the acousmatic sound can evoke visual images in the listeners, depending on the contextual relationships.

The networks and links between people, histories, technologies, and ideas are potentially infinite. All these objects "are embedded in even remoter references" (Harman, 2002, p. 32) allowing us to further analyse and synthesize the objects we perceive. The occupancy of sounds in space is complex, and the given size of the object is deeply integrated into the landscape that we perceive the object to belong to. We can find carefully constructed spaces formed by the sound material of "tiny sounds building up an impression of space, then you can begin to imply space within which these sounds should live" (Otondo, 2007, p. 13). We can use the sounds, with their possible references to spaces, to design a sound landscape or we can attempt to model an existing spaces as in Fernando Lopez-Lescano's *Space S[acred|ecular]* (2014/2016),[7] which sets out to create a sound landscape by modelling the acoustics of the Hagia Sophia in Istanbul, Turkey.

This section has presented a short discussion on the term sound landscape, and how it has been adopted for use here. Wishart's reference to the *imagined* source of sounds in acousmatic listening is a starting point for the continued discussions in the next sections where issues of the site-specific and of spatial composition will be discussed.

3.3 Location – site

Localizing a sound is, given the previous discussions, related to the problem of localizing the *source* of the sound (Chion, 1994, p. 69), as the sound both exists *in* and embodies a place. In earlier discussions, the sound object has been presented as consisting of three components, the shape of the sound, the (assumed or apparent) site/location the sound belongs or refers to, and the model we use to analyse and synthesize the sound. The sound landscape is the imagined source of the sound, a container in which we create *externalized* sound experiences, which explore the complex interactions between sounds and space. To pursue this further, I want to incorporate questions of site-specificity from the visual arts into the understandings of the sound landscape.

The term *site-specific* is often considered literally, where an artwork is installed or made for a specific space purely by the merits of "fitting into" in the space. This has been rather uncritically adopted as a genre category by institutions and discourses (Kwon, 2002), and this perspective on the site-specific makes no effort or attempt to negotiate or interrogate the space itself, or the situatedness of the work in question. The site is a given and the work is made to exist within the physical confines afforded by the walls, floor, and ceiling. The term is most often associated with the visual arts and it is within the visual arts that the term arguably is the most sophisticated (Kim-Cohen, 2016b). Works of site-specific art have been produced for a specific site and this is also the site where the work is experienced. In 1971 the French artist Daniel Buren argued that artists should abandon the studio to work site-specifically so that the art work would not be separate from the site of reception (Bishop, 2005). Within this discourse, the site is physical and "on site" refers to the place where the work is to be experienced. If a work is removed from this site, it would collapse and reconstructing the work in

[7] https://ccrma.stanford.edu/~nando/music/space_sacred_secular/

a different space could never function.

From the gallery arts we have both an *installed* exhibition, how objects (e.g. paintings, sculptures, photographs etc) are presented to an audience in an exhibition space, and an *installation*, which addresses and encompasses the viewer as part of the space:

> What both terms have in common is a desire to heighten the viewer's awareness of how objects are positioned (installed) in a space, and of our bodily response to this. However, there are also important differences. An installation of art is secondary in importance to the individual works it contains, while in a work of installation art, the space, and the ensemble of elements within it, are regarded in their entirety as a singular entity. Installation art creates a situation into which the viewer physically enters, and insists that you regard this as a singular totality. (Bishop, 2005, p. 6)

An installation displays multidimensional features in its interactions with the artwork itself, the spectators, and the space, but most importantly we can explore morphological features of artworks by the changes of our vantage points and our orientations in the space, in ways that often are difficult in concert settings. The practices of installation art Claire Bishop refers to are usually confined to gallery or museum spaces, where the notions of site-specificity opens up a wider discourse. With the site-specific, the idea of a "neutral" space is challenged, from the white cube in the visual arts, the black box for theatre to the concert hall for music. Here an engagement with site is confined to the "neutral" and local, whereas site-specific practices look outside these confines.

Here we will pick up a thread from the introduction, from the discussion on sound art and on the situatedness of sound works. Brandon LaBelle cites, like Alan Licht, *I Am Sitting In A Room* (Lucier, 1981) as an example of sound art, which is heard to *reveal* sound as site-specific (LaBelle, 2019). The reasons for deeming this work site-specific is its integration into, dependence upon, and revealing effects of the space that it is situated in. This is still only site-specific in the most literal interpretation of the term. The site of *I Am Sitting In A Room* is a given (where ever it is performed) and does not enter into any form of negotiation with its site. The score specifies: "Choose a room the musical qualities of which you would like to evoke" (Lucier, 1995, p. 312), and the evocation of the room, that is the audible changes to the recorded phrase, is due to the resonant frequencies of the space given the relative placement of the loudspeaker in relation to the floor, walls and ceiling.[8] The site and the space is clearly integral to the work but it is in no way specific to those sites or spaces.

[8]The spaces which this piece has been recorded in was discussed in the album notes: "Lucier's first recording of the piece was made in the fall of 1969 in the Electronic Music Studio at

The term site-specific was introduced and promoted by the American installation artist Robert Irwin. His practice revolved around architectural interventions that alter the physical, sensory, and temporal experience of space (Bishop, 2005). Defined from the perspective of the visual arts, he proposed four categories which a site-specific work might fall into:

1. Site dominant: A work which falls into the monumental category, of permanence and transcendent historical content.
2. Site adjusted: A work which is made elsewhere (the studio) and then transported to and installed at the site.
3. Site specific: A work which is made for a particular space, where the work is integrated into the surroundings.
4. Site conditioned/determined: All cues for the work are drawn from the site and the surroundings, order, uses, distances, senses of scale, organisation, systems of order. (Irwin, 1996)

Irwin's practice deals with phenomenological perception, where he sought "embodied perception over intellection" (Bishop, 2005, p. 57). Through this approach, he wanted to open the viewer's eyes to the world as it already existed. Irwin's notes on the site-specific from 1985 was greatly expanded by Kwon, where Irwin's categories are collapsed into one of "phenomenological site-specificity".

In her book *One Place after another: Site-specific art and locational identity*, Miwon Kwon has offered one of the most comprehensive discussions, genealogies, and overviews of site-specificity in the visual arts, which she divides into three categories (Kwon, 2002):

1. Phenomenological - which responds to the physical realities of the space in which the work is encountered.
2. Institutional - goes beyond the space itself considers the agency of the place where the work is experienced (in Kwon's terms this is the gallery or museum, but it could relate to any space).
3. Discursive - goes beyond the institution and looks at site as a product of

Brandeis University during his last days teaching there; it was, in his own words, 'harsh, strident'. The second attempt, which became the definitive version for concert use for the next decade, was recorded on March 10th, 1970 in a small rented apartment at 454 High Street in Middletown, Connecticut, where he had just accepted a faculty position at Wesleyan University; it was 'beautiful'. The recording on this CD was created on October 29th and 31st, 1980 in the living room of Lucier's house at 7 Miles Avenue, Middletown, where he had lived for ten years. *I Am Sitting In A Room* is inextricably linked to notions of 'home' – of a room rather than a concert hall, of sitting rather than laboring, of speaking rather than singing, of literally being in the right place at the right time". (Lucier, 1981) intersecting narratives, debates and practices.

The site is not merely the location where a work exists nor is it only the product of discursive relationships. The site is something that both makes and masks the location and the discursive nature of the work, and so all three of the categories fold into each other as overlapping matrices of reference, again as fabrics of citations (Kim-Cohen, 2016b). When these categories fold into each other, they negate divisions and rather assumes different positions along several axes.

The musical examples cited so far often have a strong component of field recording as a material basis for transformations and composition. Field recording happens in response to a specific site, and (often) to significations with the places recorded. If sound is recorded in one place, treated in the studio and re-introduced to this same place, then there is a transposition in time where the listeners meets a "historical" version of the same place, for example in the record *In St Cuthbert's Time*, which is called "A 7th Century Soundscape of Lindisfarne" (Watson, 2013). The record explores the region around Lindisfarne in an attempt to reflect the daily and seasonal aspects of the ambient sounds that would have accompanied life in the late 7th and early 8th century but presented today in an attempt for the listener to listen back in time.

We can view the earlier example of Alvin Lucier in the light of Jonathan Harvey's *Mortuos Plango, Vivos Voco* (1980) (Harvey, 1999). Where Lucier's seminal piece responds to the resonant frequencies of a room by playing and re-recording his own voice multiple times, Harvey's *Mortuos Plango* used the tenor bell from Winchester Cathedral and the voice of his son (Harvey, 1981) as raw material. This locates the piece within the physical and conceptual confines of the cathedral and draws upon the sound, effect, and memories of the bell, the boy soprano, and thus extending the composition far outside of the confines of the cathedral. As such the bell is a *soundmark*, the auditory counterpart of a landmark, and a religious *earcon*, the auditory counterpart of an icon, for those who attend religious service in that particular place (Blesser & Salter, 2009). Where *I Am Sitting In A Room* could be located on Kwon's *phenomenological* site-specificity axis, then *Mortuos Plango, Vivos Voco* would be on the *discursive* site-specificity axis.

In discussions of the work by the band Billy Bao[9] in *No depth: A call for shallow listening*, Seth Kim-Cohen traces the origins and existence of the record *The Lagos Sessions* (2015) directly to a site-specific reading of the work as a response to and an origin within the capital city of Nigeria, Lagos. In Kim-Cohen's reading of the music, the record does not respond to any *one* site within the city of Lagos or the surrounding areas. Rather the record responds to multiple geographical and

[9]The band consists of Nigerian expatriate Billy Bao and Basque musicians Mattin and Xabier Erkizia.

historical locations as in:

> Lagos's relation to its site as a center of the slave trade cannot have been generated entirely within the site of Lagos but is largely generated in Britain and America. Additionally, a site may have been generated locally, exported, modified and reimported as a new, recombinant site. One might think of Nigerian music traveling to America, only to mutate and return to Lagos as James Brown's funk before becoming the site of Fela's afrobeat. In shallow listening, there should be no confusion: What we are hearing is not sound-in-itself, nor the sound-of-Lagos-in-itself, but the sound of an intention to represent the fabric of citations that constitute Lagos. (Kim-Cohen, 2016a, p. 135)

Given this analysis of *The Lagos Sessions*, we can draw out this idea of the fabric of citations that follow the works we experience. A work always responds to more than just itself and the space it occupies. At the outer limits, a work always responds to the site of history, its traditions, its canon, and the context from which the work both exists and grows. As such it is a transposition, changing its identity across multiple instances of knowledge. Acousmatic music responds both to its own materials and technologies and to the histories of practice contained within the field, and the subsequent fields it references. Viewed in this perspective, there is no room for a "purity of sound", nor to the "hegemony of the visual" or the desires for immersion. All these experiences respond to a the wide and complex interactions of the meshes between people, histories, technologies, and ideas. In discussions about sites, places, and locations we often encounter a discourse to be literal, yet the preceding readings of *The Lagos Sessions* show that we respond to these types of sites in many other ways than just in the literal.

The installation *Superimposed Landscape (Lista)* discussed in case study 3, in section 4.6, responds to Irwin's fourth category as site-conditioned/determined, and to Kwon's first category as phenomenological site-specificity. The installation uses an algorithmically determined sampler and only uses sound materials gathered and live-sampled from the immediate location around the "gallery" (a bunker from the Second World War). The installation responds to the physical and changing realities of the site and all cues in the realisation is determined by the surroundings.

Natasha Barrett's *Trade Winds* (Barrett, 2007) is inspired by the sea. The piece uses recordings from under and over water, in harbours, shores, and open ocean across the world and takes its site between nature and culture as a cyclic narrative, with the central *Mobilis in mobili* drawn from the narrative worlds of Jules Verne's *Twenty Thousand Leagues under the Sea*. Like *The Lagos Sessions*, it references multiple geographical locations and a historical context within a 100-year-old sailing ship. This discursive site-specificity treats the ocean and our experiences with it as objects within a complex mesh of interactions.

Exploring the depths of the oceans in a different manner than both *Spring bloom in the marginal ice zone* and *Trade Winds*, we find Lionel Marhcetti's *Planktos 2015-2020*. Totalling an epic 3.5 hours, this piece was inspired by the poem "Planktos" by Régis Poulet[10] and brings the listener on a musique concrete journey to explore the depths of the oceans and its lurking mysteries. This journey starts out in the microscopic, the first movement is entitled "Balbutiements des bactéries", the "first faltering steps of bacteria", before it describes marine life like whales and jelly fish, and ends in a 45-minute movement describing the ocean itself. Given the inspiration from the poem where some lines read: "let's dive deeper/ let's dive to outrun/ the horrific times of the whale hunting/ industry and all its massacres", this long meditation on the ocean, its inhabitants and mysteries places the work in an axial relationship between the phenomenological and the discursive site.

Predating *The Lagos Sessions* by 95 years, Varèse's orchestral work *Ameriques* (1918-1920) is a work filled with noises, movements, and a central role written for an industrial siren. This was written after Varèse came to New York and experienced the sounds of the city with its cars and dense population. The work responds to the experience of a modern city and the use of the siren not only binds the piece to evocations of the modern industrial age but also resonates with anyone who has experienced air-raid sirens during wartime. The presence of this siren in the work binds the work to the fabric of modernity and raises it out of the site of reception in the concert hall and into a larger site of the experience of the city, industry, war, and history.

The sounds of the modern age was also the sounds of the machines, not just of industry and war, expressed through the mechanical productions brought on by the industrial revolution, the steam engine, and the car. This noise revolution was celebrated by the Futurists, and notably through the manifesto *The Art of Noises* by Luigi Russolo and his work on the development of the *Intonarumori*, noise-intoners, hand-cranked instruments meant to harness the sounds of the modern age (Russolo, 1913). This is particularly evident through the composition *Risveglio di una città* (Awakening of a city) (1913),[11] which was a call to awaken our ears to the dawn of a new music.

In the previous chapter I made the claim that objects are placeholders, containers of conceptual structures; as we saw earlier in this chapter, the landscape is a container metaphor, bounded by ontological metaphors, which are "ways of view-

[10] https://grmnetwork.blogspot.com/2020/05/lionel-marchetti.html
[11] https://ubu.com/media/sound/russolo_luigi/die_kunst/Russolo-Luigi_01_Risveglio.mp3

ing events, activities, emotions, ideas, etc., as entities and substances" (Lakoff & Johnson, 2008, p. 26). I will contend that the previous examples presented confirms that the three-part conceptual structure of a sound object links the intrinsic and extrinsic features of the object to structural spatial features and contexts. The discursive site-specificity of *The Lagos Sessions* treats the city of Lagos as an object, and as an object Lagos is a part of a larger fabric of citations that allows us to see the city not just as a physical site but also as the present socio-political and socio-historical conditions that define the site.

These different perspectives of the site-specific also connect to lines of thought which substitute "site" for "field", as a means of discussing field guides. By citing Susanne Ewing on discussing the role of site in architecture, Shannon Mattern draws out a contrast on "the *site* ('physically delimited, culturally- and historically-situated') with the *field* (a 'cloud-like set of social, cultural, economic, and nonhierarchically networked conditions of reality')" (Mattern, 2016, np). The arguments around the substitution of site for field is further contextualized in the extended uses of the *field guide* and that of *fieldwork*.

Two groups of artist-as-researchers-as-ethnographers, Unknown Fields Division and Dark Ecology,[12] conduct projects in the far north and south that explore clandestine and enigmatic landscapes as some forms of documentation of their field work:

> Both groups cultivate an epic, enigmatic aura around their work through the use of atmospheric post-rock or noise soundtracks, close-ups of heavy machinery and scarred landscapes, and low-angle shots that aggrandize the bravery and skill of valiant fieldworkers, their faces weathered through prolonged exposure to the elements. Unknown Fields mixes in blurs and glitches and strategically washed-out or grayscaled imagery, suggesting the precarity and ruggedness of these terrains (and the lengths to which the team had to go to document them). And both groups juxtapose gothic or sublime landscapes with shots of technical instruments and interfaces, or they animate and annotate the landscape with data visualizations. The overall impression is of a dark, unknowable field that resists de-clouding but is ultimately aestheticized, "fastened down," made knowable through technical mastery. (Mattern, 2016, np)

There is a fantasy associated with this type of practice, where the artist can assume some sort of moral high ground of the oppressed and absolve themselves of the guilty conscience of privilege (Kim-Cohen, 2016b). The work that some of these people do is glamorized through a macho-aesthetic valorisation of field

[12] http://www.unknownfieldsdivision.com/ and https://www.darkecology.net/about

work, a fetishisation of Indiana Jones-like transgressions and neo-colonialism. The broader conceptualisation and awareness of the social and political implications of their practice is not really addressed and the *medium* of the performance, trekking to the ends of the earth, is enough of a justification of the project. Unlike *The Lagos Sessions*, these fieldwork projects fail to recognize the political, historical, social, philosophical, and economic relationships of the projects and whatever cultural critique might have been part of the visions of the organisers is lost in a colonial narrative of the priviledged few.

3.3.1 Sites and non–sites

Taken as an inspiration for the notion of a *transposition* discussed in section 1.2, Michael Schwab draws on Robert Smithson's concept of the *non-site*. The transposition is a means of articulating the movement of research within art, to show how something changes its identity when moving across different instances. With *non-site*, Smithson developed a series of geologically and geographically based works that began in 1968, which consisted of photographs, rocks, sand, soil, maps, and containers, presented in a gallery. The objects became representations of the site from where they were taken, becoming a non-site and legible as sculpture when separated from the natural context from where they existed. Through photographs, maps, and charts the audience can experience the site where these objects originated, and if desired can seek out the sites themselves. The non-site is a decoupling of the object from its origin: the rocks from the work *Non-Site, Franklin, New Jersey* (1968) are separated from the site from where they belong and brought into a gallery where they are presented along with a photograph that details their original sites, and are representations of the site of origin.

The non-site is a three-dimensional picture that resembles the site but has abstracted the site. It is "a 'logical picture' that differs from a natural or realistic picture in that it rarely looks like the thing it stands for. It is a two dimensional *analogy* or *metaphor* - A is Z" (Smithson, 1996, p. 364). The map, diagram or street plan is always a representation of the reality of the site it depicts, "maps are selective representations of reality" (Black, 1997, p. 11), and in Smithson's work on the non-site the morphogenesis (see section 2.4) between the pattern formations of the terrain and the representations of shape opens up a "new sense of metaphor" (Smithson, 1996, p. 175).

When removing solid objects, such as sand and rock, from one site and presenting it elsewhere, we are faced with the object itself and cannot see its origins or surrounding context. There is no way to know the rest of the terrain from where the sample was taken apart from the photographs and maps chosen by the artist to represent the terrain. In the practice of soundscape, this dislocation, decoupling, of a sound from its origins, is what Schafer referred to as *schizophonia* (Schafer, 1994).

3.4 Spatial attributes

Where the previous sections have discussed music, sound, and spatial practice broadly from a practice point of view, the remaining three sections highlight the features of sounds through contextual presentations using a focus on practical topics of sound localisation and psychoacoustics, before moving on to discussions on sound spatialisation technologies and the ambisonic soundfield. These topics will be important for the discussions in chapter 4, on both the typomorphology and perspectives on spatial authoring through spatial audio.

Composition strategies for spatial sound have previously been provided by Barrett, as four approaches to space and how the composer can work with spatial audio.

1. Illusion of a space or spatial location of an object.
2. Allusion to a space or spatial location of an object.
3. Simulation of the three-dimensional soundfield.
4. Spatial possibilities contingent upon temporal development. (Barrett, 2002, p. 314)

The first approach, is for the composer to create a space that is closely matched with a real space. Secondly, space is implied by the listener's assumption of certain sounds; for example, wind in the trees implies an outdoor scene. Third, the construction and synthesis of a complete soundfield as a descriptor of place. Fourth, our perception of the soundfield and the unfolding events are reflected over time.

A terminology for describing spatial attributes in electroacoustic music has been provided by Kendall (Kendall & Ardila, 2008), based on Rumsey's terminology for reproduced sound (Rumsey, 2002):

1. Dimensional attributes:
   - width
   - distance
   - depth
   - direction.
2. Immersive attributes:
   - envelopment

- presence.

These elements can be combined in different ways within an sound image, from an individual source to clusters, or groups, of sources (Kendall & Ardila, 2008, pp. 128–130). The "scene" adopted from Rumsey's model of source-ensemble-room-scene is not unproblematic is this context, given its basis in the evaluation of acoustic music in traditional concert hall settings. Yet, we must be mindful of the separation between an acoustic source signal and the source's perceived image and defines four frames of reference for "source":

1. Source signal, the representation of the acoustic signal.
2. Source image, the source that has spatial attributes in the (auditory) scene.
3. Conceptual source, the object the listener identifies with the sound independent of spatial attributes.
4. The listeners spatial schema. (Kendall & Ardila, 2008, pp. 131–132)

When discussing the spatial attributes of sounds and sound perceptions for multichannel audio, I will propose that we discuss real-world experiences, sound perception, and localisation, abstraction of objects, relationships between objects, and the perception of space through the mass and size of objects as:

1. Perception of sound as a whole, through object cognition and smearing in time and space.
2. Immersion in the sound, the perception of not only the listening space but also the inherent spatiality of the sounds and its external references.
3. The perception of multiple locations and distances and the proximities between sounds. This is essential for the understanding of the relationships between sounds in a space.
4. The perception of space through size and mass of objects.

By "real-world" I mean that which can be sensed from our surrounding world, either directly through our biological sensory apparatus or through microphones, sensors or other data collection methods. Through "sounds as a whole", we gather some form of impression of the supposed origin of the sound, regardless of its origin. This can be especially misleading in cases where we used synthesised sounds. Once we work with (any) sounds as our material, then the origins of the sounds are no longer of any concern, yet the supposed origin of a sound still becomes important in our conceptual perception of the sound.

As we discussed earlier, Bregman notes that frequency proximity, spectral similarity, and correlations of changes in acoustic properties provide certain clues to the listener as to the correct grouping of sonic features in an auditory scene (Bregman, 1994). In complex soundfields, listeners can still identify and segregate sources into discrete events and from different directions. These involuntary perceptual groupings of sensory input through the principle of proximity dictates that stimuli occurring, either spatially or temporally, close together are perceptually fused, and we experience a smearing of sounds where we perceive multiple sounds to be one. Sounds that have similar spectral content will be perceptually grouped, through the principle of similarity. Good continuation states that sounds with similar amplitudes or spectral envelopes will be grouped and perceived as a single source.

3.4.1 Spatial hearing and psychoacoustics

Spatial hearing refers the study of perceptual cues and the mechanisms of human sound localisation. At its simplest, our hearing serves a very primitive function as a warning system (Moore, 2003), so that we can quickly identify the location of a sound and make judgements whether the source of the sound represents a danger to us or not. A majority of the events we attune to are in the horizontal plane and we localise sound better horizontally than vertically, likewise we localize sound better to the front than to the back. Localization vertically is influenced by spectral differences due to the shape of our ears (Blauert, 1997). Primarily, our abilities to localise is to shift our visual attention to whatever direction we perceive a source. The uses of spatial parameters to exploit and explore human sound localization through compositional activity is dependent on the available technology, how these attributes are explored is covered in the following section 3.5.

Sound perception can be divided into two categories: (1) free field conditions, and (2) enclosed space conditions. A free field condition can be an open field, with no buildings, mountains, or trees (Howard & Angus, 2009), and because of this there are no reflections of the sound emitted from the source. Anechoic chambers (echo-free rooms) attempt to model free field conditions. Enclosed spaces are what we most frequently experience, and in such spaces reflective surfaces are where a portion of the emitted energy from the source is reflected back from the surrounding surfaces (Rumsey, 2001). In room acoustics, Griesinger has found that the visual impression of a concert hall "often dominates the aural impression" (Griesinger, 1997, p. 721).

If listeners are unfamiliar with the sources and the space, localisation can be difficult but the auditory system quickly becomes familiar with both the sound sources and the room conditions (Plenge, 1974). In the practice of stereo diffusion, discussed in first pages of the introduction, a composer can use the loudspeaker orchestra to enhance spatial motion, contrasts, and articulations contained in a work and can draw on the multitude of psychoacoustic criteria in the performance of a work. For example, a composer can vary the apparent source width for individual sounds and with this the spatial impression of the performance space.

When we hear a sound emitted in an enclosed room, we hear a single fused sound that consists of the direct sound as emitted by the source along with a series of reflections from different surfaces in the room. The ratio of direct-to-reverberant sound is important in distance perception of a sound source, as well as early reflections, high-frequency attenuation, and air-absorption (Moore, 2003). An important aspect of spatial hearing is the difference in time of arrival at the two ears by the wavefront emitted by a source, first described by Lord Rayleigh as the *duplex theory* (Rayleigh, 1909). A sound emitted by a source that is located on the median plane will arrive at the ears at the same time, but moving the source off to one side (towards one ear) will cause the wavefront to arrive differently at the ears, representing the *interaural time delay*. Sound level differences between the ears are called *interaural level difference* or *interaural intensity difference* (Moore, 2003). When a sound travels towards the head from the left, it hits the left ear first and then is diffracted around the head to the right ear. The amplitude of the sound will be less in the right ear than in the left ear, because of both the obscuring effect of the head and the extra distance travelled. The shape of the head and the difference between the two ears produce differences in frequency response, called *head-related transfer function*. The way we can move our heads to change the interaural time delay and interaural level difference along with the differences in head-related transfer function between the two ears is another important aspect of spatial hearing (for more in-depth discussions on these different attributes see (Blauert, 1997; Howard & Angus, 2009; Moore, 2003)).

Dimensional features in spatial sound are impressions in terms of *spatial extent* (width, depth and height), *distance* and *direction* (Roads, 2015), and immersive features such as *presence*, *envelopment*, and *engulfment*. In their normal usage, these attributes describe spatial and musical perceptions, and how the human mind makes sense of these experiences. However, these attributes can provide us with insights into the situatedness in the practices of spatial audio through the sound landscape, specifically, as we will see later, how the identification, classification, and description of sounds can be made through the typormophological framework.

The sound pressure of a wavefront decreases with distance, and the main perceptual cue to indicate the distance of a sound in free field conditions is the source's amplitude (Blauert, 1997). Distance perception appears more effective when the listener perceives a number of sounds and can compare the signal level between the sources (Mershon & King, 1975). As sound travel longer distances the higher frequencies are attenuated through air absorption (Moore, 2003).

Presence is defined by Rumsey as "the sense of being inside an (enclosed) space" (Rumsey, 2002, p. 662), meaning that the listener can sense the boundaries of the surrounding space and feels present in the space. An important criterion for presence is the subject's awareness of the auditory background. Others have discussed *spatial presence* as a replacement for immersion and is regarded as an experiential state in film sound (Mera, 2016).

Envelopment is a subjective attribute of the enveloping nature of the sound. Rumsey distinguishes between *environmental* envelopment (sense of being enveloped by reverberant or environmental sound) and *source-related* envelopment (being enveloped by single or groups of sources) (Rumsey, 2002).

Engulfment is proposed as a unique 3D audio attribute of the listeners being "covered over" by sound (Sazdov et al., 2007), rather than being surrounded by sound.

Externalisation is related to spaciousness in how the sound is perceived to be outside the head rather than constrained to a region close to the head or inside it (Rumsey, 2001).

Apparent source width relates to how much space a source appears to occupy. In smaller rooms apparent source width has little relevance because it is difficult to distinguish if the sound image is wide or just diffuse (Rumsey, 2001). Apparent source width is closely related to Interaural Cross Correlation Coefficient (Ando, 1985), which measures how closely the two ear signals are correlated. The Interaural Cross Correlation Coefficient is used to determine the spaciousness and envelopment of concert halls, where a value close to zero will produce a diffuse and spatially large sound image and a value close to one will produce a narrow sound image (Potard & Burnett, 2004).

Spatial impression is used to describe whether a space is perceived to be large or small, and *spaciousness* describes to what degree we perceive we are in a large and enveloping space (Griesinger, 1999). The terms spaciousness, spatial impression, and envelopment are interpreted variably in the literature and spatial impression has often been used as a 'cover all' term (Rumsey, 2002). Several researchers equate spaciousness with *apparent source width* (Griesinger, 1997), but spaciousness has no bearing on the perceived size of the source, "a concert hall can be spacious, the reverberation of an oboe can be spacious, but the sonic image of an oboe cannot be spacious" (Griesinger, 1997, p. 721). The perceived spatial impression is dependent on lateral reflections between 125Hz and 1000Hz and is a function of the performing level and will be higher with larger ensembles (Barron & Marshall, 1981). The combination of early and late arriving energy determines the magnitudes of spatial impression, apparent source width, and listener envelopment (Bradley et al., 2000), and if reflected energy arrives within 50ms of the end of the sound event, the spatial impression of a small room is created (Griesinger, 1996). However, to explain spatial impression, both the frequency and level-dependent aspects of the music that arrives at the listeners ears have to be linked (Lokki & Pätynen, 2020).

These psychoacoustic criteria describe our perceptions of sound and space, as defined through a wealth of literature. These criteria provide us with an interface to evaluate spatial settings and experiences of music in the concert hall. These criteria also afford us to evaluate the sound landscape, which has been discussed at length in this chapter. The landscape makes our perception of the world "out there" (Cosgrove, 1998) and is a dynamic and motion filled social morphology, this makes it part of what is lived in our daily lives. This creates a situatedness in the practices of spatial audio which makes this into a production of space and a contributing factor to the contextualisation of space through subjective human experience.

3.4.2 Time

Music exists on multiple timescales, and sound can be spatialised over multiple timescales. Contrasting relationships of foreground and background sounds can create timescale phrasings, as in Karlheinz Stockhausen's *Kontakte* (1960). Varying topographies and spatial depths can create juxtapositions and different perspectives in the perception of individual sound objects and the relationships between them. Dennis Gabor suggested that all sounds could be decomposed into functions of time and frequency, called *acoustical quanta* (Gabor, 1947). Based on these theories, Iannis Xenakis composed and synthesized elementary sonic *grains*, tapping into the world of microsound (Xenakis, 1992).

Considering music that exists on extreme timescales, we can look at John Cage's ORGAN2/ASLSP (As SLow aS Possible),[13] the performance of which began in a church in Halberstadt in Germany in 2001 and will last for 639 years. The *Dream House* by La Monte Young and Marian Zazeela, ran as a continuous installation of sound and light between 1966 and 1970. Karlheinz Stockhausen's opera *Licht* spanned seven days and nights. In *9 Beet Stretch* (2002) by sound artist/composer Leif Inge, Beethoven's 9th symphony has been stretched to a duration of 24 hours with no pitch shift or distortion.[14]

Considering different timescales both in terms of human attention spans, musical duration, and the interactions between structures, Curtis Roads has proposed a temporal hierarchy of timescales, divided into nine sections (Roads, 2004, pp. 3–6). A shorter list was introduced by Trevor Wishart, where time-frames are divided into five different sections (Wishart, 1994, pp. 16–19). Of particular relevance is also the figure from (Roads, 2004, p. 5) that shows the segmented time domains.

A sound object exists devoid and separated from the sound source, but it is not timeless. If we listen to it once, then we do perceive and understand something

[13] https://universes.art/en/specials/john-cage-organ-project-halberstadt
[14] http://www.9beetstretch.com/

Table 3.1: Timescales of music as defined by Roads and Wishart, from the infinite to the infinitesimal.

Roads	**Wishart**	**Description**
Infinite	-	Idealised realm of Fourier series
Supra	-	Months, years, decades, centuries
Macro	-	Musical architecture, minutes or hours
Meso	Phrase	Divisions of form, minutes or seconds
Sound object	Continuation	Basic unit of musical structure
Micro	Grain	Sound particles, to the threshold of perception
-	Wavecycle	Single wavelength of a sound, shape made by samples
Sample	Sample	Atomic level of digital audio
Subsample	-	Fluctuations down to one billionth of a second
Infinitesimal	-	Idealised delta functions

of the shape of the sound. When we listen repeatedly our understanding of the sound changes over time and we become more and more aware of the shapes and surfaces of the sound and its relationship to the temporal event that caused the sound. The temporal event, though, is not the same as the source of the sound.

From Table 3.1 we can draw out the timescales that are the most relevant for the description of music, where somewhere between *wavecycle* and *micro/grain* a sound becomes audible. The boundaries between the timescales are perceptual, and where some of them can be psychoacoustically defined, others are more culturally defined. Through the sound object as a basic unit of musical structure, we can create sound masses such as "[O]ne type of sound mass is a cluster of sustained frequencies that fuse into a solid block. In a certain style of sound mass composition, musical development unfolds as individual lines are added to or removed from this cluster" (Roads, 2004, p. 15). These practices display the boundaries between the *sound object* and *meso* timescales.

Likewise, Wishart also pays attention to the division between perceptual boundaries:

> Just as in traditional musical practice, the boundary between a long articulation and a short phrase is not easy to draw. This is because we are no longer dealing with clear cut *perceptual* boundaries, but questions of the interpretation of our experience. A trill, without variation, lasting over four bars may be regarded as a note-articulation (an example of continuation) and may exceed in length a melodic phrase.

> But a trill with a marked series of loudness and speed changes might well function as a musical phrase (depending on context). (Wishart, 1994, p. 19)

From these divisions of timescales, we are identifying four terms that are relevant to the concept of musical and spatial form, shown in Table 3.2.

Table 3.2: Simplified timescales of music, based on the previous table, suitable for the articulation of musical and spatial form.

Scale	**Description**
Macro	Musical architecture
Meso	Groupings of objects into phrases and gestures
Sound object	Basic unit of structure
Micro	Sound particles and grains

The different timescales are usually superimposed as different layers, depending on the compositional intent (for detailed discussions on the time scales ranging from macro to micro, see (Roads, 2004, pp. 11–28)). The important aspects of these interrelationships of timescales are that they afford us to consider the individual fragments of sounds, along with the sound as a whole to the sound in spatial context. These issues are discussed further in chapter 4.

3.4.3 Montage

Derived from cinema, montage refers to processes of cutting, splicing, and other film editing operations to assemble a sequence of shots to create a film, similar to the tape-montages before the computer in *musique concréte*, acousmatic, and electroacoustic music. Assembling a composition from many different sound objects and parts on a computer is also a montage. In reference to the work of Horacio Vaggione, Roads describes "micromontage" as the process of placing individual sound particles on "the canvas of time" (Roads, 2005, p. 299), and uses the example of the pointillist painter Georges Seurat and his technique of breaking down a scene into thousands of small dots and brushstrokes of colour. A strong aesthetic sensibility combined with mathematical analysis created paintings with dense seas of brushstrokes and points.

However, it is not only in respects to granular synthesis that we can consider the work of Seurat to be of importance. In the montage of thousands of individual dots and brushstrokes to create a holistic visual scene, as for example in *A Sunday Afternoon on the Island of La Grande Jatte* (1884-1886),[15] the painting style allows the viewer to blend the colours optically, rather than having the

[15] https://www.artic.edu/artworks/27992/a-sunday-on-la-grande-jatte-1884

colours physically blended on the canvas. The approach and perspective is also important in terms of *spatial montage*, and the assemblage of a soundfield. We do not encounter individual, discrete point-sources in real-world sounding contexts, rather we should construct and develop holistic spatial scenes that can contain dozens to hundreds of individual sound objects. This is the basis for the thinking around the sound landscape, it is not only the *imagined* source or location of the sounds but also the design and montage of the sound scene, the landscape, that is important. This moves from a micro scale, of sound particles organised in time and space to a macro scale of the structural relationship between the sounds and the designed spaces.

3.5 Spatialisation approaches

Pierre Schaeffer and Pierre Henry's experiments with spatially distributed sound for the performance of *Symphonie Pour Un Homme Seul* (1949/50) was based on mono sound. They used three horizontal and one elevated speaker, each with a dedicated tape track and a specially designed *potentiometre d'espace* by Jacques Poullin was used to move sound between the four speakers (Palombini, 1993). Sound from a fifth track was distributed by the *potentiometre* (Harley, 1998).

The role of space in composition is dependent on the available technology. As discussed in section 1.1, the difference between techniques and technologies is one of compositional process on the one hand and the physical technology and technical implementations on the other. There has been development from mono to stereo and on to a range of other stereo-derived spatialisation technologies. Stereo is based on amplitude panning (Roads, 1996), and the frontal localisation of sound is based on amplitude differences between the two channels (Rumsey, 2001).

Two approaches to stereo microphone techniques, the coincident (Blumlein, 1931) and the spaced (Steinberg & Snow, 1934) microphone techniques are used today and are the foundations for the techniques and technologies of Ambisonics (Blumlein) and Wave Field Synthesis (Steinberg and Snow) (Peters, 2011). Stereophonic reproduction refers the use of two speakers, where the speakers are located at $\pm 30^\circ$. If the listener is located at the centre of the *sweet spot*, then the sound is perceived as coming from a phantom image between the two speakers. *Pairwise amplitude panning* refers to the control of relative amplitudes between two speakers, direction and distance can be created by panning between the left and right speakers and depth can be created with amplitude control, filters, delays, and reverbs (Roads, 1996). The position of the panned sources between two speakers can create a sense of space that is relative to the size of the stereo setup. The phantom image in stereo is unstable, where a slight shift in position of either listener or loudspeaker will cause the image to break down. This is called the *precedence effect* (the law of the first wave front) (Haas, 1972), and if the interaural time delay between the speakers is greater than 1ms, the sound is perceived to come from the speaker closest to us. If the speakers are too far apart, then we will experience the “hole in the middle”, where there is an audible dip between the speakers where the phantom image breaks down (Rumsey, 2001).

Two notable applications of stereophony as a multichannel format can be found in Fantasound and Delta Stereophony. The first multichannel experience in cinema came with Disney’s 1940 film *Fantasia*, through the Fantasound stereophonic format (Klapholz, 1991). The format used a pan-pot to move sound around the room with constant fades in a left, center and right speaker configuration. The three frontal loudspeaker channels are still part of surround sound applications to day.

As an auditorium sound reinforcement system in the large hall of the Palace of the Republic in Berlin, the Delta Stereophony System (Fels, 1996) applies a differentiated delay to all loudspeakers in a room the time of arrival of the wavefront will be the same for all relevant listening positions. This approach effectively eliminates the precedence effect.

Traditional stereo diffusion practices over loudspeaker orchestras is a well-established practice of “converting” a work from stereo to multichannel format. The technique is to disperse the sound over a series of different types of loudspeakers, their positions chosen for their different timbral characteristics, each speaker is a “spatial projector” (Bayle, 2007). Among the many loudspeaker orchestras in the world, the first concert of Acousmonium at the GRM took place in 1974 (Bayle, 1993). The Gmebaphone was conceived by the Groupe de Musique Experimentale de Bourges and saw its first concert in 1973 (Clozier, 2001).

Quadraphonic setups employ four speakers, located 90° apart, giving an audible “hole in the middle” effect where the phantom images break down if sound is panned between the speakers (Baalman, 2010). The situation is somewhat remedied by Octophonic setups, where eight speakers are located equidistant in a circle but has a very narrow sweet-spot.

ITU-standards[16] such as the 5.1-format uses a traditional stereo frontal setup, with two speakers at $\pm 30^\circ$, one speaker at 0° and two surround speakers at $\pm 110^\circ$ (Rumsey, 2001). The 7.1 system extends the setup from 5.1 by adding two additional speakers. The 5.1 system was developed by Walter Murch as part of the sound for the 1979 Francis Ford Coppola film *Apocalypse Now!*.[17] The setup was based on the quadraphonic setup, but changed the angular locations of the speak-

[16] https://www.itu.int/rec/R-REC-BS.775-3-201208-I/en

[17] Murch describes the process and details around the development of 5.1 in a short video clip: https://www.webofstories.com/play/walter.murch/93;jsessionid=C2EB17C3BA9EC304B8E2BAFF97B4DE7C.

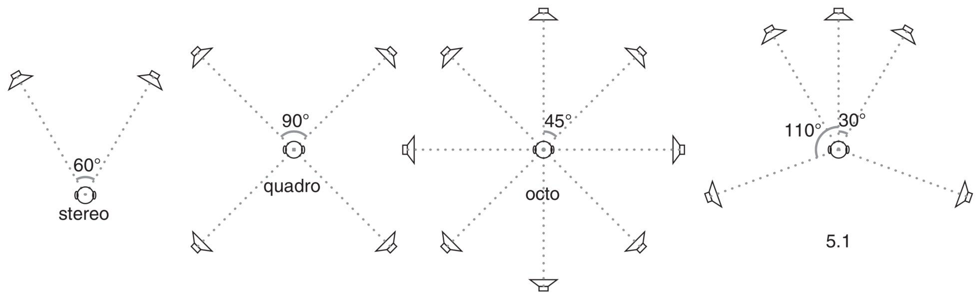


Figure 3.1: Overview of different stereophonic configurations (left to right: Stereophonic, quadrophonic, octophonic, 5.1 surround. Illustration from (Baalman, 2010).

ers and added a centre speaker reserved for dialog to create a physical and not a phantom centre. Along with this a Low Frequency Effect (LFE) was included as a dedicated sub-channel.

Positioning virtual sound sources in arbitrary locations using a 3D loudspeaker setup can be achieved using Vector Base Amplitude Panning (Pulkki, 1997, 2000). Sounds can be positioned anywhere within the bounds of the loudspeaker array. In 2D implementations the location of a source is determined according to angle, and amplitude difference adjusts localisation between two speakers (pairwise panning). In 3D approaches a sound is panned around the array and weights calculate the gain factors between triples of loudspeakers to localise the source at any point in the soundfield. Vector base amplitude panning is a very flexible method for multichannel audio but has a major drawback in its "point-source" approach (discussed in detail in section 3.6.2.

Extensions to equal intensity panning is a well-established technique for positioning mono sources in a stereo mix, and can be found in Distance Based Amplitude Panning (Lossius et al., 2009). This is a panning algorithm that distributes sound over an *N*-dimensional space. This approach makes no assumptions about the positions of the loudspeakers in space or in relation to each other. This is very flexible in an exhibition context so that the audience is free to move around the space.

Likewise, Wave Field Synthesis (Berkhout, 1988; Berkhout et al., 1993) is a technique for extending the listening area and "exhibits no pronounced sweet spot and the sampling artifacts are rather evenly distributed over a large listening area" (Geier & Spors, 2012, p. 4). Unlike stereophonic and ambisonic techniques, Wave Field Synthesis can overcome the limitations of "one 'sweet spot' and can provide a good perceptual localisation in a relatively large listening area" (Baalman, 2003, p. 1). With wave field synthesis it is possible to create a physical reproduction of a wave field, and draws on Huygens' principle that when you have a wave front, you can synthesize the next wave front by using an infinite amount of small sound sources and whose waves together will form the next wave front (Huygens, 1690). Wave field synthesis offers new possibilities for acousmatic music and sound art, and new a project involving this approach will be discussed in section 5.

Over the years many different approaches to sound spatialisation has been developed and explored for various artistic reasons. John Chowning's early experiments with moving sound sources using a digital computer and the use of four-channel movements based on Lissajous curves with Doppler shift for amplitude and frequency control (Chowning, 1971), was an early and sophisticated approach to spatial composition. Knut Wiggen's five studies *Sommarmorgon* (1972), *Etyd* (1972), *Resa* (1972), *Massa* (1974) and *EMS för sig själv* (1975) were composed with his MusicBox software and used the same streams of numbers to generate the musical material and the spatialisation over four channels (Rudi, 2018).

Many other approaches, which include spectral, granular, and decorrelation methods, have been introduced. Spectral approaches to real-time spatial distribution using frequency domain processing (Torchia & Lippe, 2004) allows for individual spatial trajectories of spectral bands, and this is extended as spectral spatialization with boids and particle systems (Kim-Boyle, 2008) which allow for a more algorithmic approach to the distribution of spectral sound masses. Boids is a flocking algorithm which simulates the behaviour of birds (Reynolds, 1987). Spectral sound diffusion can also be considered as spatialised partials (Parry, 2014), where individual partials can occupy distinct spatial positions. Normandeau has introduced *timbre spatialization* (Normandeau, 2009) where the use of bandpass filters splits a sound source into different components and uses the room (and the listener's ears) as "summation". In addition, through spectral splitting (S. Wilson & Harrison, 2010) a source is split into different frequency components that are distributed in space.

Granular approaches involve Spatialisation with Particle Systems (Kim-Boyle, 2005), Swarm Lab (Davis & Rebelo, 2005) and Spatial Swarm Granulation (S. Wilson, 2008). Among techniques related to decorrelation, are decorrelation of audio signals (Kendall, 1995b), and sub-band decorrelation (Potard & Burnett, 2004).

The Manyfold-Interface Amplitude Panning (Seldess, 2014), is an implementation of the Meyer Sound SoundMap (Ellison, 2013) that uses barycentric coordinates to spatialize sound on a dome.

A range of different libraries for spatial sound is available, including the IRCAM *Spatialisateur*, or *Spat*,[18] which is currently one of the most comprehensive libraries for spatial sound, written for MaxMSP (Carpentier et al., 2015; Jot,

[18] https://forum.ircam.fr/projects/detail/spat/

1999). A strong advantage of the library is the extensive sets of functions for panning, analysis, and building spatial systems. Version 5 of the library implements the o.dot-library (MacCallum et al., 2015) which provides an effective way of formatting all control communication using Open Sound Control (Wright, 2005).

HoaLibrary from Paris 8[19] implements ambisonic processing in two and three dimensions (Sèdes et al., 2014). A series of graphic interfaces allows flexibility for filtering, distortion, synthesis, and processing of soundfields. A separate processing thread is available in the library, where the user easily can build FX-processing or synthesis in the ambisonic domain, for example [hoa.fx.grain∼] provides granular processing on an incoming signal in the spherical domain, while [hoa.syn.grain∼] is an ambisonic granular synth.

The ICST ambisonic externals[20] provide a small but effective set of functions for ambisonic processing (Schacher, 2010; Schacher & Kocher, 2006). Apart from flexible encoding/decoding functions, the library offers a simple way of working with trajectories through the [ambicontrol] function. This allows the user to write and record trajectories of sound objects through the soundfield.

Offered for two different systems, the Ambisonic Toolkit[21] implements an ambisonic workflow in first order (Lossius & Anderson, 2014). The library exists as plugins for Reaper or as UGens for SuperCollider. The UGens offer a wide range of soundfield analysis tools.

Ambisonic decoding is offered as a separate set of tools from the Ambisonic Decoder Toolbox (Benjamin et al., 2010; Heller et al., 2012) which provides flexible tools for decoding to irregular arrays. *Spatium*[22] offers a modular set of spatialisation tools for amplitude panning and ambisonics through spatialisation renderers and interfaces, along with plugins and Max objects (Penha & Oliveira, 2013). The modular implementation is modelled on a stratified approach to spatialisation (Peters et al., 2009).

The IEM-plugin suite[23] offered as a set of VST plugins are flexible tools for encoding, transformation and decoding to higher order ambisonics from digital audio workstations such as Reaper.

Offering a large set of tools for the mainpulation of spatial scenes, the Coding and Multidirectional Parameterisation of Ambisonic Sound Scenes (COMPASS)[24] tracks the time and frequency parameters in ambisonics recording to consider the sound components that comprise a sound scene. Importantly, the

[19] http://hoalibrary.mshparisnord.fr/en/
[20] https://www.zhdk.ch/en/research/icst/software-downloads-5379/downloads-ambisonics-externals-for-maxmsp-5381
[21] http://www.ambisonictoolkit.net/
[22] https://ruipenha.pt/spatium/
[23] https://plugins.iem.at/
[24] http://research.spa.aalto.fi/projects/compass_vsts/plugins.html

ambient (background) component in the spatial scene can have directionality, offering a great deal of flexibility for the end-user in manipulating the differences between foreground and background.

Although not a spatialisation algorithm in itself, the Spatial Sound Description Interchange Format (SpatDIF)[25] is an project that aims at providing a syntactic and semantic specification for the storing and transmission of spatial scenes (Peters et al., 2012, 2013). SpatDIF provides a concise syntax for describing spatial sound scenes, and is aimed at projects where the authoring and rendering of spatial scenes occur at different times and places. A multi-layer, stratified, approach to sound spatialisation has also been proposed (Peters et al., 2009) which considers layers from authoring to physical devices, where each layer has a particular role to play. Ambisonics, discussed in the next section, affords a flexible solution to the "portability" of spatial music through its independent encoding and decoding process. However, SpatDIF can offer flexibility for a number of different fields including composition, installation, research, engineering, virtual reality, and sound design among others.

3.6 The ambisonic soundfield

Developed in the 1970s, Ambisonics[26] use spherical harmonics to encode the directional information of sound sources within a three-dimensional *soundfield*. As a "surround sound system", ambisonics consists of the *encoding* of a signal, the performance of a series of transformations on the encoded signal, before decoding to speaker feeds. The flexibility and portability of ambisonics is due to its loudspeaker-independent approach, where the physical positions of the loudspeakers are used to decode the encoded signal in the best possible way. Unlike fixed-channel distribution formats like ITU 5.1 and ITU 7.1, in ambisonics decoders can be designed for different speaker arrays (Lossius & Anderson, 2014), which enables users to "compose" space independent of the physical loudspeakers. Ambisonics uses all the available speakers of the array to recreate the recorded or synthesized soundfield.

The approach taken in recreating a soundfield was described by Gerzon as:

> For each possible position of a sound in space, for each possible direction and for each possible distance away from the listener, assign a particular way of storing the sound on the available channels. Different sound positions correspond to the stored sound having different

[25] http://spatdif.org/index.html

[26] A short documentary on the early development of ambisonics at the University of Oxford: https://www.youtube.com/watch?v=X23hZNoSkUs

> relative phases and amplitudes on the various channels. To reproduce the sound, first decide on a layout of loudspeakers around the listener, and then choose what combinations of the recorded information channels, with what phases and amplitudes, are to be fed to each speaker. The apparatus that converts the information channels to speaker feed signals is called a "decoder", and must be designed to ensure the best subjective approximation to the effect of the original sound field. (Gerzon, 1974, p. 484)

At the decoding stage, the soundfield can be optimized to psychoacoustically satisfy localisation criteria (Gerzon, 1992). Ambisonics can be decoded to different loudspeaker setups, including horizontal only (2D) and full periphonic reproduction with height (3D), as well as mono, stereo, 5.1, and so on (Rumsey, 2001). Unlike more conventional surround sound methods and stereo diffusion, ambisonics does not treat individual speakers as a projector of sonic images or as instruments (Bayle, 2007). Ambisonics is still in active development, for a current state of the art see (Zotter & Frank, 2019).

The term *soundfield* refers to the set of signals that carry the directional information that constitutes the sphere being reproduced. Sound sources can be used in two ways: either captured from an ambisonic microphone, know as *A-format*, by the likes of a Soundfield microphone,[27], a TetraMic,[28] Eigenmike,[29] or you can use mono or stereo sources and encode these as *planewaves*. The first two microphones consist of four matched cardioid capsules, which when encoded, make a *first order signal*. This is comprised of four channels, W, X, Y, Z. Where W is the omnidirectional signal, which essentially is the sound signal; and a set of three figure-of-eight patterns, X, Y, Z, are the directional pressure-gradients for left-right, front-back and up-down (Malham & Myatt, 1995). The encoded signal is referred to as *B-format* (Craven & Gerzon, 1977). The last microphone in the list above consists of 32 capsules located on a full sphere and encodes up to a fourth order signal.

Ambisonics is a hierarchical format, where an increase in order of decomposition improves spatial resolution. The abovementioned first order format ambisonic signal consisted of four channels. The spherical decompositon of a soundfield can be extended into higher orders, which provide more detailed information about the soundfield, increases the localisation of discrete sources and extends the listening area sweet-spot. In very high orders ambisonics can be equated with wave field synthesis (Daniel et al., 2003). By an increase in orders comes an increase in the amount of spherical harmonics, for a 2D-ambisonic signal the channels are cal-

[27] https://www.soundfield.com/products
[28] https://www.core-sound.com/TetraMic/1.php
[29] https://mhacoustics.com/products

culated as $M = 2N + 1$ and for 3D-ambisonics (with the height component), the channels are calculated as $M = (N+1)^2$, where M is the number of ambisonic components and N is the ambisonic order. Then, a fifth order, 3D higher-order ambisonic encoded signal will consist of 36 channels (for presentations and discussions on normalisation and component ordering, see (Zotter & Frank, 2019)). This also adds demands on the loudspeaker setup for decoding, as an increase in order also requires more loudspeakers to fully and accurately represent the encoded signal.

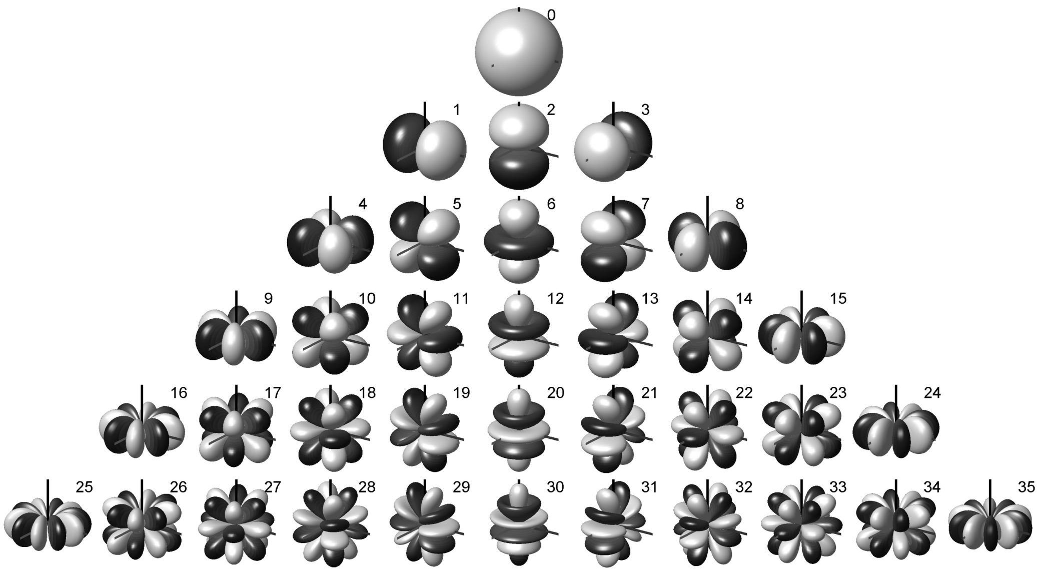


Figure 3.2: Spherical harmonics up to fifth order ambisonics. At the top, index 0 is the omnidirectional signal, and each subsequent row shows the growing complexity and improved resolution of the soundfield. Image from https://commons.wikimedia.org/wiki/File:Spherical_Harmonics_deg5.png

A series of transformations can be applied to the encoded soundfield, including *rotation* on the horizontal axis, *tilt* on the X-axis and *tumble* on the Y-axis (Malham & Myatt, 1995). *Beamforming* can be used to isolate certain sounds in a complex soundfield (Van Veen & Buckley, 1988); for example to attenuate reflections in encoded reverb (Ahrens, 2012, p. 254). Transformations of the image can be made with the *Focus*, *Press*, *Push*, *Zoom* and *dominance* functions, where dominance increase the gain towards one region of the soundfield and attenuates to the opposite side; push and press operate by an omnidirectional response and push or press in a target direction, and focus and zoom emphasize sounds in one direction through a cardioid pattern (Lossius & Anderson, 2014). The *proximity* effect is an important contribution to a perception of nearness of sounds in the soundfield (Lossius & Anderson, 2014), and for listeners of natural soundfields, the *near field effect* is perceptible through an emphasis on interaural level difference (Daniel, 2003), and through the application of filters to the encoded signal this can create virtual sound sources close to the listener (Adriaensen, 2006). The *blur* function, enables a transformation of an *N*-order ambisonic encoding to a lower order (Carpentier, 2017), creating a "bluriness" of the spatial image.

Before the sounds can be played back over a speaker array, they have to be *decoded* into speaker feeds. This means converting the encoded signal into a signal that is appropriate for the available speaker layout so that the spatial information contained in the encoded signal is represented as good as possible (Heller et al., 2012). The decoder generates a matrix of gains that contain the directional information of the entire soundfield and where the sources are located. A decoding matrix will contain the amount of channels you are decoding *to* based on the number of ambisonic components. If you have a fourth order 3D encoded soundfield and decode to an array of 34 speakers, the matrix will be 34*x*25. Each of the 34 decoded channels will refer to a matrix of 25 speaker gains.

Rather than the sound-to-speaker paradigm mentioned earlier, where individual speakers are treated as voices or instruments, in ambisonics "the sum of all the loudspeaker contributions leads to the reconstruction of the target sound field at any point in the reproduction area" (Nicol, 2017, p. 285). The strength in an ambisonic system is the flexibility it offers the production and reproduction of a recorded or synthesized spatial scene, by also encoding the directional information with the sounds and making it possible to reproduce the scene by decoding to individual speaker arrays. A soundfield *recording*, made with a Soundfield microphone, captures a full three-dimensional sound scene, whether this is a forest, a city, or a coffee cup. With soundfield *synthesis* the challenge is the creation of a spatial scene that has some relevance externally, it should have a relevance outside of panning and the novelty of multiple speakers. The main problem with soundfield synthesis is in its lack of *externalisation*, when the spaciousness of the sound is perceived to be 'out there', and this often rests on the point-source. Sound waves emitted from a source interact with the surroundings and with the listener in different ways: we do not hear "the source", we hear the sound that is the result of the sound propagation and a series of reflections. By analysing how we hear sounds in the real world and how they interact with the spaces around us, we can synthesize this to create a holistic soundfield with abstract sounds that still sound like something from the real world. This we will discuss as *spatial authoring* in chapter 4.

3.6.1 Ambience labelling information

When we listen to sound, we are surrounded by it and even if the source of the sound can be precisely localised in one direction, the sound reaching our ears is the result of a series of reflections from the surrounding space. By only using hearing, humans can localize the directions of sound sources (Pulkki et al., 2011). In the real-world, sound does not exist as a point. Sound produced by a source will propagate outward to the surrounding space and we will experience different frequency reflections and time decays from a series of surfaces in the space surrounding us.

There is “a class” of information that is necessary when creating sound environments - that of the *background*. The sonic background we experience in the everyday of our lives is something that we normally do not focus on nor necessarily notice. This is, however, the context from where we hear the sonic foreground and the sound objects that occupy the sound environment: “In sound environments, ambience labelling information *is* the context” (Lennox et al., 2001, p. 2). The background is the context from where we read and make sense of the sounds that surround us. In the real-world sounds of the everyday, the background sound is of ventilation systems, distant traffic, airplanes flying overhead, trees, birds, and other diffuse sources. Many of these sounds can be difficult to localise, but as listeners, we try to understand our context dependent on the perspective from which we experience our situation, by attempting to recognise patterns and make connections between what we hear.

All sound sources have complex radiation and directivity patterns and the importance of this in soundfield synthesis and auralisation has been the focus of studies, notably in (Noisternig et al., 2011; Okamoto et al., 2011). These complex directivity patterns of sound sources combined with a potentially complex set of reflections from the surroundings, highlights that a single point in space will not suffice to create effective sound landscapes for the use in composition, installation, simulations, education, or empirical research.

In soundfield synthesis, approaches that do not use soundfield recording techniques, can be encoded as mono or stereo soundfiles into planewaves, which is a classic ambisonic panning technique (Lossius & Anderson, 2014). This process will place a sound object in space and at a particular location, as a point in space.

In the concert hall, Griesinger has found that the sonic background of a performance space can have unique timbral and spatial qualities and properties (Griesinger, 1997), which can introduce different timbral colourations to the sound as it is experienced.

RT_{60} (reverberation time) is a measure of how long it takes for the sound pressure level to drop by 60dB (Howard & Angus, 2009, p. 301). This is an easily understood parameter but it says nothing about the amount of reflections, arrival times of these or their strength, which cause rooms with the same RT_{60} to sound very different (Halmrast, 2015). In enclosed spaces, such as a concert hall or gallery, the reflections can also be very complex but unlike the outdoors, easier to control. We experience the room effect as the direct sound, followed a series of

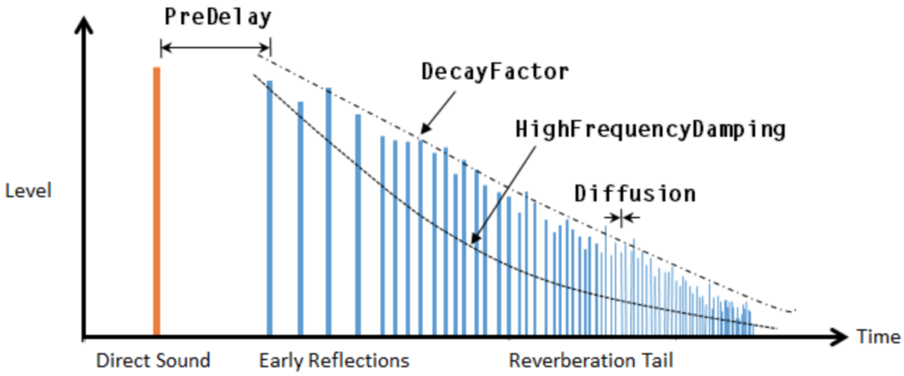


Figure 3.3: Generalized view of reverberation and decay. The illustration shows the direct sound (orange), followed by the early reflections, reverb tail and the diffuse reverb (blue). Through the decay of the sound there is also attenuation of the high frequency content in the sound. The DecayFactor is also called Reverberation Time. Image from https://se.mathworks.com/help/audio/ref/reverberator-system-object.html.

early reflections and (late) reverberation. If a reflection arrives later than 30ms after the direct sound it is perceived as a distinct echo. According to Rumsey, early reflections of 50–80ms after the direct sound can have a broadening of the spatial effect of the sources (Rumsey, 2001). The early reflections can cause interference effects in the direct sound, particularly from comb filtering (Halmrast, 2011) and all interferences from the surrounding space introduce colouration in the sound, which is experienced as changes in timbre (Halmrast, 2000). These changes in timbre influence the way we perceive the sounds that arrive at our ears and the grouping of these sounds into a coherent whole.

By drawing on these basic insights, we can use ambisonics as a very effective means for the design of spatial scenes that transgresses the point in space paradigm in ways that other spatial audio approaches do not. This can offer us methods and perspectives for fully exploring the ideas of place, space, and site in spatial, acousmatic music.

3.6.2 From point-source to sound image

With an increase in ambisonic order comes an increase in "spatial resolution", this means that the increase in resolution provides a larger listening area for the audience as well as a more precise localisation of sound sources. In a low order encoding, the sources are more diffuse and tend to occupy a larger area of the soundfield. However, given the preceding discussion, this is not without its problems.

A sound source encoded into seventh order ambisonics (64 components in 3D) provides a higher degree of spatial resolution allowing for a more precise localisation of the sound source than a sound source encoded into third order (16 components). As we discussed previously, in the real world sounds are not points. We never hear just the direct sound emitted from the source. Rather, the sound that arrives at our ears are combinations of the direct sound and a series of reflections caused by the surrounding topography. Indeed, it is rare to only experience *one* sound in any given environment. Therefore we should move *from* a paradigm of the single point-source *to* the composition of spatial sound images. One suggested approach has been the ambisonic 'O' format, which incorporates the three-dimensional radiation characteristics of a sounding object (Malham, 2001a).

Related experiments have been made using decorrelation techniques in rendering apparent source extent (Potard & Burnett, 2004), where a high degree of decorrelation of point sources causes the auditory system to perceive them as one auditory event. If the same source has been routed to multiple loudspeakers, for example, different degrees of decorrelation of each source will help to diffuse them slightly and will not create a phantom image between the speakers (Leonard, 2017) but will rather increase the extent of the source image.

Newer technical approaches to the design of spatial sound images, or sound scenes, have been proposed and implemented as audio effects for multi-directional decomposition of sound scenes in ambisonics (McCormack et al., 2021; Politis et al., 2018), which allows the user to divide the soundfield into source and ambient components after a set of spatial parameters and spatialise them separately. This approach offers the user an effective way of considering the relationships between the foreground and background, as discussed in the previous section.

Concerning approaches to spatial image formation in compositional practice, Brümmer proposes multiple applications for applying spatial components in composition by, among others, using motion to create realistic acoustic presentations, variations of speed in sound sources to create gestural quality (Brümmer, 2017). For example, the placement of sound in space can create geometric shapes, and the variability of motion can create expressive gestural qualities. Similarly, through use of Schaeffer's typormorphological criterion of *gait*[30] Barrett proposes a framework to create spatial sound images based on high order ambisonics and sound processing to create complex images through changes in image size, perceived motion, and changes to image width and depth (Barrett, 2019).

We saw in section 1.3 that the sign is a linguistic unit and the combination of a sound image and concept. The term sound image has become important in acousmatic music discourse and in its simplest term it refers to mental representations

[30] In French this criterion is named *allure* and means to walk.

evoked or motivated by the heard sound material. The *sound image* is dependent on the intrinsic and extrinsic features to function properly and the structural relationships between of the sound objects.

The thinking around the composition of sound images, rather than the single point in space, is integral if the uses of spatial audio is to have any relevance outside of "novelty". Throughout this chapter we have looked at questions surrounding spatial orientations within the world of visual arts, through notions of the site-specific and discussions around space and place, leading to the ideas of the sound landscape. For example, Smalley talks about "the acousmatic image" and "spectrally clear or blurred images" (Smalley, 1997, 2007) and these images are not isolated points in one direction, rather these images considers the totality of the interaction between sound and space. Therefore, soundfield synthesis and applications to acousmatic music should always draw on these extended notions of space, place, and site to display features from the real world, even though the sound material might be completely abstract, that is, non-representational.

The point-source, or the point-in-space, is often dominant in spatial audio practice. The advanced development of spatialisation and 3D spatial audio software still encourage composers, designers, and artists to work with a point-in-space paradigm rather than the construction of spatial sound scenes. In terms of spatialisation, a *point-source* also refers to a process where you would tie one sound to one speaker only and in that way creating a "stable" location for this sound. This refers to Roads' dictum "put a speaker there" (Roads, 2015), if you want the sound to come from a specific location. This perspective refers to the idea of the speaker as a "voice", or "instrument" (Desantos et al., 1997).

To design sound images, we have to consider the spatial identity of the sound, what it is doing in space, what its behaviour is and these different considerations lead to the conclusion that images are not points because they have a size and a dimension, and can occupy space in different ways. You cannot take the recording of a voice and change it to a whisper by just lowering the volume and positioning the sound source closer to the listener. The world is sonically complex, and we can distinguish thousands of sounds effortlessly and we are acutely aware of small sonic changes in our surroundings (Schnupp et al., 2012). The sounds we encounter in the world exist in a causal relationship with their surroundings and to each other and we can, as we have seen through previous discussions, interpret these sounds and relationships in many different ways. All sources radiate sound in different ways but we should not only consider how the sound is radiated from the source-body, we also must consider how the sound from the source interacts with its surroundings, how it is reflected and what happens to the sound as it is received by the ears.

All these issues will be returned to in chapter 4 in discussions on Schaeffer's typomorphology and the five case studies. The discussions in chapter 4 will be centered on how we can use the different categories from the typology and morphology when we are creating spatial sound images and how the treatment of sound and timbral characteristics is not a simple feature of a sound source or object, but rather is structural to our experience of sound.

3.7 Summary

In this chapter the *structure* is used as an organising principle and as an activity of perception. The objects we experience are perceived as objects within a structure, and these are contextualised through the complex interactions between sounds and spaces. In this chapter this interaction has been brought under the heading of the *sound landscape*, a term borrowed from Wishart and used to discuss the apparent and imagined source of the acousmatic, loudspeaker-mediated sounds (Wishart, 1996).

Through discussions on the site-specific and the differences between space and place, spatial audio practice has been contextualised into a wide mesh of interactions where the discursive elements of the resulting composition belong to sites, narratives, technologies, and people. The basis for the sound landscape is the *imagined* source of the sounds, this can help the audience in finding something to "hold on to" when listening (Landy, 1994). This refers to a signifier of some sort, where the experiences can be related. Likewise, for the composer, by thinking about a landscape it can aid in the structuring of spatial sound images, and not rely on the point-in-space.

The landscape denotes an *external* world, a structural feature that is mediated through subjective experience. This externalisation is important because, first, it relates what we experience of something "out there", that is outside of our heads and belonging to the world around us. Secondly, the macro to micro perspective folds into the structures of sound objects presented in chapter 2, where each individual sound object contains a shape, a site/location, and a model. Each sound object we treat and examine belong simultaneously to the perceptions of the listener and to the external world. A series of examples from artistic practices have been presented along with the different topics discussed in the text, and these examples all contextualise the topics seen from practice.

It is important to iterate that the move from considering the point-in-space to spatial sound images is a move away from treating technologies of software and loudspeakers as novelty, and instead focusing on the possibilities afforded by these approaches through the design of holistic sound images synthesized from analysis of the real world. The engagement with site, with landscape, and with place is an attempt to take the considerations on the possibilities and potentials of how spatial sound images can be designed and given a wider context.

The next chapter will draw on all the different ideas presented and discussed in these two preceding chapters. Through five case studies, the focus will be on practical applications and discussions on spatial authoring and how the different ideas on site, space, and place can be explored practically.

Chapter 4

Typomorphologies

> In an often quoted statement, Pollock remarked that "new needs need new techniques, and the modern artists have found new ways and new means of making their statements: the modern painter cannot express this age, the airplane, the atom bomb, the radio, in the old forms of the Renaissance or of any other past culture".
>
> *Matilde Marcolli*[1]

In the previous chapter we looked at the structural relationships of objects and the features associated with their spatial interaction, in relation to the sound landscape, the site-specific, and the place-bound. The structure of an object has previously been described as consisting of three parts: a shape (the overall envelope profile, the spectrum mass), a site/location (the reference to a site, a real, imaginary or imagined place), and a model (used for analysis–synthesis). The *sound landscape* denotes an implied or imagined source of the sounds through acousmatic listening. Here, the sound landscape has been taken to also refer to the physical space created through the uses of multichannel audio and how a listener can localise sources and draw connections between them when experiencing multichannel acousmatic sound. This chapter will present Schaeffer's typomorphological framework and suggest how this complex feature space can also be used to discuss spatial features. This will be contextualised through five case studies that

[1]Marcolli, 2020.

will illustrate the ideas and the concepts already discussed.

The typomorphology will be briefly presented before being contextualized to the earlier discussions (for detailed discussions and presentations around the typormophology and outlines of the “concrete music theory”, see (Chion, 2009; Godøy, 2021; Manning, 1993; Schaeffer, 2012, 2017)). Alongside this, there will be discussions on timbre and mapping as important topics to consider alongside the possible spatial correlates of the typomorphology.

In chapter 2, I voiced a critique of the terms *stream* and *event* due to the disregard of the *object*. I posited the object as a basic building block, not just of sound but also of perceptions that are not merely temporal. However, in chapter 3, the ideas surrounding the event still became an important tool in storytelling for the acousmatic composer, where the *context* in which the spatially composed work exists is that of a *sound landscape*. In section 2.6 we saw that the sound object is *not* the sounding object, the physical phenomenon, a snatch of a recording, a symbol on a score, nor the mood of the listener (Chion, 2016, pp. 171–172). The sound object is a multidimensional, ontologically complex “thing”, but when the sound object is conceptualized as the way the sound exists in the auditors mind, it can rightly be criticized from a perspective of *sound-in-itself*. This is due to the often-misunderstood practice of reduced listening, one part of Schaeffer’s model of listening discussed in section 2.7.1. The critics often treat reduced listening as a single end-goal, as a state-of-mind, or as an activity in and of itself, that is, the references to “pure reduced listening” or the “ultimate mode of perceptual contemplation” (see section 2.7.1). Nevertheless, as we saw from Chion’s discussion of listening practice, reduced listening offers a temporary suspension of our knowledge of the outside world in order to shift our intentional focus to the different parts of what we hear. This also includes Smalley’s *technological listening* (Smalley, 1997), not as a separate listening mode but rather one perspective of intentional focus. In order to judge the degree to which something can be used for sound production we need to listen to it for itself.

While the listening to, and analysis of, sound objects is an activity directed at the intrinsic and extrinsic features of the sound, we must also look to the contextual features of the sound. We discussed this in the previous chapter under the concept of the *sound landscape* and as a necessary affordance of spatial audio. Where synthesized sound does not belong to a particular *site*, the perception of the sound and spatial montage in an acousmatic context still requires an *external* relevance and relationship. I stated in the introduction to this thesis that ecologically valid explorations of these practices has been favored over controlled laboratory experiments, particularly as there arose a need to relate the practices of spatial audio to something outside of *itself*. In regards to visual perception, Bregman stated that “the term ‘ecological validity’ means that it tends to give the right answers about how the visual image has probably originated in the external world” (Bregman, 1994, p. 13). This is also relevant for sound. More so than any other format for sound reproduction, spatial audio affords us models and a set of approaches to communicate artistic and scientific intentions, concepts and the wider structure of composed works, installations, simulations, and the like. We must always strive towards the exterior, as the "elusive inside world of sound - the sonorous, the auditory, the heard, the very density of sonic experience - emerges and becomes perceptible only through its exteriors" (Sterne, 2003, p. 13). There is a certain vagueness to our sense modalities, expressed by Merleau-Ponty as "between my sensation and myself, there is always the thickness of an *originary acquisition* that prevents my experience from being clear for itself" (Merleau-Ponty, 2012, p. 224). The sensations and experience belongs to something that is outside of ourselves, something that belongs to an *accessible* exteriority.

This chapter looks into the typomorphological framework presented by Pierre Schaeffer under the name TARSOM, *Tableau récapitulatif du solfège des objets musicaux* (Summary Diagram of the Theory of Musical Objects), and we will look at the relationships this classification system can potentially have to spatial parameters. There are many diagrams presented in Schaffer's work, and an attempt at an all-encompassing diagram of a typology of sound objects was presented as TARTYP (*Tableau Récapitulatif de la Typologie* (Summary Diagram of the Typology of Sound Objects). In section 2.4 we saw that Schaeffer did not consider space as necessarily relevant in itself, rather that *time* is the space where the object exists, despite performance research that utilized spatial technologies like the *potentimètre d'espace*. Schaeffer and colleagues did not have access to the technological tools we have today, and it is fruitless to spectulate as to how the *spatial parameter* would be incorporated into his work. However, in the *Outline of a concrete music theory*, with Abraham Moles (Schaeffer, 2012), the two authors defined "25 initial words for a vocabulary" (p. 191–194), where words 23–25 are defined as *spatial music*, *static spatialisation*, and *cinematic spatialisation*. Spatial music is any music "that is concerned with the localisation of sound objects in space when works are being projected to an audience"; static spatialisation is defined as static sources in space, locatable to a point; and cinematic spatialisation refers to "projection that makes sound objects move in space at the same time as they move through time" (Schaeffer, 2012, p. 194). Harley has referred to this as *trajectories sonores*, and building on this, Xenakis formulated concepts of *stéréophonie statique* and *stéréophonie cinématique*, referring to sounds that are distributed over loudspeakers as points or where the sound sources are mobile and moving (Harley, 1998). We will return to these perspectives on sound and space when discussing classifications in section 4.1.

Before continuing, I want to return to a previous discussion, from section 2.6.4, on the linear and non-linear - discussed previsously as *anamorphosis*. The terms linear and non-linear are not easy to define but in the context of sound and audio programming we find that in a linear system we can multiply a signal by a constant, for amplification or attenuation of the signal; and in a non-linear system we can multiply a signal by another signal, as in amplitude modulation (S. W. Smith, 1999, p. 95). Linear relationships we can plot in a straight line, "even complicated linear systems can be divided into modular parts. That is, they can be taken apart and put back together again unchanged" (Roads, 1996, p. 887). Non-linear systems, however, "are not strictly proportional. One can think of them as having internal thresholds; when these thresholds are crossed, they switch into another mode of behavior" (Roads, 1996, p. 887). This description of non-linearity is similar to the perspectives afforded by morphodynamic theory, discussed in section 1.3.

The term anamorphosis refers to a distorted image that appears normal when viewed from a particular angle; for example, a circle is seen as an ellipse when viewed from an angle. This represents a non-linear relationship between the image and the viewer, and is dependent on the viewer's position to see "the image within the image". With anamorphosis, the same laws of perspective are used to stretch or distort an object on the picture plane (Topper, 2000). It is the position of the viewer that dictates which part of the image is seen "correctly". The sound signal is a carrier of information but it is not the information itself (Garnett, 1991), and sonically the "distortion" associated with anamorphosis is an alteration of the original *shape* of the sound. The questions surrounding the distortion by anamorphosis is a change in the morphology space of the sound, which is the perceived sound as we *hear* it. The sound object originates externally but exists in the listener's consciousness based on a listener's specific perspective and listening situation. Our unfamiliarity with the sounds and with the spatial conditions can quickly change, given listening intention. We try to understand our context dependent on the perspective from which we experience our situation by attempting to recognise patterns and make connections between what we hear (Kendall, 2010).

Faced with the acousmatic listening experience, listening is our primary mode of analysis and the parsing of the incoming auditory stream is "a multimodal, embodied experience of objects and actions" (Kendall, 2010, p. 68). The control of relative amplitudes of distributed audio material to a set of loudspeakers generally does not address issues of precedence (Kendall & Ardila, 2008); if the signals played back over the speakers are different, they will be perceived as different sources comings from different spatial locations and our spatial segregation will work more clearly. Indeed, Halmrast has pointed out that "the average person perhaps does not notice such colouration of the sound because often there are several surfaces at different distances that more or less randomise the comb filters, or perhaps the person has not experienced the original/direct sound by itself, unmodified by the reflection" (Halmrast, 2015, p. 257). The salient properties of a sound that is examined and studied in the studio can quickly be changed when the sound is perceived in a different space. Then we experience that “no sound event, musical or otherwise, can be isolated from the spatial and temporal conditions of its physical signal propagation” (Augoyard & Torgue, 2014, p. 4). We can then equate the composition and design of spatial behaviour, as composed and designed morphology. This will discussed further in the next section as we move on the the typomorphological framework.

4.1 The typomorphology briefly explained

At the opening of this thesis Edgard Varèse was cited for his topological vision of sounds in space. These same perspectives have been formulated by Jean Petitot, in the context of morphodynamic models and their unfolding as bifurcating, non-linear dynamic systems:

> The phenomenological description of sound images, sound structures and sound organizations is very diverse; it includes forms, figurative salience, clear and fuzzy contours, attacks and fronts, not to mention deformation, stretching, mixing, stability and instability, rupture, discontinuity, harmonic clouds, crumbling and deviation of figures and so on. (Petitot, 1989)

The bifurcations that Petitot describes are related to Varèse’s topological and spatial metaphors of colliding masses, shifting planes, projection, transmutation, and the like - and to the previously discussed anamorphosis. A bifurcation is a point where something divides into two parts (or branches) and is a model of transition (Strogatz, 2015). The typomorphology provides a framework for understanding such transitions in sound perception, and it can also provide a framework for understanding spatial transitions.

It must be stressed that I will not propose some form of “spatial category” to be added to the TARSOM, rather I want to discuss how the existing categories of sound classification and characteristics could provide us with a spatial perspective that is not dependent on the creation of new spatial *categories*. Between the topological metaphors employed by Varèse and Petiot to the discussions around space, place, site, and location in sections 3.1 to 3.3, to the psychoacoustic attributes discussed in section 3.4.1, and in the discussions to come, we already have at our disposal a multitude of potential spatial “categories”. The object–structure defined in chapter 1 describes a relationship between the intrinsic and extrinsic features of a sound object, its contextual and structural relationships to space, place, and site and how we analyse and synthesize these relationships.

The questions addressed by Pierre Schaeffer in his *TOM* on the correlations between the worlds of acoustics and the engineer with that of the listener are still valid today. More so, as music-making and the technologies become more and more sophisticated, a rigorous methodology for the correlations between the technologies and the ear become very pertinent. The complexity of Schaeffer's theories should not be underestimated, nor should the rigour in the examination of sonic matter: "The various types, classes, species, genres of objects are summarized in a huge TARSOM (Summary Diagram of the Theory of Musical Objects) which the author presents as a 'tool for investigation' and not as a table of results" (Chion, 2009, p. 100). Which in turn is further emphasized as: "the general procedure in this music theory is to move forward in a series of approximations rather than in a straight line" (Chion, 2009, p. 100). Then, the general idea in this music theory is a series of approximations through a process of analysis–synthesis, as we discussed first in section 1.2 and again numerous times in the previous chapters. These approximations are non-linear in the sense highlighted by Petitot previously and by Varèse before him. The purpose of the typomorphology and the classifications of sounds is always "What are we hearing?" in order to gain knowledge about what we are listening to (analysis) and to use this in approximating something we are moving towards (synthesis).

As a precursor to the typomorphological framework, Pierre Schaeffer cited Luigi Russolo and *The Art of Noises* (Russolo, 1913) as a forerunner through his classification of noise sounds (Luening, 1964). Russolo implemented his catalogue of noise sounds through a collection of noise-making instruments (see section 3.3).

The typomorphology is a descriptive inventory that precedes musical activity, it is the initial "phase in the programme of musical research" (Chion, 2009, p. 124). The typology is a "first sorting" according to the overall shape of the sound and the morphology looks at the internal characteristics and features of the sound object. The tasks of the typomorphology are identification, classification and description, and it is divided into three parts (Chion, 2009, p. 124):

1. Identification of sound objects (typology).
2. Classification by type (typology).
3. Description of characteristics (morphology).

Identification and classification of sound objects is a procedure which consists in isolating and cutting out sound from all possible contexts, and then arranging the sound objects by type. This sonic examination is based on subjective judgement and is done in terms of reduced listening, and, as stated several times before, involves a temporary suspension of our knowledge about the world and about the sounds we are listening to. The typology starts by identifying sounds into three different categories based on the dynamic envelope:

1. Impulsive.
2. Iterative.
3. Sustained.

The dynamic envelope is one of the possible *shapes* of the sound. From this we would only retain their most general characteristics of the sound (Schaeffer, 2017). The morphology "concerns the description of sound in terms of its internal structure" (Chion, 2016, p. 177), and is concerned with what is at work within these shapes. As we have already extended the notion of a sound object to contain a shape, a site, and a model, then we can now say that the shape of the sound is both the typological classification and the morphological description, its relationship to a sound landscape and *how* we listen to it, analysing and synthesizing, to gain knowledge about what we are hearing. Reduced listening is not "one" thing, but a result of the interrelated modes of the "circuit" presented in section 2.6 and of our shifting intentional focus. The sound object then contains its own intrinsic and extrinsic features as well as relations to the structures that contextualises it.

After this initial identification, the sounds are then classified into *pairs* of typological criteria, where they are used to give approximate distinctions between objects (Chion, 2009, pp. 134–137):

1. *Mass/facture: mass* relates to how the sound occupies the spectrum, with the matter of sound; *facture* is related to its form, the shape of the sound over time. *Mass* might be fixed, with or without identifiable pitch, and it can vary from organised to disorganised; *facture* can be continuous, instantaneous, and iterative. *Facture* signifies the assumed sound generation envelope.
2. *Duration/variation:* these criteria are tightly linked, where *duration* is the time as "psychologically experienced", and the *variation* is the experience over time. These are distinguished into short, medium, and extended *durations* and into non-existent, reasonable, and unpredictable *variations*.
3. *Balance/originality:* this pair deals with the structure of the object, and distinguishes between objects that are redundant, balanced, eccentric, or too complex. The *balance* of the object is a compromise in the *facture*, between being too structured and too simple, and is redundant if it has a "non-existent originality". Eccentric if it is "unpredictable" in its originality.

Sound objects are referred to as *suitable* when they appear more appropriate than others for musical use (Chion, 2009, p. 106). To be suitable, they must fulfill certain criteria, they must be *balanced* typologically, simple, original; and not be too anecdotal or loaded with emotion or meaning, that is, they must lend themselves easily to reduced listening. They must be able to produce an easily identifiable musical value. I will, like Thoresen, disregard these normative dimensions of the TARTYP, and remove the distinction between the *suitable* and the *unsuitable* objects (Thoresen & Hedman, 2007), which in effect means that we should also disregard the criteria pair of *balance/originality*.

It is difficult to defend a stance, today, that a sound can be too complex to be used for composition, as in acousmatic music, where *all* timbres, pitches, and complexes of morphologies are possible, and should be explored. There should be no doubt that the criteria defined in the typomorphology are the result of years of musical experimentation in the studio, but unfortunately do not always correlate with what we have described as the ecologically valid - experiments made *outside* of the studio and laboratory. Acousmatic music is a music where all levels of complexity of space are possible, and we should not consider any spaces to be redundant or too complex since our reactions to spatial environments are learnt with regards to how we associate different sounds with different types of locations and objects. Schaeffer's experiments, as all uses of technology, are part of a continuum and the technological affordances of tools guide the continued uses of the typomorphology and our intentional focus.

After the typological classification there are several possible morphological criteria that allows us to describe sounds that have already been classified in much more refined ways. The seven criteria that are defined in the morphology overlap with the criteria already defined in the typology; this system of classification, identification and description is a questionnaire, where we are listening to attain knowledge about the sound objects we are studying. Cross-referencing of criteria is therefore natural, on our path towards approximations. Chion describes the uses of these criteria as:

> The concept of morphological criterion, which is more general than value, is essential if we want to build a general Music Theory of the sound-world and must give up using the concept of timbre and traditional musical values, which are only relevant to the particular field of Western classical musics. Indeed, the concept of timbre is bound up with instrumental identification as a synthetic perception of a certain number of associated sound characteristics, rather than an aid to describing and perceiving these characteristics is themselves. Now, with studio music, there is no longer an instrument. Similarly, musical values are bound up with the traditional system of notes, and without this they lose their meaning. (Chion, 2009, p. 160)

Here Chion illustrates the problem with timbre as it has been defined and also sets out the purpose of the typomorphology. Timbre will be discussed further in section 4.1.1.

The morphology is divided into seven criteria of *mass*, *dynamic*, *harmonic timbre*, *melodic profile*, *mass profile*, *grain*, and *allure* (gait/oscillation). Again, the aim of the typomorphology is not to identify abstract values such as pitch classes, but rather to classify and to *understand* sound in its possible diversities. The following is a brief summary of the criteria of the morphology, based on (Chion, 2009, pp. 158–187).

1. *Mass* refers to how a sound occupies different pitch regions and tessituras, it describes how the sound object occupies the pitch-field, whether there are one or several distinct and locatable pitches to be heard. *Mass* also refers to colour and thickness of the sound, and is a meeting point between the old and new musics. *Mass* is complemented by *harmonic timbre*, which specifies the colour of pitched sounds, and in this way refers to the timbral dimensions of the sound's shape.

2. *Dynamic* describes the temporal profile of the sound, which is also related to the shape of the sound - and of specific importance is the attack (see section 2.6.2 on the cut bell). This criteria is especially connected to the energy articulation of the sound and is a criterion of form. This criterion exists only in time, and is considered to be a study of the attack phase of the sound and concerns the perception of variations in sound intensity, in other words, the sound's envelope.

3. *Mass profile* and *melodic profile* are linked as temporal variations, where *mass profile* refer to the internal variations of the sound mass, and *melodic profile* is a displacement of the entire pitch-field and describes the overall trajectory of the sound mass - it is the sound itself that moves, not its internal developments.

4. *Grain* is a microstructure in the sound, which can be fine or coarse, and refers to a perceived surface of the object and its tactile texture. It can refer to a very rapid *gait*, a rapid *variation*, or an accelerating *iteration*. A rapid succession of impulses stops being perceived as impulses but becomes a continuous sound with a characteristic *grain*.

5. *Gait*[2] refers to an undulating movement or fluctuation of sound objects, which can also be described as an oscillation. The oscillation of *gait* can both be in terms of duration and motion. The *gait* of a sound can be seen as a "signature" of its source (Thoresen, 2015, p. 282) and it can be divided

[2]Gait is a suitable translation of the French word *allure*, which means to walk, or a way to walk (Thoresen, 2015). In English, allure has a different meaning - "to entice by charm or attraction" (https://www.merriam-webster.com/dictionary/allure).

> into mechanical, a living agent (person), or a natural phenomenon. *Gait* is a criterion that refers to the causality of the sound.

Each of these criteria are part of the multidimensonal model discussed throughout this thesis, where the criteria are subdivided and sub-subdivided in a top-down, subjective exploration of feature categories; needless to say, Schaeffer's system is complex and it is not possible to do justice to its depth here. The system does, however, provide opportunities for the exploration of spatial features. This system provides the listener with a framework for exploring the intrinsic and the extrinsic features of sound objects. Particularly through the applications to spatial features as it does not consider space as some abstract entity but rather analyses sounds for their features, shapes, and motions. The perceptions of spatial environments depend on the listener's accumulated knowledge of the physical (and external) world:

> When sensing a spatial environment, an individual builds a cognitive map of space using a combination of sensory information and experiences accumulated over a lifetime. The cognitive map of space in our consciousness is subjective, distorted and personalized - an active and synthetic creation - rather than a passive reaction to stimuli. (Blesser & Salter, 2009, p. 46)

This reference to the cognitive maps we use to sense spatial environments aligns with the message from Schaeffer's musical research, and it is through our subjective and attentive perception of the world and the sounds contained within it that we make sense of what we are experiencing. For the identification of a sound object, the "identification is done by reference to a higher level of context which includes the identified object, as an object in a structure" (Chion, 2009, p. 61). It becomes clearer when examining the different criteria from the typomorphology, that sounds have a relationship to the external world, and it is the sound's morphological criteria that provides us with clues as to how it existed spatially and how we can make it exist spatially. When we practise reduced listening, we can examine the sound's features, and through identification, classification, and description we give it external relevance both to the sound's own "signature" and to its imagined place in a sound landscape. That being said, the taxonomical tour de force developed by Denis Smalley in his article "Space-form and the acousmatic image" can serve as exemplary in analytical terms for discussing the potentials of spatiality in the typomorphology (Smalley, 2007), based on our definition of the object-structure. Before continuing this discussion with case studies and other spatial topics, I want to discuss timbre in the context of spatial, acousmatic music and how this is seen through the lens of the previous discussions.

4.1.1 Timbre

Previously Chion was cited as saying "the concept of morphological criterion, which is more general than value, is essential if we want to build a general Music Theory of the sound-world and must give up using the concept of timbre and traditional musical values" (Chion, 2009, p. 160). Given the narrow definition of timbre from the Acoustical Society of America standard 11.09,[3] which defines how sounds of the same pitch and loudness can be told apart, it remains one of the most unclear and poorly defined parameters in music, psychoacoustic, and acoustic research. This definition of the term can in no way provide a fruitful starting point for discussing sound qualities that extend outside the confines of Western instrumental music. It is an "illdefined wastebasket category" (Bregman, 1994, p. 92) and "a catch-all term for all those aspects of a sound not included in pitch and duration. Of no value to the sound composer!" (Wishart, 1994, p. 135). Timbre is a "multidimensional property" but there is little scientific agreement on what these properties are (Roads, 2015). Rather than offering my own definition of timbre, this section aims to discuss this term with a little wider context than that of a musical instrumentarium and perhaps offer some perspectives on how we can consider timbre from the viewpoint of spatial audio.

Owing to the restrictive definition of timbre, the call for the dismissal of the term all together rests on its inability to convey the complexities of sounds outside the traditional instrumentations which it points to. Chion remarks that "what does the expression 'a trombone's timbre' mean once one strikes the instrument rather than blowing through it in the traditional fashion?" (Chion, 2011, p. 238). An instrument does not have "a" timbre, rather "it is misleading to suggest that one sound-producing object or instrument yields exactly one timbre. Contrary to parlance of 'the bassoon timbre', there is no single timbre that fully characterizes the bassoon. The timbre of a bassoon tone depends on pitch, playing effort, articulation, fingering, etc" (Siedenburg & McAdams, 2017, p. 3). Timbre is a perceptual attribute that tells us something about the identity of the source and of the event that caused the sound. It can also tell us something about the context of the source.

The insights learned from the experiments with the cut bell, as discussed in section 2.6.2, describe how the removal of the attack of a sound changes its timbre, how a bell with its attack removed sounds like a flute (Chion, 2009, p. 13). Again, this shows how a perception of timbre is not restricted to the narrow definition of how sounds with the same pitch, duration, and amplitude are dissimilar, but rather

[3] https://asastandards.org/Terms/timbre/. Also the American National Standards Institute definition of timbre is "Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar" (1960).

timbre points to something designed and at the same time describes the *shape* of a sound. For example, in Horacio Vaggione's *PianoHertz* (2012) (Vaggione, 2017), piano sounds are processed and granulated, then spread across the stereo field. At times the glissandos on the piano are recognizable as a distinct piano, at other times they are smeared and stretched in time and frequency. However, there is still a correlation between the processed and altered sounds to the dynamics of the instrument and the abstract and granularized sounds are still part of the piano's perceptual timbre.

In *PianoHertz* the reference back to the piano is still audible, and this does not refer to "the" timbre of the piano, but rather to the sound of *a* piano, or the colour of the sound. Hermann von Helmholz introduced the term *klangfarbe* to describe sound quality (Helmholtz, 1885/1954), indeed, "timbre refers to the 'color' or quality of sound, and is typically divorced conceptually from pitch and loudness" (Wessel, 1979, p. 45). In other perspectives "timbre appears as the global character of something deliberately designed, that is, as the perceptual dynamical properties of a composed, formed, sound object" (Di Scipio, 1994, p. 135), which indicates that timbre is only perceivable when something is composed or designed. However, it is our ability to differentiate between sounds that defines timbre: "Once we have equalized loudness, spatial location and pitch, we call the perceptual quality that still differentiates between sounds timbre" (Schnupp et al., 2012, p. 101). Nevertheless, sounds arriving at the ears will have been reflected over multiple surfaces and these cause a colouration on the original sound, which is a *change* in timbre (Halmrast, 2015). Thus, a perception of sound quality, sound colour, or timbre is not solely dependent on the sound "itself", rather the spaces also contribute. Different rooms sound differently and "the background can have its own spatial and timbral properties" (Griesinger, 1997, p. 725).

However, just discussing different perspectives on what might constitute timbre and how this can be defined (or re-defined) is not enough. We have already established that there is not one timbre to an instrument, rather it consists of many different possible timbres, based on register, how it is played and the like. If a sound is made in a reverberant room, then the tone is sustained and masks the start of the next tone (Halmrast, 2018). In a concert hall, the first reflections, especially through interference by comb filtering, lead to a change in tone colouration and "image shift" (Barron, 1971). Likewise, our spatial perception can be influenced by echo disturbances, shifts in the image of the apparent source, shifts in spatial impression, and from different modifications of timbre (tone colouration) as functions of differences in intensity and arrival time (Kendall, 1995a). If we then look back to Schaeffer's initial categories of identification based on envelope, an *impulsive* sound can in a given space, be classified as *sustained*.

Let us return briefly to a discussion in chapter 1, where Brün was cited as stating that we should aim for the composition *of* timbre, rather than composing *with* timbre. Given the preceding statement by Di Scipio, we can then see that the creation of timbre is contingent on the combinations of sound materials and is in no way merely a property of acoustic instruments. A space changes the timbre of a sound and timbre is then a form of the sound materials. However, can we extend this even further? Can *a place* have a timbre?

Timbre is the result of "dynamic morphologies", where "the discovery that timbre itself is partly dependent upon the evolution of spectral characteristics is our first real link with sounds of dynamic morphology, i.e. sounds in which the perceived pitch spectrum, amplitude envelope etc., all evolve through time" (Wishart, 1996, p. 64). From this statement we see that all of Schaeffer's preceding categories contribute to the understanding of timbre, regardless of the sound's origin. These dynamic morphologies extend outwards - not just to the sound of the materials but also to the sound of a space, to the sound of a place. Complex spatial organisation can also be understood as timbre (Gottfried, 2016) where the complexities of spatial organisation of sound material contribute through a shared perceptual morphology. The potential morphologies contained in a performance space and how the sounds are changed depending on the architecture and the reflections imposed on the sounds.

In section 2.7.4 we discussed the soundscape in approaches to listening where a *soundmark* defined as the auditory counterpoint of a landmark: a sound or sound event, that is of particular importance in the community. It is not only something that is important to a community, but also "refers to a community sound which is unique or possesses qualities which make it specially regarded or noticed by the people in the community" (Schafer, 1994, p. 10). A soundmark makes a community and a place unique, it is in this sense the timbre of this sound, the character of this sound, that makes the place unique. In this sense, timbre is not, as Di Scipio argued, an intentionally formed sound object, but is part of the complex mesh of interactions between people, histories, technologies, and ideas.

In the 19th-century French countryside, "self-esteem, emotional well-being, civic pride, and territorial identity all depended on hearing the town bells. When citizens heard the chiming of the bells, they felt rooted within a cultural geography that could easily be walked" (Blesser & Salter, 2009, p. 30). The citizens heard the timbre of the bells, and felt connected to a specific place through the *soundmark* these bells represented, and the distinct sound of one village's bell could be differentiated from another village. When the sundial and hourglass was superseded by clocks that chimed the hours, time was broadcast at punctuated intervals. This created an acoustic arena, which replaced this small visual arena (Blesser & Salter, 2009), and this acoustic arena contributed to making the aural place important.

These examples would suggest that the sound space is important for our perception of the sound, as it provides some background context. Traer and McDermott have found that the brain separates sound into contributions from the source and the environment, which contributes to a robust recognition of the spaces around us (Traer & McDermott, 2016). This was touched upon in section 3.6.1 where we discussed that, through ambience labelling information, the background is the context (Lennox et al., 2001) and a perceiver's viewpoint of the surrounding space is as "relative units", depending on "the perceiver's assessment of the importance of various features of an environment" (Lennox et al., 1999, p. 132). As we saw in section 3.1, space is a social morphology that is at the same external to us as it is something we are part of (Entrikin, 1991). We belong to the changing sonic characteristics of a place and this "thingification" (Merrifield, 1993) of place makes up a net consisting "of a potentially unlimited number of interconnected nodes" (Trochimczyk, 2001, p. 50). Timbre, then, is the sound colour, the sound characteristics, and the morphological characteristics that makes up a sound, a space, and our subjective perception of it.

From these discussions of the identification, classification, and description of sound as well as timbre, we will turn to some applications of these ideas in composition and sound design in terms of space and morphology. Before discussing the five different case studies, we will first look into notions around mapping functions, and specifically discuss how mapping can be applied in discussions around spatialisations. This section looks past the music technological perspective on mapping and considers mapping of and to a landscape. Then, a brief section will discuss some different audio effects programming experiments aimed at spatial audio applications. These next two sections will provide some additional context to the case studies.

4.2 Mapping functions

In section 1.3.3 we looked at a methodology for the three-dimensional representation of knowledge in cartography, through the (Cartography)[3] (MacEachren, 1994). This representation of knowledge is aligned with the discussions on semiotic morphisms, transpositions, structure-mapping, and cross-domain mappings in the same chapter as movements of knowledge along axes. In this section we will continue these discussions through the issues of *mapping*. This section will first present mapping from the traditional music technology perspective before bringing mapping into a wider spatial audio-relevant discourse by drawing on the preceding discussions around space, place, site, and our physical location. For the contexts of the present discussion, the instrument discourse *mapping to* should be preceded by the step *mapping of*, for an evaluation of mapping in spatial discourse. This has perspective has been a methodological focus in approaches to eco-structuralism (Opie & Brown, 2006, 2009), among others.

Mapping is a central term in everything from music technology, mathematics, linguistics, dance, performance, to almost any conceivable field of study. It is a concept that describes the process of taking data from one domain and transferring it into another, where each element in a set is associated with one or more elements in a second set. These concerns have been discussed many times throughout this thesis, especially through semiotic morphisms and the actor-network theory. A map is a representation of data, showing for example, the positions of the stars, spatial arrangement over an area, a sequence of genes, or a representation of land. In music, it refers to the designed link between an instrument's playing interface and the sound source (Hunt & Wanderley, 2002), or to the "transformative method to link performer actions to composer object parameters" (Winkler, 2001, p. 190). Indeed, mapping also refers to the capture of gestures and then bringing the data into the computer for some form of processing (Miranda & Wanderley, 2006).

Digital instruments can quickly become more complex than acoustic instruments and the paradigm of mapping in these contexts have certain limitations, especially if the instrument is non-deterministic: "Mapping describes the way a control is connected to a variable. ...the concept of mapping becomes more abstract and does not describe the more complex realities of electronic instruments" (Chadabe, 2002, p. 4). In this sense, mapping is inextricably linked to a practice tied to acoustic instruments and how we model digital instruments around the same set of playing gestures.

At the heart of mapping strategies are coupled components. From acoustic instruments we learn that the relationship between the bow, the strings and the soundbox is inseparable and these act as both the control mechanism and the sound generation (Hunt et al., 2003). These types of mapping schemes can be extended beyond this, and there have been proposed four classic categories for mapping strategies:

1. One-to-one.
2. One-to-many.
3. Many-to-one.
4. Many-to-many. (Hunt & Wanderley, 2002)

Musical parameters are, as defined by Josef Häusler, all sound or compositional components that can be isolated and ordered (Landy, 1991). Several people have argued that digital instruments imbue some agency on the users, where other technological systems do not (Bown et al., 2009; Magnusson, 2009); however through the ecosystemic interface (Di Scipio, 2003) and the improvising machine (Lewis, 2000), agency is extended from merely considering the coupling of the performer and the instrument, but rather to the broader scope of interactive systems. Yet, this does not at first view tell us anything about how *mapping* relates to issues of spatial audio.

How does the question of mapping translate into spatial audio? Can we use the same mapping schemes as described in the music technology literature? Can we "map" sound to space? Can we translate human gestures into spatial trajectories, map movement of people to relative positions of sound sources or through some other rule-based system for the translation of a control input to a spatial output (Van Nort et al., 2014)? Certainly as a control paradigm, but if we seek to address the complexities of spatial-audio mediation that we have been discussing so far, I will propose that we look beyond the simple mapping paradigm as it exists in relationships of musical instruments and extend into physical, spatial considerations akin to what we have done so far with the landscape, place, and site.

Mapping in the traditional music technology sense is ultimately tied to musical instruments and our interactions with them. In the introduction to chapter 2, I stated that acousmatic music affords the listener a wide register of sound directly related to human emotions, perceptions, and memories through the use of sound material that is related to natural, technological, and human relationships. But as noted in chapter 1, acousmatic music affords a decoupling from the human body and its limbs as the gestural references to sound-making are removed. That is, the external references of instrument performance are removed. However, we are not decoupled from the *experience*, and we *embody* the sound we hear as it surrounds us and acts on us. This spatial occupancy locates the listener at the centre of the sound and each listener is the centre of their own listening experience. Chion reminds us that it is not "the psychology of the auditor that matters, it is the particular spot where the latter is positioned that does" (Chion, 2016, p. 172).

The privileged listening position can be considered as equivalent to the privileged viewing position in Renaissance perspective painting, which places the viewer at the centre of the world depicted in the painting (Bishop, 2005). This is similar to the perspectival view in the concert hall, a privileged listening position with a "frontal" perspective. In the concert halls there are places that function as acoustic sweet-spots from where you will have the best listening experience, and in multichannel music this position is centrally located among the speakers.

Through installation art, and in extension interactive environments, the idea of the single, privileged position is abandoned in favour of a view of the dislocated individual through a post-structuralist view that

> the correct way in which to view our condition as human subjects is as fragmented, multiple and *decentred* - by unconscious desires and anxieties, by an interdependent and differential relationship to the world, or by pre-existing social structures. (Bishop, 2005, p. 13)

This dismissal of the "centring" model of Renaissance perspective affords the possibilities of multiple perspectives as there is no one right way to view the world.

To explore this further through mapping and spatial audio, we will start with the body and move on to spatial understandings through cartography.

The choreographer, dancer, and theorist Rudolph Laban studied human movement and introduced the term *kinesphere* to denote the construct of our awareness of the area that the body is moving within, the surrounding environment, and the others we interact with. The kinesphere is "the sphere around the body whose periphery can be reached by easily extended limbs without stepping away from that place which is the point of support when standing on one foot" (Laban, 1963, p. 10). With the kinesphere, Laban supplies a geometrical framework for describing movement in connection with the environment, with spatial patterns and lines of spatial tension through which all axes pass (Maletic, 1987).

William Forsythe takes the kinesphere and decentres it. This frees it from its anchor at the centre of the body and extends it so that the new, decentered kinesphere can be the centre of the ear, the elbow or between two limbs: "I began to imagine lines in space that could be bent, or tossed, or otherwise distorted. By moving from a point to a line to a plane to a volume, I was able to visualize a geometric space composed of points that were vastly interconnected" (Forsythe & Kaiser, 1999, p. 64). This decentred kinesphere is not tied to a central position, rather it responds to changes in perspectives and to changes in and of scale. Forsythe multiplies the centres within the body, and transposes them into the space surrounding the body, this is the "creation of a many-timed body, as opposed to a shaped body, folding and unfurling towards and against itself" (Cvejic, 2015, p. 138). Forsythe's malleable geometrical shapes extends from the human body, and exists as a long lineage in modern dance and choreography of using geometric abstraction, which imagines lines in space and around which the dancer rotates. This offers multiple possible perspectives that find a reference in the previously discussed concept of *anamorphosis*.

Anamorphosis was discussed in section 2.6.4 as *warping*, where a picture within a picture can be revealed by a shift in viewing position. But importantly, anamorphosis denotes a change in the original shape. Forsythe's decentered kinesphere is a change in the perceived shape of the body, an installation is the perceived change in the shape of the viewing/listening space, and the sound landscape is a change in the perceived *source* of the sounds we hear. The landscape is "a way of seeing the world" (Cosgrove, 1998, p. 13). It is through the landscape that we will continue to discuss the issues of mapping.

A large part of Marcel Duchamp's artistic and art historical reputation is through the "readymade", an art-object made from mass-produced objects[4], notably his *Fountain* (1917), a urinal placed on its back and signed by the artist with the false name, "R. Mutt". Despite the notoriety of these objects, they contain references

[4]https://www.tate.org.uk/art/art-terms/r/readymade

to a cartographic coherence, observed as:

> Duchamp's readymades engage analogy, humor, and shifts in scale to translate elements of the human made urban landscape into the interior landscape of the studio. Such shifts and translations parallel the physical and conceptual transformations of landscape into cartographic representations, or maps. (Housefield, 2002, p. 478)

These objects defined a mapping of landmarks from the city of Paris to his New York studio, where the objects *Bottle Rack* (1914) and *Bicycle Wheel* (1913) are representations of the Eiffel Tower and the Ferris wheel above Champ de Mars (Housefield, 2002). These objects are visual representations of landmarks in Paris through a mapping of formal geometric forms to existing objects. Through this mapping, "cartography translates physical and social forms alike, using codes of reference that remain internally consistent within a single map and throughout a series of related maps" (Housefield, 2002, p. 478).

Ultimately, what do these things mean and why is this relevant to the discussions on spatial audio? It is important because it represents a shift away from formal analysis and aesthetic appreciation to that of interpretation and criticism, and a move towards conceptualism, as we saw in the discussions of space, place, and site in chapter 3. Through the discussion on the non-site in section 3.3.1, we saw that a map is a selective representation of reality (Black, 1997), where we choose what features to represent. Through an example by Bruno Latour, these representations occur through what he calls "cycles of accumulation" (Latour, 1987). His example is that of the cartographer. The cycle starts with an explorer travelling to distant lands with instruments and equipment, with the aim of drawing a map of a remote land. Here the cartographer meets people, draws, and notates in his sketchpads and notebooks until he leaves this land to returns home with a map. Later, a new explorer is sent out, not only with equipment and ships but also with the maps of the previous expedition. This explorer returns with what should be a better map. This new map is then added to the context of the previous maps. The process of making science is, according to Latour, the result of these repeated cycles of accumulation. These iterative processes are also the process of art-making, through repeated mapping and re-mapping of objects and structures.

The map is, like our axial feature space, a field of possibilities and a field of concepts. Maps are made "not of wildlife, earthquakes, hurricanes, mountains, canyons, birds but of *signs* – these themselves composed of marks and concepts" (Wood & Fels, 2008, p. 190). The maps we use are multidimensional feature spaces.

Latour and his colleagues state explicitly that "maps are interfaces to datasets" (November et al., 2010), meaning that there is not an isolated mapping between a model and a representation. Instead there are multiple mappings of correspondence between the different maps. A map (a scientific visualization) is only meaningful when presented in context to other mappings; for example, theories, instruments, texts, charts, equations, and citations etc. Each inscription is one step in a chain of reasoning to reach a conclusion. Remove any one element from this cascade of inscriptions and it looses much of its meaning because it looses its context. Each of these inscriptions are one step in a chain of reasoning to reach a conclusion, which creates a cascade of meaningful mappings that as a whole presents, preserves, and conveys a meaningful content. New mappings are based on previous mappings, which creates new forks and again creates new mappings.

So far we have followed the mapping concept from initial discussions on the relationships between gestures and musical instruments to a spatial orientation that seeks to break away from a fixed perspective orientation to a malleable decentring of the human body and our experiences. This decentring is further explored through geometric abstractions where the body can be tossed, distorted or bent and represents a shift in perspective on how the body is seen in relationship to the surroundings. This warping of our perspectives is then brought into physical objects through the representation of Parisian landmarks to the readymades of Duchamp. In section 2.4, we discussed that spatial understandings maps onto an understanding of objects and shapes, where the understandings of spatial configurations are contingent upon modality-specific translations (Landau & Jackendoff, 1993), which have also been defined as cross-domain mappings (Lakoff & Johnson, 1999) earlier in this text.

Through these perspectives on mapping I want to bring the focus back into what has followed us since the start of this thesis on the re-definition of what constitutes a sound object, namely, that the objects we interact with contain three parts: a shape, a site/location, and a model. If we shift our focus from aesthetic appreciation to interpretation and criticism, we should also shift our focus from mapping *to* something, to a mapping *of* something. In this sense, we can only map *to* space if we have already done a mapping *of* space. That is, we have created a *situatedness in space* and the mapping *to* space is an extension of the discussions of chapters 2 and 3. The three-part object affords us an interface to analyse and synthesize this approach, and the typomorphological criteria affords us a methodology for approximation and exploration.

This *situatedness* rests on a social, geographical, or cultural context to create meaning of spatial relationships. In the case of Duchamp's readymades, where "common objects of everyday use only slightly modified, if at all, by the artist but turned into art by selection and relocation alone" (Cosgrove, 2005, p. 38) illustrates the idea that a mapping of these sites in Paris can be mapped to a collection of related objects set to represent these actual places. Or, indeed, as experienced in Trond Lossius' *Edgelands* (2019), an ambisonic composition which maps "political, societal, economic and environmental issues”[5] through field recordings of suburbia, “those liminal zones where human activity fades away and nature takes over”[6] to invite the listener into multiple modes of listening to understand and appreciate these sonic environments. The mapping *of* has preceded the mapping *to*, where the superimposed layers of field recordings are complemented by ambient textures which creates recurring motifs of birds and traffic, passing trains, helicopters and planes, and distant sounds of cities.

The next section looks at some approaches to audio effects processing for exploration of spatial morphologies. These explorations have been inspired by existing applications and have been part of a series of explorations that look at the interplay between the object and the structure. None of the material in the next section has been published elsewhere, and will be further developed into a software library at a later time. The discussions will be limited to the relevance of this thesis.

4.3 Morphological experiments

This section looks at a few experimental approaches to the practice of spatial audio, which are influenced by and developed in light of the case studies that are discussed in the following sections, as well as the typomorphological criteria discussed at the start of this chapter. Many experiments have been explored throughout this thesis project but only a small selection of these are presented here. These experiments were the result of an interest in using existing effects processing paradigms (Dattorro, 1997a, 1997b, 2002) as spatialisation functions and came about through a playful experimentation with the possibilities afforded by software, as already discussed as practice as research in section 1.2, where there would be no “correct” way of using or experimenting with the different processing functions. An advantage of this approach is the freedom to “stack” and embed different processing functions on top of and into each other. The experiments explore channel-fixed and one-to-one methods, as well as others that bridge into ambisonic processing. The experiments look at motion, filters, and reverberation through MaxMSP.[7] The experiments presented here are primarily in the time domain, and the results from these experiments will be compiled into a software library at a future date.

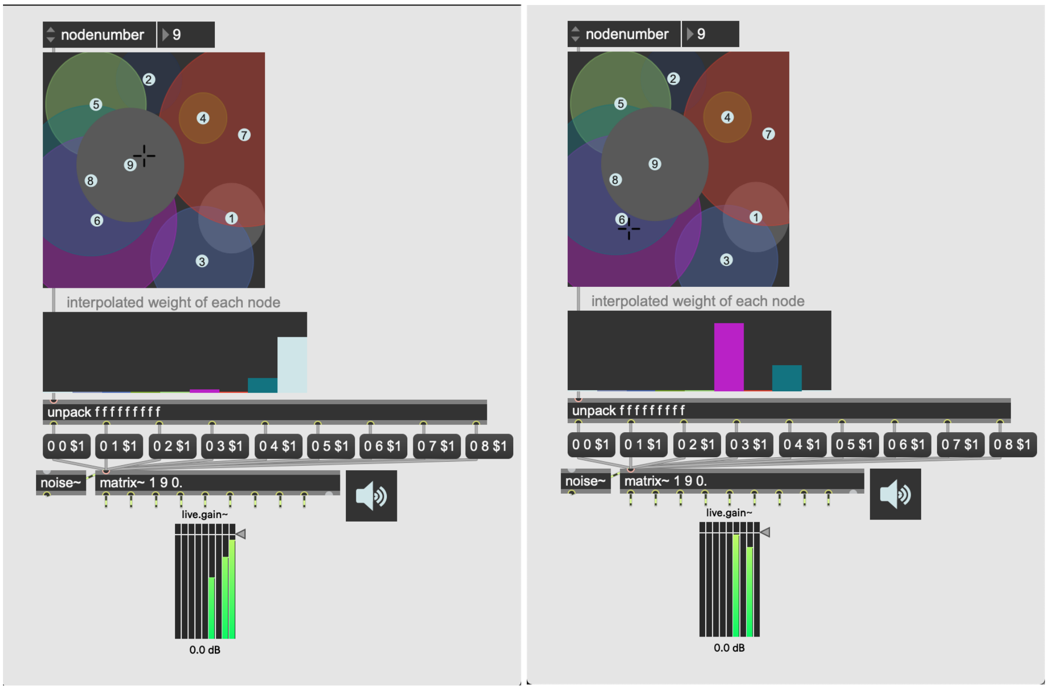
[5] Private communication, 04.05.21.

[6] From the program note: http://www.trondlossius.no/articles/1316-edgelands---new-work-at-ultima-festival

[7] Where there is a reference to a MaxMSP function, the name of the function in encapsulated in square brackets, such as [noise~].

4.3.1 Motion

As discussed earlier, sounds in the world are complex as is our interaction with them, therefore it is a problem when sounds are spatially represented as point-sources with only a direct sound. We need to overcome some of these limitations when working with spatial audio to attempt to represent sounds in more holistic ways, and we could consider this from the idea of motion processes, as stated by Smalley: "sonic motion can suggest real or imagined motions of shapes in free space" (Smalley, 1997, p. 110). In section 4.5 we will discuss an approach to creating stochastic motions of sounds though an irregular setup in an exhibition space, which was setup to to avoid circular trajectories. This stochastic approach affords an unpredictable *gait* of the sound over *N*-loudspeakers.



(a) (b)

Figure 4.1: Two [nodes] functions with weights. Depending on the placement of the cursor, each of the "blobs" will contribute different weights depending on area of activation. The cursor can be automated or user controlled and can effectively control relative amplitudes of sound sources in an encoded soundfield or loudspeakers placed in a listening space.

Two approaches to create sound motions was experimented with as part of the work discussed in section 4.5, and are briefly presented. The first looked at the native [nodes] function in MaxMSP, that is a graphical interface object which allows the user to interpolate data graphically. This affords a "localizing" of speakers in an irregular two dimensional space - an exploration that was inspired by Distance-Based Amplitude Panning (Lossius et al., 2009).

The [nodes] function is a visual approach to interpolated data, where the user can map a series of coloured blobs to 2D space and with a cursor can navigate this same space. The function outputs the weighted values of all blobs from 0. to 1. At the center of the blob the cursor is at 1., but several different size blobs can be overlaid, which can create complex interactions. In Figure 4.1 the weighted values control the amplitude of nine channels of a [matrix∼], and is an illustration of the possible uses to create motion through a dense two-dimensional space. The cursor can be mapped to the output of an envelope follower (discussed subsequently) whereby the motion between the different channels is directly proportional to the magnitude of the signal. This approach then maps the *dynamic* of a sound directly to the *gait* of a sound as represented in space. Given the flexibility offered by numerous libraries, and with a little extra programming, this [nodes] function could easily control the relative amplitudes of nine different sound sources in an encoded ambisonic soundfield, allowing the user to easily navigate through the different variable amplitudes of the sources, as for example demonstrated by one approach to imaging (Barrett, 2019) or as a method for using wave terrain synthesis in spectral spatialisation (James, 2015).

In the installation *The Exploded Sound* (2012), Nye Parry used spectral decompositon techniques to individually spatialise partials which were located to individual loudspeakers, yet the perceived motions and rotations of the sound were created by the audience walking around in the room (Parry, 2014). Parry's approach drew on spatio-operational spectral synthesis (Topper et al., 2003), where the authors have found that rotation and motion of synthesised sound increases the awareness of individual partials and their location in space.

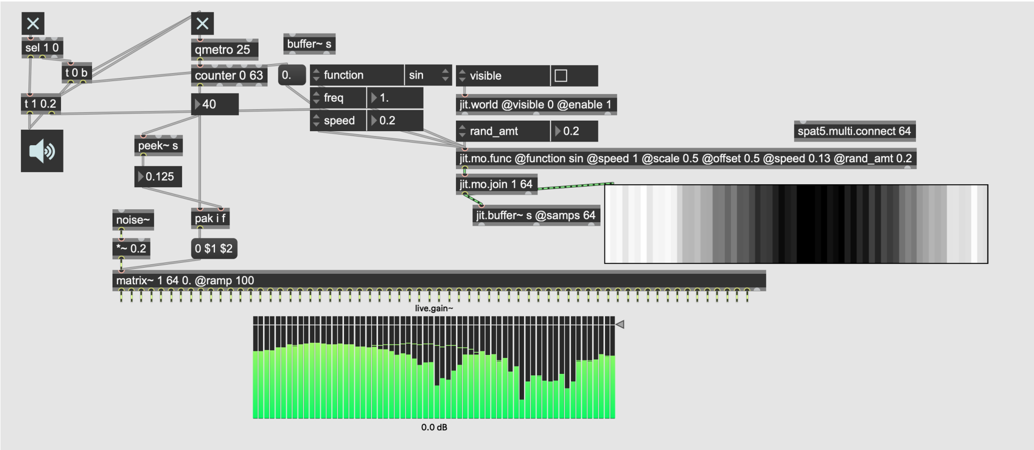
Parry's approach is not unlike the realisation of David Tudor's *Rainforest* (1973), where suspended objects are resonating in the space through uses of contact microphones and transducers, and the audience was encouraged to move through the space to experience the spatial effects of the work.

Using acoustic instruments, electronic processing, and spatialisation to creating motion with both static and moving sound sources has been explored by the improvisation quartet Lemur in the piece *Lemuria*[8] (2017) among others, where the piece was realized as six "islands" where six performers are located. The sound from the performers is amplified and spatialised around the room and the audience can walk around the performance space freely. This consideration in motion both of sound sources and of the audience was the intention in the case study *City Dwellers II*, in section 4.7, where motion is three-part, first as point-

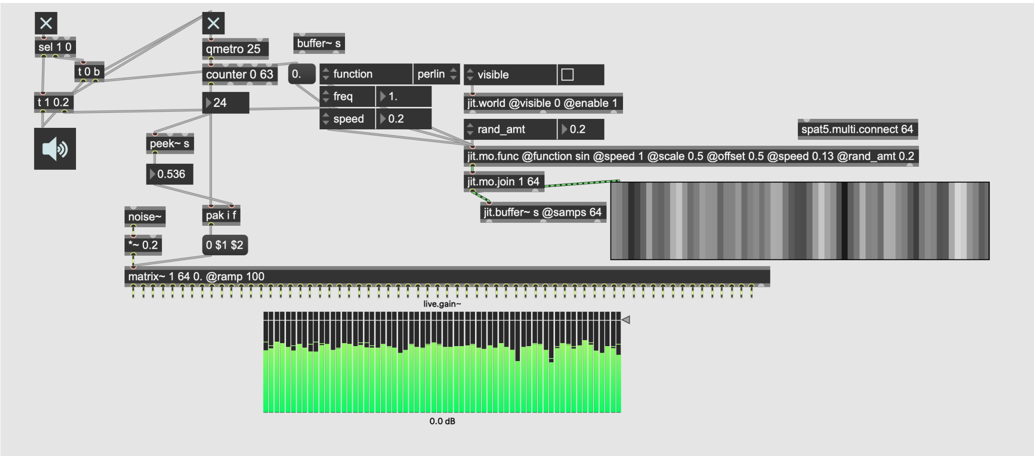
[8] http://www.lemur.fm/lemuria/

sources from each individual physical object located in space which are sounding at different times, the surrounding sound design which creates a bounding, moving surface around the audience, and the audience itself.

An inherent drawback in using this graphical function for spatial motions is one of directionality. Each blob is circular, where the weighted values are equal in all directions. A loudspeaker does not radiate sound in all directions and these blobs will not accurately represent the spatial distribution of sound as it is interpolated between the different speakers.



(a) Matrix controlled by *sin* function



(b) Matrix controlled randomly

Figure 4.2: Sixty-four channel output matrix controlled by *sin* and *random* functions using Jitter matrices.

A different approach to creating motion using the same methodology, as seen in Figure 4.2, is through jitter matrices. The [jit.mo.func] function generates a single dimensional matrix using a specified signal function, similar to that of a sound oscillator. Used in conjunction with a [jit.mo.join], which streams and output a matrix, this is stored in a [jit.buffer] with 64 samples per buffer. To control the amplitudes of a [matrix∼] this is read back from the [jit.buffer] using a [peek∼] function to control the mixing matrix. This is a simple way of creating spatial complexity using one or multiple sound inputs into the [matrix∼] and is an easy method for applying the *gait* of the fluctuations in the matrix to the *mass* and *dynamic* of the sound.

An envelope follower tracks the shape of a sound's envelope in real time (Reiss & McPherson, 2014). This produces a *dynamic* control signal that can be applied to a different sound by multiplying the second sound by the envelope, This *imposes* the envelope of the source sound on the second sound, as illustrated in Figure 4.3. However, it can also be used to control the cutoff or centre frequency of a filter, to control fluctuations in delay times or it can be used to move sound sources around in a three-dimensional soundfield dependent on the *dynamic* profile of the sound.

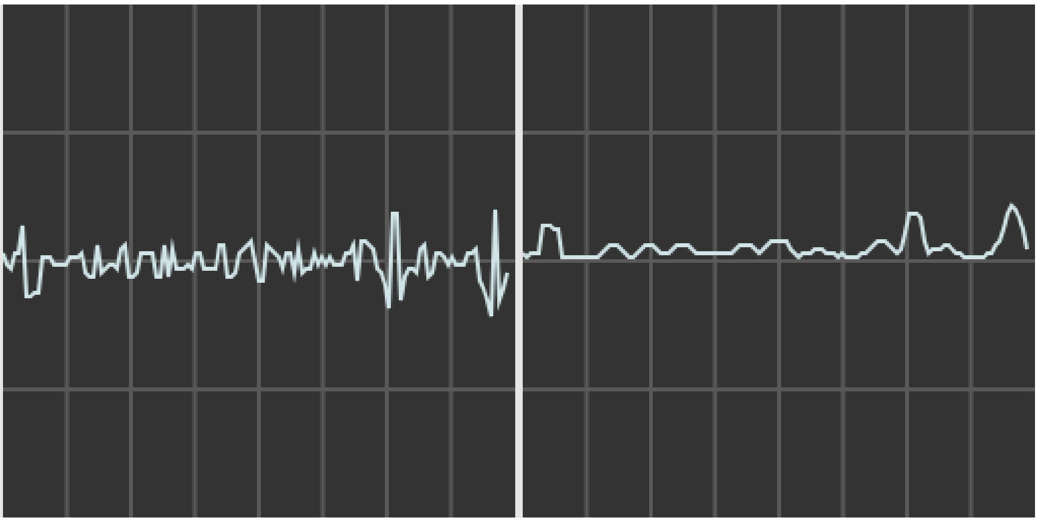


Figure 4.3: Two graphs from an envelope follower, on the right the amplitude curve of the input sound is displayed, and on the left the amplitude curve as applied to a 200 Hz sine wave.

These are three simple approaches to creating motion of sound sources in spatial audio applications. These can all be applied in an amplitude panned scenario as approached in *Hot Pocket* or to directly control sound sources in an encoded soundfield. All three approaches are interfaces to articulating the *gait* and *dynamics* of sound sources and to build complexity of the sound *mass*. In section 2.4 we saw that morphologies in nature contain traces of motion and the events that caused them. These shapes are articulated by motion and materiality, by their radiating patterns and by interaction with their surroundings. The motion processes discussed here, in combination with the following examples play into this. The remaining experiments in this section are not concerned with motion but rather with a sense of articulating space through the timbral dynamics of *mass* and *grain*.

4.3.2 Distance filter

A distance filter was designed to add a sense of "realism" to approaches of creating dense ambisonic soundfields. Ambisonic encoders allow the users to specify angles and relative gains of the sources, and we can add as many sources to a scene as our system will allow in terms of processing power. However, by only using adjustments of amplitude, there will be very little manipulation of depth of the individual sound sources. A distance filter can then be used to mimic air absorption, so that amplitude and the higher frequency content of the sound is attenuated as the source moves further away. This is illustrated in Figure 4.4 using noise. When the source is a certain distance away from the centralized listening space, it will be silenced. For example, people sound "duller" when speaking at a distance, because humidity, smoke particles and other impurities in the air will absorb higher frequencies (Howard & Angus, 2009). Where reflections cause comb filtering of the signal and change the timbre of the sound, the effects of air absorption on a sound is determined by the distance it has travelled rather than by its interactions with surfaces (Halmrast, 2011).

This distance filter is designed according to a formula by based on Richard Furse's software *VSpace*[9] that implements a 1-pole lowpass filter, where the cutoff frequency is related to the distance of the source to the centre of the sweet-spot. This approach to distance filtering has previously been implemented in the WFSCollider software[10] from the Game of Life[11] and is used as emulation of point-source distance. The IRCAM Spat-package, mentioned in section 3.5, implements a distance filter through the objects [spat5.air∼] and through the "perceptual control interface" [spat5.oper]. Likewise, version 3.0 of the ICST Ambisonics package, see section 3.5, implements distance filtering for sources inside and outside the center zone (see the *Distance_encoding* tab in the [ambiencode∼] helpfile for further information.

The availability of different distance filter emulations in a variety of software packages and libraries would indicate the wide-spread use and popularity of this technique among composers, sound artists, and sound designers.

[9] http://redmine.spatdif.org/projects/spatdif/wiki/VSpace
[10] https://github.com/GameOfLife/WFSCollider
[11] https://gameoflife.nl/en/about/about-wave-field-synthesis/

In MaxMSP the [biquad∼] object implements a second-order low pass filter that has been used in this instance, rather than the native [lores∼], which is a resonant low pass filter. Richard Furse specifies that "air filtering is less than first order so anything like this has to be an approximation to some degree".[12] A first order 1-pole lowpass filter can be built by introducing one sample of delay to the input signal (J. O. Smith, 2007).

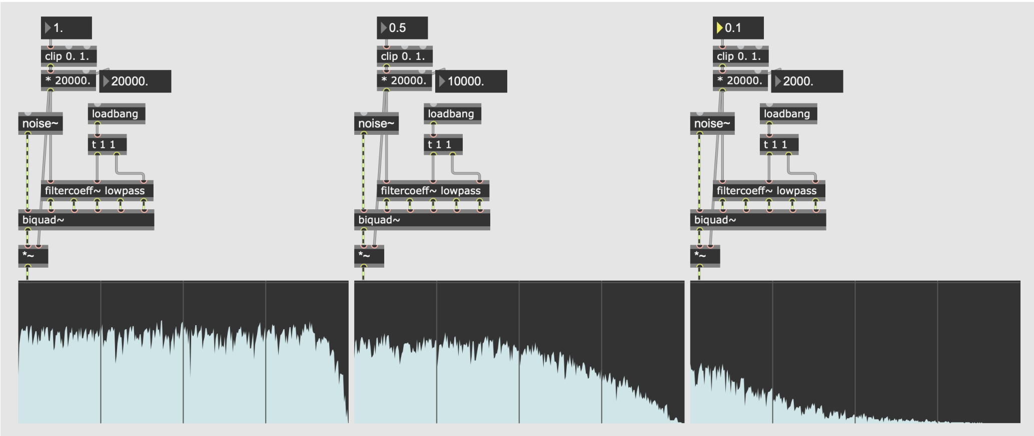


Figure 4.4: Three settings of the distance filter to roll of high frequency content, at $1. = 20000Hz$, $0.5 = 10000Hz$ and $0.1 = 2000Hz$.

All the adjustment we could make to a signal, as well as any medium the signal passes can be considered a filter, "when you think about it, everything is a filter" (J. Smith, 1985, p. 70). If no distance compensation is applied to sound sources apart from amplitude adjustments, they will generally be perceived as just being sounds of different loudness and not provide any depth to the sound image, or effective distance cues (Moorer, 1985).

4.3.3 Bandpass filters

In section 3.5, Robert Normandeau's approach to timbre spatialisation was briefly discussed. This is an approach where bandpass filters split a sound source into different blocks within a specified frequency range, these blocks are then sent to different speakers where the room (and the listener's ears) is used as summation of the resulting signals (Normandeau, 2009). A similar approach has been under development for a collaborative project, but has been postponed due to COVID-19 restrictions. This project sought to use live inputs from modular synthesizers that are then encoded in ambisonics and spatialized in real-time. To avoid the

[12] Private communication, 09.01.2019.

point-source problematic that has been discussed several times previously, Normandeau's approach to timbre spatialisation was reconsidered towards A-format processing. An input would be split into four discrete signals, and encoded using an A-to-B-format encoder, to mimic the encoding of files recorded with a soundfield microphone. A bandpass filter is a filter implementation which attenuates the frequencies above and below a defined passband (Reiss & McPherson, 2014). In Directional Audio Coding (DirAC), an input sound is analysed through division into discrete frequency bands through bandpass filterbanks in an effort to mimic the resolution of human spatial hearing (Vilkamo et al., 2009). The aim of DirAC analysis is to measure the direction and diffuseness of a soundfield in frequency bands.

This encoding would result in a spherical decomposition where the directional pressure gradients would consist of the narrowly filtered input signal. Each stereo output from a synthesizer would result in two sets of encoded B-format signals. Before decoding within a speaker array, a series of soundfield transformations can be performed on the individual B-format sets, such as rotate, tilt, tumble, mirroring, or zoom (Malham, 1998) (for more information see section 3.6 and (Hollerweger, 2005)).

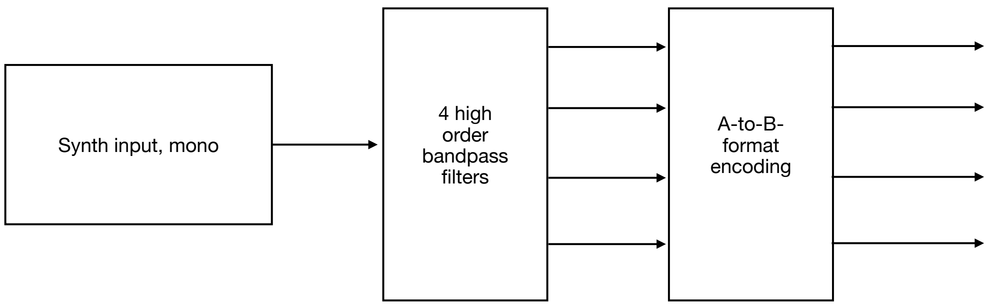
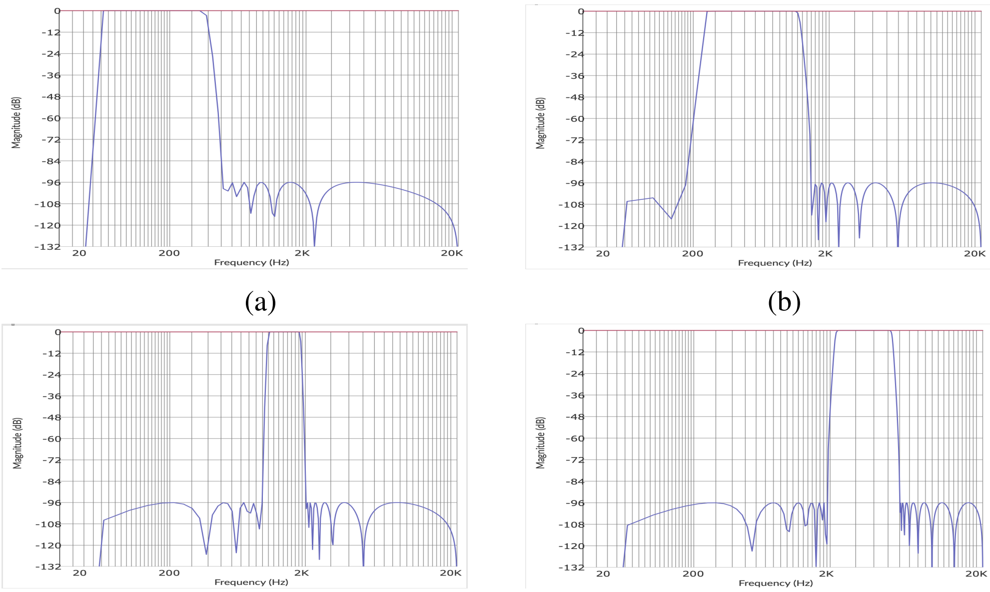


Figure 4.5: Monophonic input signal from a modular synth split through a bank of four high-order bandpass filters, before being encoded into spherical B-format.

In MaxMSP, higher order filters can be created through the [filterdesign] object, which creates a dictionary filter specification for use with the [cascade∼] object, which stacks a series of [biquad∼] to create the higher order filter. The [filterdesign] object implements specifications for, among others, Chebyshev filters. The Chebyshev specification creates a steeper roll-off by allowing ripples in the frequency response (S. W. Smith, 1999). In Figure 4.5 the filters are specified at fifteenth order and are close to brick-wall filters, meaning that the roll-off between the pass-band and the rejection-band is very steep. At this order there are visible ripples in the rejection band, but the ripples are at $-96dB$, which means that they are outside of the audible range.



(a) (b)

(c) (d)

Figure 4.6: Plots of four different high-order bandpass filters with Chebyshev topology for splitting a monophonic input sound into an "A-format" signal. The filtered signals will correspond to the W, X, Y, Z signals in an A-to-B-format encoding.

A similar approach to effects processing of encoded ambisonic signals has been proposed through the Ambisonic ToolKit (Lossius & Anderson, 2014), where four-channel, encoded B-format soundfield recordings can be decoded back to A-format for effects processing. The decoded signal can be manipulated using FIR, IIR filters, delays, or other linear processes. The four-channel signal can then be encoded back into B-format for further soundfield transformations.

This approach to soundfield synthesis is inspired by the "*idiomatic* approach" (Lossius & Anderson, 2014, p. 1339) to "think ambisonically", which fosters a *sound-field sound-image* approach rather than the discrete placements of sounds in a sound-space. To think ambisonically, we must consider the context where the sound exists. As ambisonics affords spatial thinking without considering loudspeakers, we are in a situation where the affordances of the sound landscape can guide the ways in which we approach spatial synthesis. The sound landscape is the imagined source and location of the sound, and through ambisonics we can represent this as a three-dimensional enveloping landscape for the listener. With a landscape comes morphological features, changes in structures and patterns of sounds and their locations. In this sense we can use the typomorphological framework to analyse and synthesize new sound objects and new soundfields based on our approximations of the sound landscape.

4.3.4 Reverberation

Reverberation is a naturally occurring acoustic effect, where an emitted sound is followed by several thousands of reflections from the walls, ceiling, floor, and furniture before reaching our ears and can be “the boon or bane of the acoustical performance of a room” (Long, 2005, p. 585). Approaches for the evaluations of reverberation are many, and several of the psychoacoustic criteria discussed in section 3.4.1 can be found in the literature discussing architectural acoustics.

However, this short section does however not deal with measured building acoustics, rather it looks at some experimental approaches to artificial reverberation. There are countless reverberation plugins on the market, but we will instead look at a few experimental approaches of how this can be explored through spatial audio.

IRCAM’s spatial audio library *Spatialisateur* (Spat) is built around a feedback delay network reverberation engine, encapsulated in the [spat5.spat∼] object (Jot, 1999). The reverberation engine implemented in Spat offers an interface for placement of sources in existing or virtual spaces. This can be controlled through the [spat5.oper] object, the “Perceptual control interface”, and provides control factors such as “source presence”, “room presence”, air absorption, along with reverberation, room size, and control over early and later reflections (this is discussed further in section 4.8).

Early experiments with artificial reverb were conducted by Manfred Schroeder and explored different approaches to the creation of artificial reverb; first a parallel set of four comb filters, which were summed into two cascaded allpass filters, as well as and approach which used a set of cascaded allpass filters (Schroeder, 1962). The parallel comb filters produce a series of echoes that when fed through the two allpass filters in series create decorrelations of the signal. An allpass filter passes all frequencies equally but introduces a frequency-dependent phase shift on the incoming signal (Roads, 1996). In Schroeder’s second proposal of five allpass filters in series, the echo density is multiplied by each filter.

Decorrelation creates a diffuseness to a signal through constructive and destructive interference (Kendall, 1995b) and aids in the creation of diffuse soundfields (Baalman, 2010). Uncorrelated signals played through loudspeakers tend to be perceived as discrete sources in different locations. If the signal are correlated, we experience the phantom image between the speakers, as discussed in section 3.5. The diffuse field created by decorrelated sources helps to avoid image shifts in the sound images, where an image *warps* in relation to the position of the listener (Nyström, 2018). Decorrelation of sound is used to make something sound "spacious", for example in audio production, a singer or instrumental performer can be recorded twice where deccorrelation occurs by the micro-variations between the two performances (Kendall, 1995b). Likewise, through allpass filters (Schroeder, 1962), four-channel comb filter reverberators (Stautner & Puckette, 1982), granular synthesis (Rolfe & Keller, 2000; Truax, 1998), and sub-band decorrelation (Potard & Burnett, 2004), diffuse sound images with a high apparent source width can be created. Decorrelation can also be achieved through the use of multiple delay lines to create artificial spaces (Oliveros, 1995).

In an experimental instance, a reverberation patch based on Miller Puckette's abstraction [*rev3*∼] from Pure Data,[13] has been explored. This is a 16-delay reverberator implemented as a feedback delay network, as seen in Figure 4.7, which creates a lush and spacious decorrelated sound. In the PureData-version this reverberator has four outputs, placed at the last four points in the delay chain and all the delay times are evenly spaced from 10 to 130 milliseconds. This has been used for experimentation with summations of different channels, along with point-source allocations of the different parts of the decorrelation network.

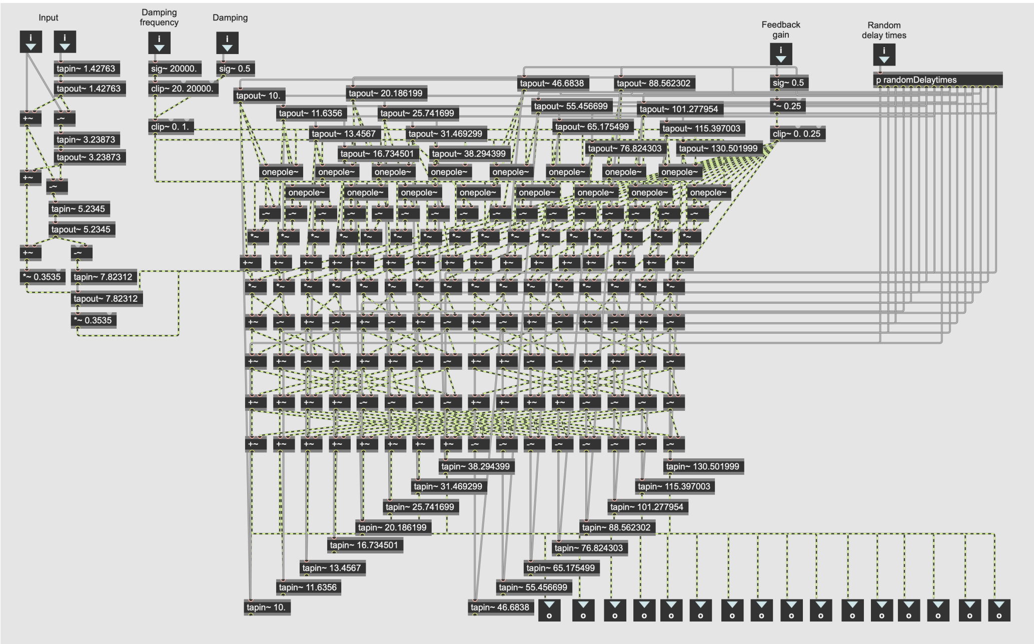


Figure 4.7: View of the decorrelation network from the ported [rev3∼]. In the top left corner is a cascade of delays is for controlling the pre-delay, before being sent to the decorrelation network. Input 3 is the damping frequency, input 4 is the amount of damping and input 5 is the feedback amount.

[13]http://puredata.info/

As the delays were evenly spaced, the amount of "interesting" information was very sparse,[14] and would often be covered over by the later reflections. I added a simple control feature to space the delays differently, as well as a randomization function so each part of the decorrelation network could be controlled independently. To exploit the uses of the decorrelation network, one output per delay chain was added, so this could easily be connected to an ambisonics encoder. In Figure 4.8, the sixteen outputs from the reverberator are distributed as individual points on a sphere surrounding the listener.

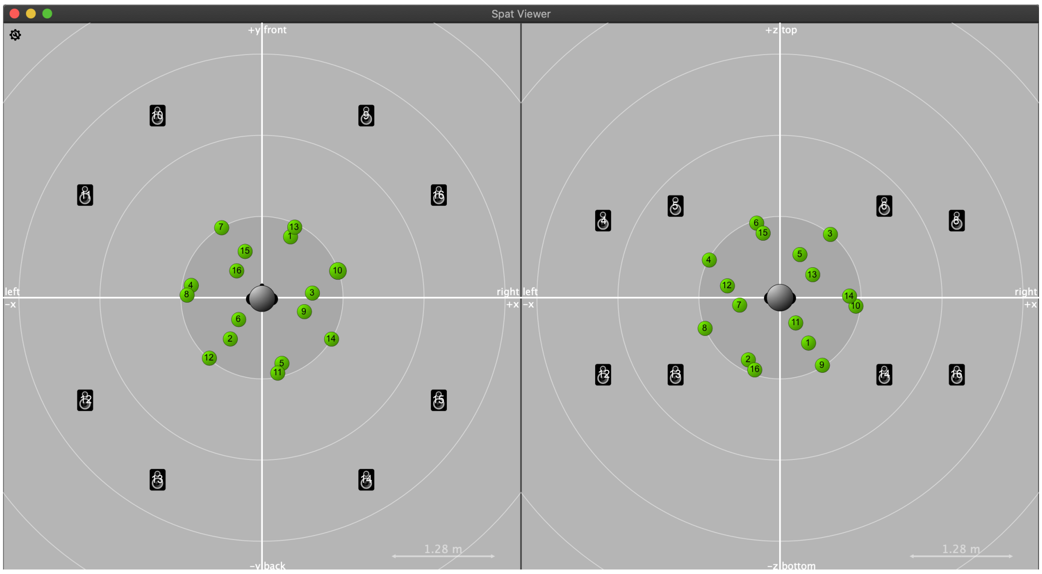


Figure 4.8: Sixteen channels distributed spherically around a listener, from [spat5.viewer]. The right image displays the distribution of sources around the x and z-axes, and the image on the left displays the distributions along the x and y-axes. In this instance each point in the decorrelation network can be separately encoded and moved around a listener as a discrete sound source.

This reverberant sound is created before encoding, which still affords the opportunity of applying ambisonic domain processing of rotation, tilt, beamforming, blur, and the like to the spatial signals. The flexibility of this extended approach allows for the treatment of the different parts of the decorrelation network as discrete sources, which can be moved around the listener in all directions. The motions of all the different components of the decorrelated, reverberant sound can be moved following the *dynamic* profile of the sound, can be clustered to increase

[14] After a private conversation with Tor Halmrast, I realized that the "interesting" information in the reverberation, the information that tells you something about the source and the space the source occupies, happens in the early reflections.

the perception of *mass*, can be iteratively moved in space to expose the *grain* of the sound, among many others.

The initial experiments with the encoding of this reverberation approach led on to a different approach to distributing delayed copies of a signal within a soundfield, through a "rain-delay". This is named after a eurorack synthesizer module called "Rainmaker",[15] which uses a 16-tap delay to create "light showers, or something more torrential". This approach uses the input of a monophonic sound, which is fed through a delay bank where it is split into 24 discrete components delayed randomly between 0. and 1000. milliseconds, illustrated in Figure 4.9. Each tap on the delay line is fed through an individual distancefilter to create depth before encoding all the signals as 24 discrete sound sources in a soundfield, similar to the spherical covering in Figure 4.8.

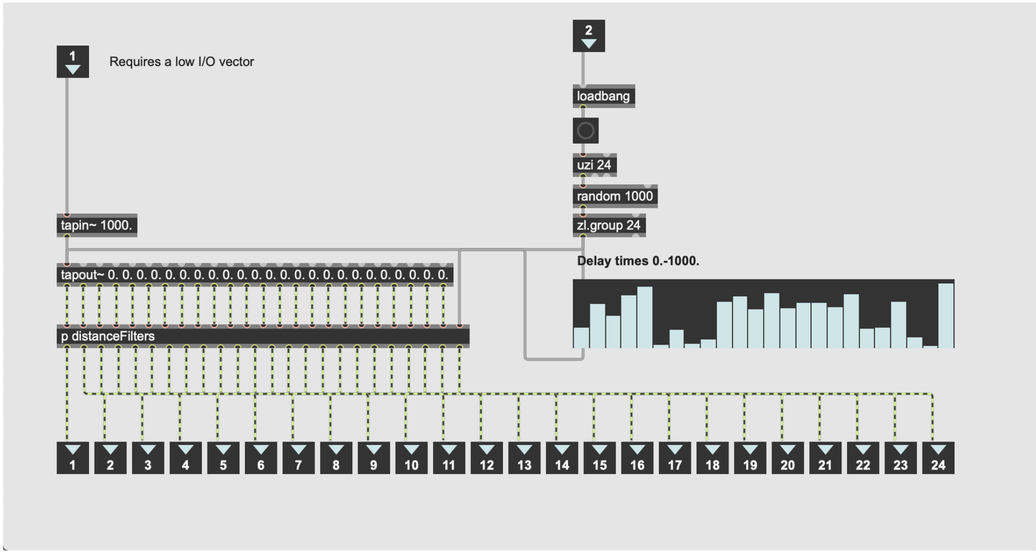


Figure 4.9: A mono sound is fed through a 24-tap delay line where each delay time can be set individually or randomly. Each output from the delay line is fed through a separate distancefilter to create depth.

The rain-delay approach creates spatial effects and diffuse soundfields without using reverberation. The delay times can of course be increased, but too great a separation of the individual sounds in time will create audible echoes and lead to image shifts.

Throughout this section we have briefly discussed some experimental approaches to spatial audio processing that covers motion, filters, and reverberation. The goal for this section has been to show, next to the case studies, that typomorphological relationships can be expressed through space and spatialisation, if we consider an

[15] https://intellijel.com/shop/eurorack/cylonix-rainmaker/

idiomatic approach to spatialisation and treats the soundfield as a sound-image, rather than discrete points in space. The individual effects parameters that are used to create the spatially distributed sources can each be harnessed for their potential morphological features, where an envelope follower can be used to control the mass/facture of the sound, a reverberation can alter the mass profile and melodic profile of the sound, the grain, and the gait can be manipulated by different motion processes and splitting of the sound. As already stated, based on our subjective listening and evaluation, these different criteria can be subdivided and sub-subdivided for the different feature dimensions.

The next five sections present different projects in which I have been engaged as an artist, sound designer, and programmer. These different projects highlight different theoretical aspects from the previous chapters by providing some insights into the practical applications of spatial audio applications. Specifically, I want to focus on the methods and technology used to create the work and their relationship to the theoretical discussions, rather than focusing solely on the artistic outcome. This will bring us into the first case study, an experimental approach to sound design that resulted in the a composition in higher order ambisonics titled *Tessallation Rift*.

4.4 Case 1: *Tessellation Rift*

This first case study discusses a process of experimental sound design which lead to the development of the composition *Tessellation Rift*, which came out of a period of sound and spatial design experiments. These experiments sought to examine ways to place the listener *inside* the sound, as a means to *dwell* inside the sound, as noted in the introduction in chapter 1, as a way of exploring the idiomatic approach to ambisonic sound images from only synthesized material (Lossius & Anderson, 2014). This project set out to examine ways where a listener can experience the shape and the morphological features of the sound, as well as hearing sounds in relation to *outward* and *spatial features*. The experiments were carried out in NOTAM's Studio 3, a 24-channel studio, over a period of 4 months which provided time for experimentation where the spatial structures could be examined over a multi-channel array and to evaluate the results and re-synthesize both the sound materials and the spatial structures, as discussed in section 1.2.1. These experiments were explored through Schwab's ideas on the *transposition*, and specifically where the experimental context was created for the sake of gaining knowledge about several instances of a project (Rheinberger, 2018), in short, making and knowing derived from practical investigation.

These experiments involved examining relationships between foreground and background sound; simple, impulse-like sound material and complex textures; and variable order ambisonics, where the material is encoded from 0th to 7th order. The composition which resulted from the experiments was performed as part of the Puls-project by the Nordic Culture Fund for concerts in 2019,[16] curated by Trond Lossius. The piece has been performed at Østre, Bergen, Norway, and FRST Festival in Gotland, Sweden, as part of New York City Electroacoustic Music Festival 2020, New York, USA, and will be performed at the International Computer Music Conference 2021, Santiago, Chile.[17]

These experiments arose from an interest in exploring the creation of timbrally complex sounds out of as simple sounding material as possible.[18] In this instance, the only sound source is the impulse. The impulse is a signal where all samples are 0 apart from one that is 1, and is a discrete time-amplitude fluctuation that we hear as a distinct "click" (S. W. Smith, 1999). An impulse generator (analogue or digital) emits a series of impulses at a specific frequency, which serve many functions inside and outside the laboratory as a means of testing the *impulse response* of a system. Grains, clicks, and impulse-like sounds do not sound like anything on their own but when sequenced in time and density they become timbrally significant (Farnell, 2010). This bears a resemblance in approach to Trevor Wishart's *Imago* (2002) (Harrison et al., 2004) and Horacio Vaggione's *Harrison Variations* (2002) (Harrison et al., 2004), which both use the same sound source, a one-second recording of two wine glasses clinking together (Roads, 2015).

However, the principal inspiration for pursuing experiments into using impulses as sound material arose from an examination of an illustration of the synthesis methods used in Karlheinz Stockhausen's 1960 composition *Kontakte*. This piece was based only on filtered impulses, which used impulse generators, band-pass filters, preamplifiers, tape recorders, plate reverberators, and feedback loops (Roads, 2004, p. 70). In *Kontakte* sequences of squared impulses were organised where the delay between the impulses and amplitudes were determined by serial permutational laws (Di Scipio, 1994). Another historical inspiration can be found in Iannis Xenakis's *Concrete P.H.* (1958), which is an early exploration of granular textures where the sounds of burning wood embers were cut into one-second fragments (Roads, 2004) and montaged to create expansive sweeps inside the curved construction of the Philips Pavilion (Manning, 1993).

[16] https://puls.nordiskkulturfond.org/en/

[17] The conference has been postponed from 2020 to 2021.

[18] This approach has also been explored through compositions such as *Ice Feathers* (2018, performed at the International Computer Music Conference (ICMC) in Daegu, South Korea), *Droplets* (2018, performed at the New York City Electroacoustic Music Festival (NYCEMF), New York City, USA), *Forest Glade* (2019, performed at ICMC-NYCEMF, New York City, USA), *Footprints of Creatures* (2019, performed at Sound/Image, University of Greenwich, London, England). These pieces were spatialised using the approaches described in section 4.5.

Similarly, in the composition *Textonics* (2018)[19] Erik Nyström was interested in composing with texture through *textons*, time-finite microsounds, and *filaments*, continuous sounds (Nyström, 2018), which create both differing spectral density and motion through combinations of sounds.

Examinations of filtered noise signals processed into impulses and textures can be found in Kerry Hagan's *Morphons and Bions*[20] (2011), where noise is not just a signal but also something which conveys meaning through textures and differences between simple impulse-like sounds and larger densities (Hagan, 2017). These experiments can also be heard in *nyx* (2017)[21].

Tessellation Rift explores an experimental spatial sound design, specifically as an exercise between the near and diffuse field synthesis of higher order ambisonics, and focuses on features between the soundfield and the sound object. These features illustrate listening to a sound from the inside (the listener being placed inside the sound object) and hearing an object from the outside (the sound being projected to the listener from an angular location). It has been stated that the nature of ambisonics affords the perception of envelopment to be extended further than other spatial audio systems through "the sensation of being 'inside' the sound" (Barrett, 2010, p. 3). However, it is difficult to confirm such a statement and there are multiple spatialisation approaches that could achieve this (see 3.5). However, one of the strengths in ambisonics is the ease with which multiple processing techniques can be combined, for example decorrelation and spectral processing can be introduced to sources before encoding.

The acousmatic situation "dictates that a sound be described less in terms of its origin than in terms of its heard morphological qualities (form, mass, profile and so forth)" (Nattiez, 1990, p. 92), a stance which should have been well established at this point throughout this text. The impulses themselves are not the focus of the listening attention, as the heard material bears little resemblance to clicks, rather the sequences in time, the associative timbral qualities of the sound, and the spatial densities of sounds are of importance. The perceived *mass* of the sound is directly proportional to the perceived spaciousness of the soundfield and to the apparent width of each individual spatially located sound source. Here, we perceive the extent of the sound landscape through the size and mass of the objects; the multiplicities of locations, distances, and the proximities between the sounds; and through the perception of the sounds and the space they exist in along with the smearing in time and space.

Implemented in MaxMSP and drawing on Stockhausen's work with *Kontakte*, the sound synthesis is based on the simple premise that a generated impulse can be

[19] https://soundcloud.com/erik-nystrom/textonics
[20] https://soundcloud.com/kerryhagan/morphons-and-bions-stereo
[21] https://soundcloud.com/kerryhagan/nyx

shaped through a filterbank of narrow bandpass filters. This signal is then delayed and split into two separate feedback paths, feeding back into the delay line. In MaxMSP the object pair [tapin∼] and [tapout∼] are the inputs and outputs to a delay line, where the first object receives a signal and copies it into the delay line, and the second object allows you to tap into the delay line at a specific time. In this instance, an external signal-rate scrubbing effect is applied to the [tapout∼], which controls the changing position of the "playhead" at signal-rate, creating glitching fluctuations and rapid changes in pitch. In audio editing, "scrubbing" is an interaction where the user drags the playhead across the waveform to hear it. Most of these functions are randomized to create unpredictable sound material.

After the generation of material, it was then subjected to repeated listening, and sections were edited down and then repeatedly listened to again to assess the possibilities of spatially composing with the material.

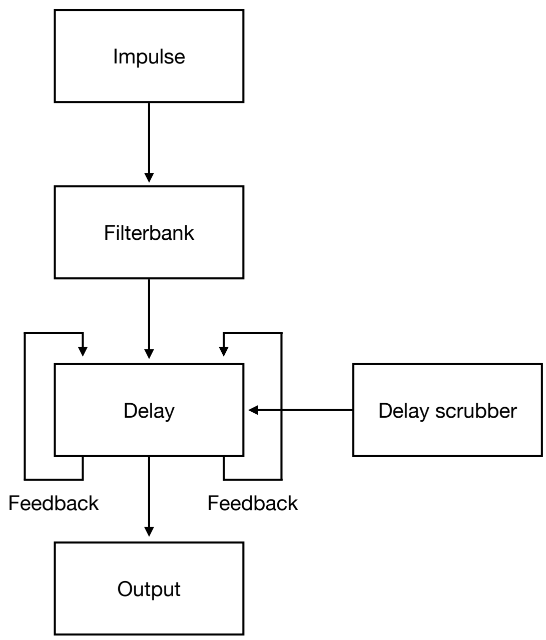


Figure 4.10: Diagram showing the processing chain of the impulse sound sources, where impulses are processed through a bank of bandpass filters, a delay line with stereo feedback and signal-rate scrubber, before the sound is sent to the output.

All the sound material was synthesised and the resulting material was abstract in a non-representational sense, where the considerations of how the sounds could be combined to create a sound landscape became of the utmost importance in the structure of the composition. As there are no concrete, "real-world" references in the sound material, the *imagined* sources of the sounds become very important through the highlighting of potential recognition of the sound sources and how they exist in space. Throughout the composition, the sounds constantly shift between the impulsive and the sustained, between the *mass* and its *duration/variation*, and the *gait* between the foreground and the background. The sounding material and its *grain* push the shifting densities of the sources around the soundfield (see figure 4.11 which illustrates the attacks of the impulses, along with the mass they produce).

The size of a sound and its smearing in time and space, requires more than simple points in space projected from a loudspeaker. The complex spatial arrangement of sound is then, as previously discussed, timbre (Gottfried, 2016). This smearing of sounds in time and space contributes to a heard morphology that is more than just the creation of diffuse borders around a sound. This smearing can also refer to the streams of auditory scene analysis (Bregman, 1994). We do not perceive individual sounds of an event but one sound: when someone is walking over gravel, we hear it as one sound despite each footstep being a separate sound. In this sense, the sound sources themselves are not of interest, rather the morphological traits of the sounds and how we perceive them in space is what we use to derive a meaningful representation from it. This contributes to the experience of a functioning sound landscape.

We cannot touch sound, but it is something that is referred to as a powerful physical being (C. Cox, 2017). This "physical" being is the sound's *shape*, the shape is one of features that makes up our cognition of the object along with its references to a site and to the models we use to analyse and synthesize the objects. Through the soundfield we can perceive different surfaces of sounds surrounding us, through the differences in microstructures of the sound, through the sound's *grain*. This can involve perceptions of sounds as "hard to soft, luminous to reflecting, illuminated to shaded, high to low reflectance, uniform to speckled reflectance, smooth to rough texture, opacity to transparency, dull to shiny, and hot to cold" (Gibson, 1982, p. 152). Gibson's list here refer to *visual* features of surfaces, but here we refer to the potential *metaphorical* descriptions of sounds as we hear them in space.

Through the sound's *grain* and its *gait* we can construct *nested time-frames*, which Wishart discusses as the basis from where we construct rhythm and large-scale form (Wishart, 1994, p. 19). However, the sound object and study of sound fragments offer a different focus from the study of large-scale form, and affords us a certain level of resolution to be part of our focus (Godøy, 2006). These nested time-frames are not simply matters to create rhythms, rather different time-frames create different types of densities, differently shaped spaces, and different types of surfaces.

Throughout *Tessellation Rift* reverberation is not used as a "space-creator",

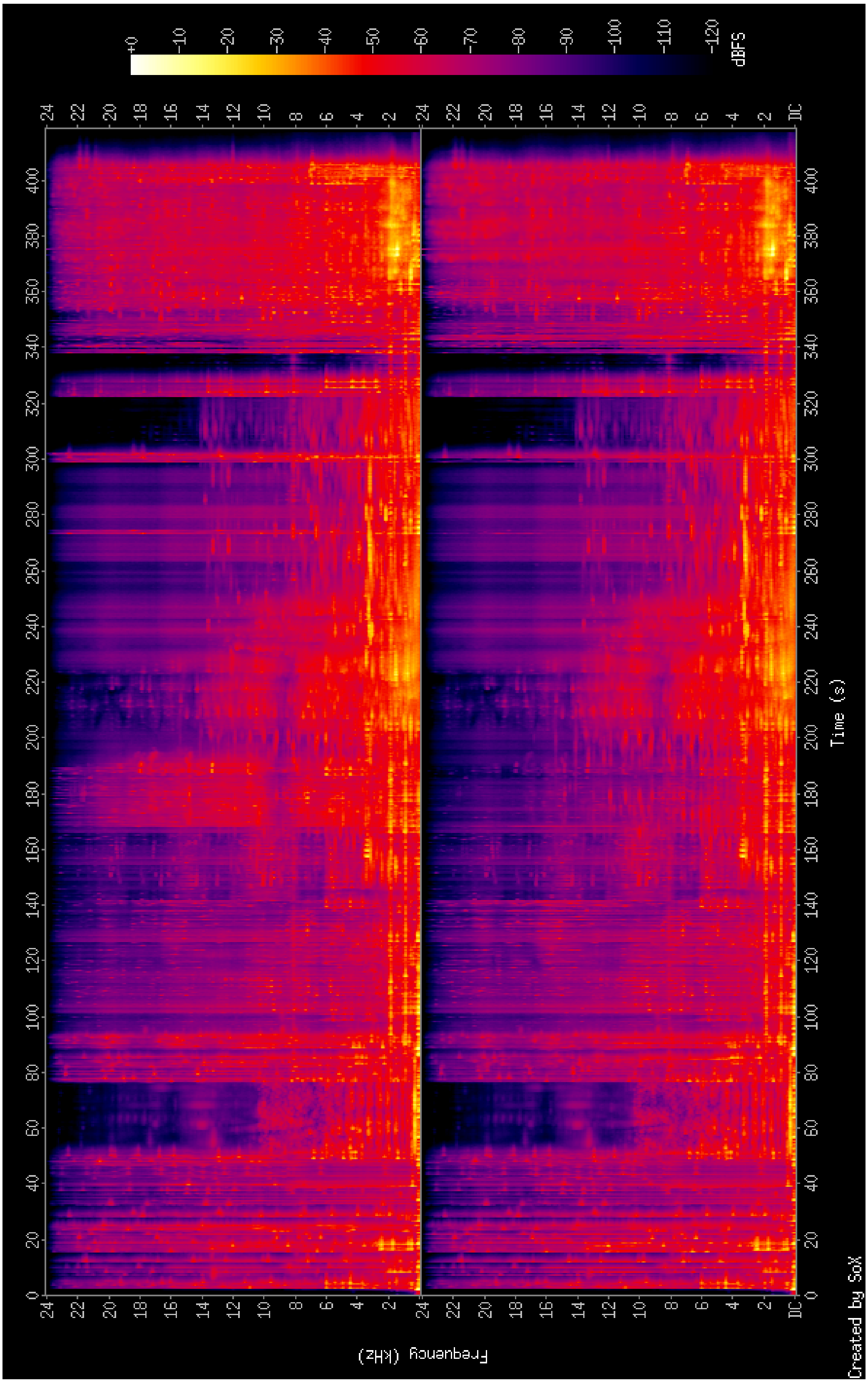
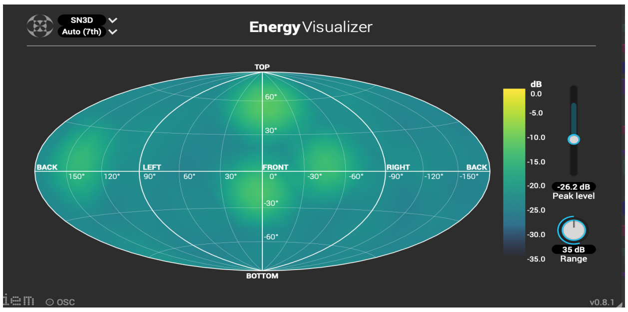
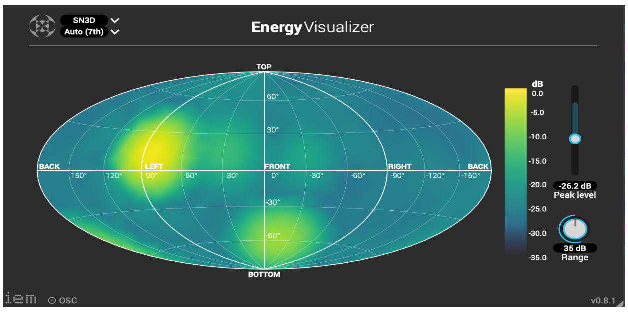


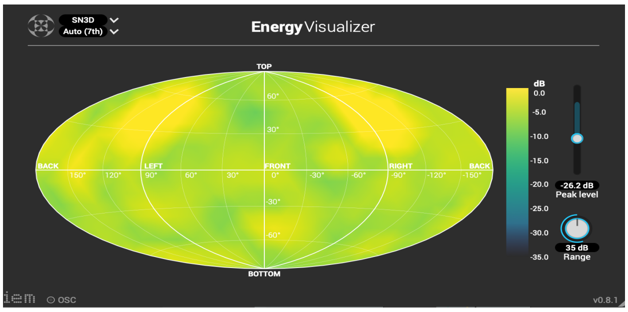
Figure 4.11: Spectrogram of *Tessellation Rift* decoded to stereo, where time is displayed in the x-axis, frequency on the y-axis and amplitude on the z-axis (colour graph).



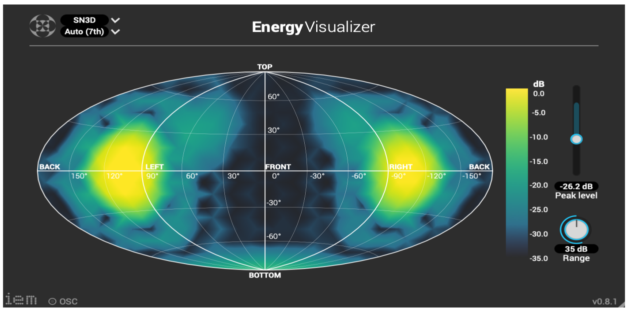
(a)



(b)



(c)



(d)

Figure 4.12: Four different soundfield energy visualisations from *Tessellation Rift*. In these visualisations the localized sources are clearly visible in the soundfield as circular points and the diffuse field caused by the reverberation encapsulates the sources and fills the soundfield. In image (a), the sources are diffuse and retreated into the background while in (b) the sources are moving into the foreground. In image (c) the background fills the entire soundfield and masks the sources, while in image (d) the sources radiate sound that fills nearly the entire soundfield. Visualisations from the *Energy visualizer* from the IEM ambisonic plugin suite (https://plugins.iem.at/docs/plugindescriptions/#energyvisualizer).

rather it is employed to create granular-like textures through which we can perceive the diffuse field, a “bounding” surface where the discrete sound sources are placed, as illustrated in Figure 4.12. Synthesised sound are not captured from a real-world sounding space and will often not contain the richness of the naturally occurring acoustical effect. Therefore “certain synthesized sounds have little or no intrinsic spaciousness. These acoustically “dead” signals can be enhanced by spatial panning, echoes and reverberation processing” (Roads, 1996, p. 472).

The final montage of the composition was done in Reaper, using the IEM-plugin suite for ambisonic encoding and decoding. Stereo-sources were encoded using the *StereoEncoder*[22] or the *Directivity Shaper*[23]. The Stereo Encoder has a *width* function which allows the user to separate the two input channels, this affords an effective way of avoiding the single point in space, where the same or similar sounds can be encoded together in the soundfield to create sources that have a larger mass and can occupy a larger area of the sound landscape. The stereo encoder was chosen because the source material was combined into stereo tracks, where the individual channels were shifted in time and frequency. The Directivity Shaper allows the user to split the input sound into different frequency bands and place them in separate parts of the soundfield, akin to the previously discussed “timbral spatialisation” (Normandeau, 2009). However, the strength of this plugin can be found in its “Order and shape” function, where the individual frequency bands can be fractionally encoded from 0th order to the order specificed in the project (in this case 7th). For *Tessellation Rift* the StereoEncoder was used to place sources in the three-dimensional space surrounding the listeners, and then processing the encoded ambisonic signal through an *FdnReverb*[24] plugin. The same sources were sent through the Directivity Shaper with orders varying between 0th and 7th.

An FdnReverb is a Feedback Delay Network reverberation that simulates reverberation through a process using delay lines, filters, and feedback connections (Välimäki et al., 2016). The delay and feedback matrix has a strong effect on the time-domain distribution of the results of the reverberator, which creates a diffuse field where the source is difficult to localise. In this instance, the diffuse field of the reverberation fills the entire soundfield, as seen in the first three images in Figure 4.12. This produces an externalised sound signal. The advantage of using this method is that we can create a blurring effect of the boundaries of the sounds without reducing the order of the ambisonic signal.

In section 3.6 we discussed the hierarchical orders of ambisonics, where a higher order encoding produces a signal of a higher resolution. This leads to

[22] https://plugins.iem.at/docs/plugindescriptions/#stereoencoder
[23] https://plugins.iem.at/docs/plugindescriptions/#directivityshaper
[24] https://plugins.iem.at/docs/plugindescriptions/#fdnreverb

sources that are easier to localise but also emphasizes the point-in-space, which we should aim to avoid. If a signal is encoded into seventh order, then by applying a *blur* to the signal before decoding (Carpentier, 2017) we reduce the order of the signal, making it *blurred* in the soundfield and ambiguous to localise and more diffusely spread (this will also be discussed in section 4.6, and is illustrated in Figure 4.20).

A diffuse reverberation that occupies the entire soundfield has a large *mass*, not just in terms of its different pitch regions but mainly in terms of the variety in colour and thickness of the sound along with the interplay of *grain* and *gait*, different textures and piercing attacks move and shift through and around a space to create textures that oscillate between the foreground and the background. The sonic features of the shapes and motions contained within this composed work can effectively be analysed and described with support of the typomorphological framework. The pattern formation we perceive when listening implies all the possible relationships the sounds have to each other, their possible sources and how we hear them in relationship to the space they occupy - the sound landscape.

The *gait* of a sound as it is moving through space is a primary focus in the next case study, as a solution to an unpredicatably moving sound source through an exhibition venue.

4.5 Case 2: Stochastic spatialisation for *Hot Pocket*

This case study discusses an approach to spatial programming for the exhibition *Hot Pocket*. Uncanny and magical, *Hot Pocket* invited the audience into a cave or the belly of some unidentifiable animal. The exhibition showed a completely transformed Museum of Contemporary Art in Oslo. All the floors and walls were covered in carpet, the light faded slowly up and down, and next to a floating surrounding sound and fluttering of birds around the ceiling, the audience could experience several objects, photographs, and videos of trolls and other mystical creatures as a *gesamtkunstwerk*.

The exhibition *Hot Pocket* by the Norwegian artist Tori Wrånes at the Museum of Contemporary Art in Oslo used a 35.4 channel setup, adding different densities of loudspeakers to all the rooms in the exhibition venue, as illustrated in Figure 4.13. The overall aim for introducing a large multichannel sound setup in the exhibition space was to “fill” the space with sound, and this was implemented to create continually moving and unpredictable spatial trajectories of sound. For this we explored a stochastic approach to spatialisation, as stochastically generated amplitude-panned trajectories. A brief overview of the project has already been published (Holbrook, 2018).

The spatialisation was implemented in MaxMSP and a single patch ran in

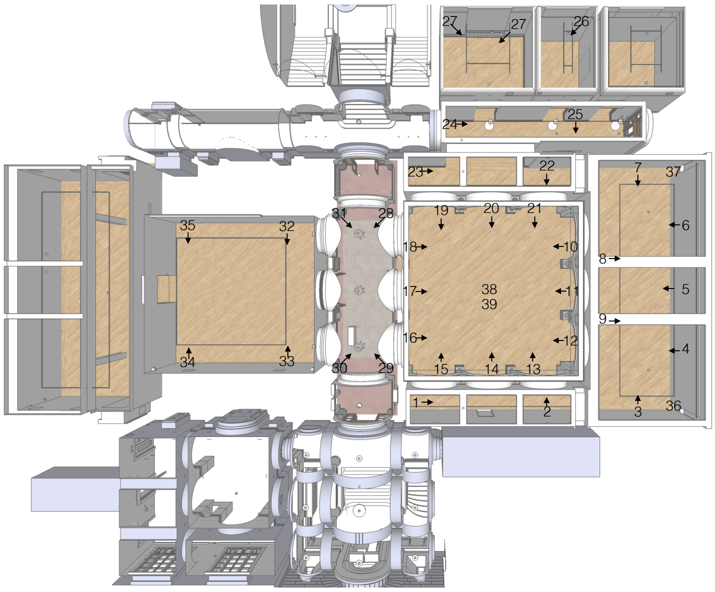


Figure 4.13: Plan view of speaker placements in the exhibition venue at the Museum of Contemporary Art. The arrows indicate the direction of the loudspeaker. Speakers 36-39 are subwoofers, where 38 and 39 were located above the ceiling.

the background taking care of sound design, spatialisation, and lighting triggers. There were certain challenges when working with a space of this magnitude: first, the exhibition venue consists of a series of rooms with dramatically varying sizes and ceiling heights; second, the audience would be in constant motion when moving around the space to primarily look at the works on display; and third, all cables for speakers and technical equipment were hidden behind carpet, so any mistakes in the initial setup would be difficult to remedy.

4.5.1 Approaches

The sound programming for the exhibition was conceptually devised with the aim of creating constantly moving sound sources throughout the room in an organic fashion, rather than creating simple circular trajectories. Several existing spatialisation algorithms were considered but none were found to provide an effective interface towards this arbitrarily arranged sound setup. Four primary approaches to spatialisation were considered: Vector-Based Amplitude Panning, Distance-Based Amplitude Panning, Ambisonics, and Wave Field Synthesis. All four were decided against, primarily due to the spread of speakers throughout the entire floor, as well as the desire for a continually moving sound "effect". The basic premise involved in pair-wise panning between two loudspeakers is a very flexible means of creating moving sounds through a space.

(a) (b)

Figure 4.14: Two views from *Hot Pocket* at the Museum of Contemporary Art: (a) is a view of Banksalen, showing the circular stage and the overhead light, and (b) is a view to the side of the circular stage in Banksalen, with parts of one of the sculptures in the exhibition. Loudspeakers can be seen on the columns on the right. Photo: Annar Bjørgli/Nasjonalmuseet (used with permission).

Of the approaches considered, wave field synthesis would prove very impractical in this regard, as the visual nature of several rows of loudspeakers would be too intrusive in the space, as well as time and financial limitations on building an array of hundred of loudspeakers. Ambisonics would also prove impractical, as the various parts of the exhibition space would need individual encoders and decoders, and there would be a problem with creating the motion around the space as needed. Likewise, speakers would often be placed far apart, elevated only, and in several places there would only be one speaker "filling" a space. Vector-Based Amplitude Panning was quickly dismissed as it is dependent on a complete circle or dome to effectively create the phantom images within the triangulation of loudspeakers. Distance-Based Amplitude Panning at first stood out as a good solution for creating rich textures and layers floating through the space. For this project we aimed for several layers of point-sources moving individually through the space, and hence deemed Distance-Based Amplitude Panning to not be effective. We aimed for, as much as possible, a sense of sounds travelling in clear and noticeable directions rather than an ambiguous diffuseness, enabling the audience to clearly localize sounds throughout the space.

A multichannel, "composed" approach to the sound design could have been easily achieved by assembling a soundtrack using a digital audio workstation with multichannel capabilities, like Cockos Reaper.[25] This approach would mean that the same sound design with the same spatial distribution would have been heard throughout the space each day. Instead we aimed for an approach that would create an ever-changing sonic appearance of the exhibition, with varying densities of sound at any given time during the day.

In this instance, the sound did not have an end-goal as a sound installation but was a supplement for the rest of the exhibition, and this posed challenges as to the placement of speakers and other technical equipment which had to be inconspicuous (Lossius, 2008). The sound design and the spatial spread of sound must not come into conflict with or disturb the art works on display and at the same time the speakers should not be obstructed to prevent the sound from radiating into the space.

4.5.2 Implementation

The setup for the exhibition was designed for an arbitrary amount of speakers and in total we ended up using 35 speakers and four sub-woofers, all produced by Genelec. The National Museum has acquired a large number of speakers through years of producing exhibitions on site and for travelling exhibitions. Although the speakers were by the same manufacturer, the collection a range of different models (6010, 8010, 8020, 8030, 8330, 7050) and from different series. The difference in frequency response from the smaller to larger speakers was considerable but for the purpose of this exhibition this was considered to be a benefit rather than a problem. The differences in colouration and dynamics between the speakers were favourable and mirrored concerns with loudspeaker orchestras. The speakers were distributed as evenly as possible through out the space to provide an even distribution of sound all around the different rooms. Two of the subwoofers (36 and 37) were placed in the exhibition space and two were placed above the glass dome in Banksalen (38 and 39) (for placements see Figure 4.13).

A Mac mini running MaxMSP was used for the central control in the exhibition. The sound was distributed to the speakers via two MOTU 24Ao soundcards, each providing 24 output channels over D-Sub to XLR. Lighting control data was sent as MIDI data to the grandMA 1 Ultra Light console via an M-Audio 2x2. Where possible, the speakers were hung at 3 metres above the floor, angled down at 45 degrees. The exceptions were speakers 10-21 which were hung at approximately 6 metres height, and speakers 28-31 which were at 2.5 metres. The area under the glass dome was fitted with 12 speakers and was seen as a "main" focal

[25] https://www.reaper.fm/index.php

point in the room around a large revolving circular stage. This stage would also be used for a series of performances throughout the exhibition period where all the speakers installed in the exhibition would be used.

As the audience would not be seated, any form of distribution which is dependent on a sweet-spot would not work. Rather, inspiration was drawn from Iannis Xenakis' approach to dynamic stochastic synthesis (Xenakis, 1992) and random walks were used as a method for creating spatial motion. The random walk paradigm allows for greater or lesser degrees of control, determined by how "random" one wants the distribution to appear: "The fluctuation speed of a parameter is directly proportional to the step size of its random walks: the smaller the steps, the slower the rate of change in that parameter" (Luque, 2009, p. 79). With a high step size the distribution would appear completely random and fluctuate around the room unpredictably. Other approaches to stochastic spatialisation have been explored as a means to spatialise texture streams through amplitude panning (Hagan, 2017).

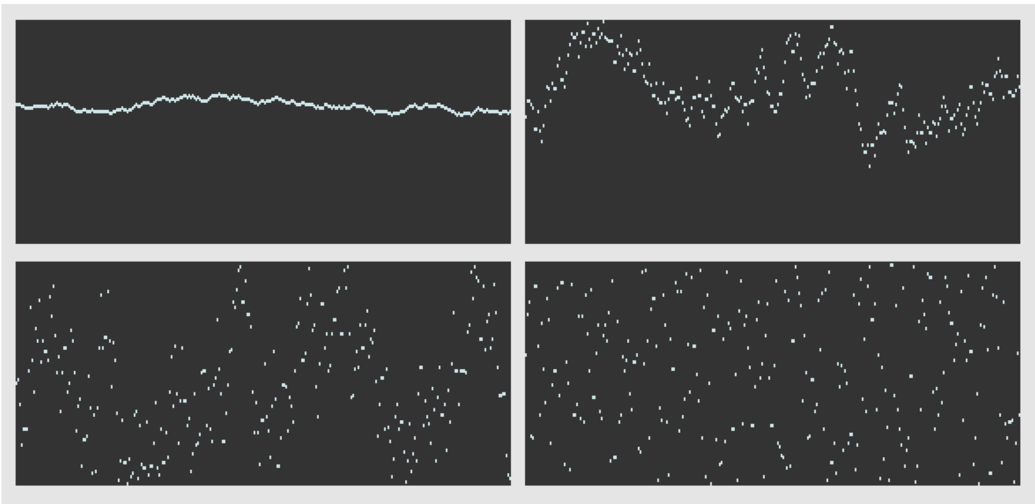


Figure 4.15: Different distributions of random walks. As the step size increases, the data becomes more randomly distributed. The differences in step size allowed for "tuning" the motions through the exhibition space, where, for example, the distributions controlling amplitude would be kept low to avoid large jumps.

To create shifting and moving sound scenes throughout the entire floor of the exhibition, the stochastic trajectories were defined from three data points: *speaker number*, *transition time*, and *amplitude* as the necessary values (see Figure 4.16). Each entry in the table is read sequentially, as a step-sequencer. A set of 256 values per data entry was created at each iteration. The amplitude specifies a value between 0. and 1., the speaker number would be between 1 and 35. Speakers 36-

39 were subwoofers and played a special role in the sound programming. The transition time defines the cross-fade time between channels. This is the time it takes the sound to move from one channel to the next; as an example, a defined transition time of 1389 milliseconds is the time it takes the sound to transition from channel 10 to 11. Once the sequencer finishes the transition to channel 11, a new transition time and channel is defined, effectively meaning that the sound is always in motion between two channels.

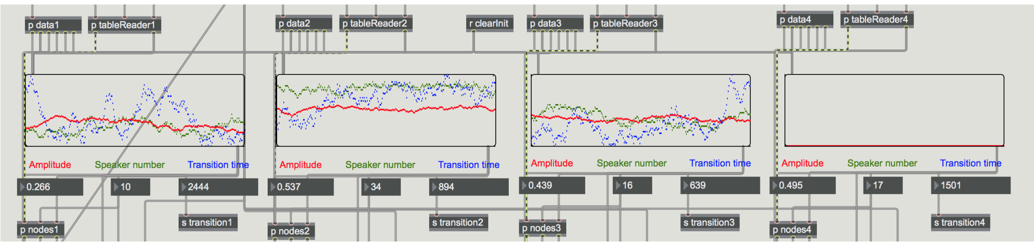


Figure 4.16: A view of the spatialisation sequencers, showing curves for amplitude (red), transition time (blue) and speaker number (green).

The spatialisation uses the native MaxMSP object [matrix∼], which operates as a simple routing and mixing matrix. The object can take any number of inputs and route these to an output of choice. In this case, there was one input per matrix and 35 main outputs to loudspeakers and four outputs for the subwoofers. The [matrix∼] object also enables a variable, linear gain for each output, defined in the range $0. - 1.$ If this is not enabled, each output is set to unity gain and an audible click is heard with each change of channel. The @*ramp* attribute defines the cross-fade time between the channels. The variable gain along with the cross-fade between channels afforded flexibility for the stochastic motions. Sound was distributed between speakers using pair-wise panning, based on stereophonic panning where the relative amplitudes between channels are adjusted to create a phantom image between two speakers. In many instances for the setup used for *Hot Pocket*, the speakers were too far apart to create a phantom image. Instead it has the benefit of creating a continually moving sound between the speakers, as the sound is always in motion between two speakers.

The sound design for the exhibition was approached in several layers. The main lights in the exhibition space faded up and down slowly, similarly the stochastically generated sound distributions moved around the exhibition space. The sound material used for these distributions were different wind recordings, which created very different timbral textures depending on the motions through the space and how many sounds were present in the same area of the exhibition space.

- At every hour, a foghorn sounded from a randomly determined speaker. As this first foghorn faded out, a second foghorn sounded from a speaker at the other end of the space as a call and response. These two foghorns were then followed by the sound of deep breathing played through speaker channels 38 – 39, the sub-woofers located above the glass dome. This breathing sounded like a large, sleeping animal, and added to the feeling of being inside a cave or the belly of a large animal. The amplitude curves of the breathing were tracked to control an orange light fading up and down, also above the glass dome.

- At quarter past every hour, a lightning strike appeared in the glass dome but carried no sound.
- At every half hour, a flapping bird circled the "main" area of the exhibition around the circular stage, in channels 10 – 21. This flapping was accompanied by a circling white light in the direction of the flapping bird. This circular trajectory was generated using a similar approach as the rest of the moving sound in the exhibition. Each time the circling bird was triggered, there was a 50% chance of the bird moving clockwise or anti-clockwise. Then the starting channel was randomly selected, along with how many channels (1 or 12) the bird should travel before the end of the 8-second sample. In addition, a one second bird sound was triggered in a randomly determined speaker every 2 minutes throughout.
- The rotating bird at every half hour was followed by the same breathing and orange light as before.
- At quarter to every hour, a randomly determined colour filled the glass dome for a short time and carried no sound.

This timed sequence along with the continuous motions of sounds created a shifting and organic changing *mass* throughout the entire exhibition space through the *gait* of the sound. The *grain* of the sound changes as the audience moves around the space and the sound moves around them. The *duration* of the individual sound components, as they are "psychologically experienced", is completely dependent on the listener's position, the sound's position and the time-point in the design sequence.

4.5.3 Discussion

In section 1.3 we looked representational processes, and the notion of "non-specific spatialisation" (S. Wilson & Harrison, 2010). This is a thinking that seeks to use spatial audio method to synthesize spaces but not to accurately reproduce any real-world context. The stochastic methodology arose from a desire to take advantage of the inherent idiosyncrasies and limitations of that the space in The Museum of Contemporary Art afforded. The museum was housed in the former central bank of Norway, Norges Bank, which adds to the rather overwhelming visual appearance of the space and consists of rooms with dramatically varying sizes and ceiling heights. The artist's aim was to transform the space completely, which was achieved in large parts by the carpets covering the walls and floors, the lighting design, and the moving sound.

The notion of non-specific spatialisation emphasises that this methodology does not seek to "simulate precise locations or directions of (usually point) sources" (S. Wilson & Harrison, 2010, p. 241). As such, in many situations one aims and wishes to recreate a specific real-world sound scene or to synthesise a "new" space to create an imaginary but "real" space. The aim for Hot Pocket was not to attempt to create a believable real space but to emphasise the experience of being immersed in a cave or a large animal of some sort. Non-specific spatialisation is helpful when considering the lack of standardization among the various setups for multi-channel sound reproduction, but specifically when attempting to design a holistic and functioning spatialisation system for highly irregular spaces.

The realisation of this approach to spatialisation came about as a solution to the specific problem posed by the artist. The advantage from this approach to sound spatialisation was the flexibility afforded by using re-generated spatial trajectories rather than resting on a "composed" looping tracks from a digital audio workstation. The spread of the sound through the uses of stochastic tables is flexible in that different sounds can be given more or less random properties; if you increase the step size of one looper, then the pattern becomes more random than a low step size. To retain a kind of "reality-equivalence" the step sizes were kept low to allow the audience to trace the sounds as they moved through the space.

For the newer iterations of the functions, the traditional amplitude differences used in amplitude panning-based spatialisation also uses a distance filter (see section 4.4). The distance filter uses a low-pass filter to roll off the higher frequency content of a sound the further away it is perceived to be. These functions are planned to be extended to include several other applications of stochastic motion of sound to create a flexible tool for spatial audio, by including, among others, time and frequency domain splitting of the input sounds for individual spatial trajectories.

4.6 Case 3: *Superimposed Landscape (Lista)*

The landscape around Lista Fyr in the south of Norway is unique. It opens up into the North Sea, the trees grow at an angle to avoid conflict with the wind, and the area is mix of flat, rocky, and rugged. It attracts thousands of tourists every summer and avid birdwatchers all year. With several conservation areas and a site of extensive research in ornithology, it is also the site of war and destruction, with traces from the Nazi occupation still visible in the landscape. As a venue for a sound installation, it is unique - a place where the landscape disappears into the sea.

Superimposed Landscape (Lista) was a site-specific installation that was part of the juried group exhibition Sørlandsutstillingen's 50th anniversary exhibition.[26] The proposal for the installation was made in response to a call for works that involved a direct negotiation and interaction with the site surrounding Lista Fyr. The installation was situated in a bunker from the Second World War, as seen in Figure 4.17, and was presented as an eight-channel sound installation in a two-dimensional higher order ambisonics, horizontal shoebox-shaped configuration. The premise for the installation was to use the sounds from the surrounding location as its only source material and to "recreate" the sounds from the landscape in a space that was part of the site. This was achieved by installing microphones outdoors in different parts of the landscape and building a sampler that would capture the sound from the microphones to disk. At the time of the proposal, the space was not clarified and in the end I had only a few days to set up the installation, connect microphones into the landscape, record, and build the software for the space. The positioning of speakers, microphones, and interactions with the space is an important exploration of the exhibition space and has to be done on-site (Lossius, 2008), leaving less time to develop the installation.

The installation used two sound inputs: 1) real-time sampling from the surrounding landscape, and 2) a bank of samples recorded from landmarks in the area surrounding the lighthouse. In section 3.3, on discussing the site-specific, we saw that a sound recorded in one place, treated in the studio, before being re-introduced to this same place, represents a transposition of the sound in time where the listener would meet a "historical" version of the same place. In this instance, the capture, processing, and presentation of the sounds happen in the same location. If we follow Smithson's notions of the non-site, the experience of the sounds the installation is still of the site.

Locating the installation in a bunker that was cut into the landscape, allowed the installation to become part of the landscape rather than to be dislocated and housed in a gallery removed from the site. Inside the bunker, the audience could experience the installation itself but the sounds from the North Sea was always present outside and this caused a continuous mixing of the sounds from the inside with the outside. The outside landscape then functions as an acousmêtre (see section 2.6.1 and (Chion, 1994)), a audible but hidden presence that has a high influence on our understanding of what is present to us.

[26] https://sorlandsutstillingen.no/utstilling/2019/. The installation has also been accepted to the International Computer Music Conference 2020 in Santiago, Chile, which has been postponed to 2021.

Figure 4.17: View of the exhibition space at Lista Fyr.

In Bill Fontana's *Sound Island* (1994), sound from the rugged cliffs facing the Atlantic Ocean off the coast of Normandy was transmitted to hidden loudspeakers lining the Arc de Triomphe in the centre of Paris (LaBelle, 2015). The natural white noise of the waves crashing against the rock on the coast masked the heavy sound of the traffic surrounding the monument and

> what stands out in Fontana's installation is the continuation of the transposition of realities indicative of soundscape composition, while add to this the mixing of visual experience with acoustic phenomena in real time. (LaBelle, 2015, p. 231)

Although *Superimposed Landscape (Lista)* did not exist in the same space as Fontana's *Sound Island*, the effects are similar but the dislocation of sound is not the same. Here a listener can be immersed in the outdoor landscape, then meet new instances of the transformed landscape inside the bunker before returning to the "original" landscape.

The technical implementation of the installation draws inspiration from examples such as *I Am Sitting In A Room* (Lucier, 1981) and *4 rooms* (Kirkegaard,

2006) where a space is recorded, played back onto itself, and re-recorded, again and again. Along with the outdoor microphones, two omni-directional microphones were installed in the exhibition space to sample from what ever sound was currently present. This also bears resemblance to the performance activity of Kaffe Matthews (Matthews, 1997, 1998), where sounds from the performance space are sampled and manipulated in real time.

In the *Audible Eco-Systemic Interface* project, Agostino Di Scipio (Di Scipio, 2003) created an interactive sonic system that is a structural coupling between machine, human, and environment. Microphones will pick up sound generated by the computer, and the microphone inputs will generate some form of control parameters to affect the audio transformations. In this project the software is continually reconfigured by its own previous output and its environmental characteristics. These different projects all have influenced the thinking and implementation of this installation.

4.6.1 Implementation

Over a three-day period, sounds were recorded from around the lighthouse and edited down as shorter samples (see Figure 4.19). The recordings were made with conventional condenser microphones and contact microphones. The installation was built around two loopers that sampled sound from live microphones and two loopers that played back from a bank of 48 samples, as illustrated in Figure 4.18.

The sounds in the installation exist on multiple timescales, and the pre-recorded sounds are items of the past, a small archive of what particular parts of the site sounded like days, weeks, and months earlier, and these are blended with sounds captured at present. The sounds exists across the macro to micro timescales, as described in section 3.4.2. One aspect of the installation used sounds from one specific point in time and the other used a real-time input from outside the bunker, to add to the temporal shift in focus, as well as the internal sampling from the room itself, which then merged the pre-recorded sound with the real-time recorded sounds into new sampled loops.

In the installation there are a series of recordings done with contact microphones on the radio mast located close to the lighthouse. Although this is a visible landmark, it cannot be considered a *soundmark* on the site, as most people will not have access to the sounds it produces without specialised equipment.

The spatial audio functions were built on the ambisonic library HoaLibrary[27] (Sèdes et al., 2014) for MaxMSP, which allows the user to synthesize, transform, and render soundfields in creative ways. The installation synthesizes mono sound sources into two-dimensional variable order ambisonics. At each recording or se-

[27] http://hoalibrary.mshparisnord.fr/en/

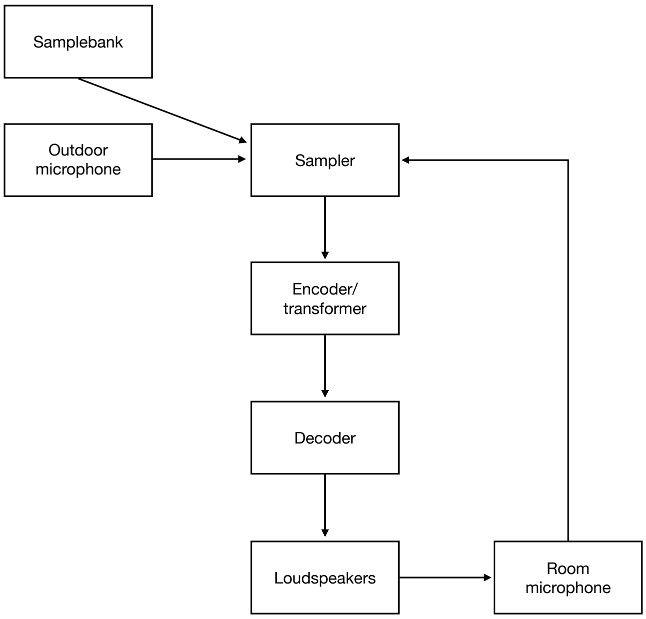


Figure 4.18: Diagram of spatial audio implementation for *Superimposed Landscape (Lista)*. A sampler captures data from an outdoor microphone and/or selects from a pre-recorded samplebank, these samples are encoded in higher order ambisonics, a frational order transform is applied to each sample and it is decoded to the loudspeaker setup. A room microphone samples back into the system at random intervals.

lection there is a probability system:

- For pre-recorded samples there is 40% chance of looping the sample again, a 40% chance of selecting a new sample or a 20% of being silent.
- For samplers there is a 50% chance of using the outdoor microphone, 30% of using an indoor microphone that points outwards or a 20% of silence.
- Every 5 minutes a microphone in the exhibition space records a sample of random length between 2 seconds and 2 minutes and plays this back into the space, and loops the sound of the space back onto itself.

When a sound file is loaded or recorded, the azimuth direction of the source is randomly determined, as well as the fractional order for each encoder. If the signal is kept at seventh order, the sound source could normally be very clearly localised to the azimuth direction specified by the random function given enough speakers to reproduce the decoded signal (in this instance a minimum of 15 speakers would

Figure 4.19: Map detailing the orientation of the installation. The red dots indicate a recording location, the green dot indicates the entrance to the bunker and the orange dots indicate the placement of microphones outside the bunker. Image from the Norwegian Mapping Authority website (Kartverket) https://www.kartverket.no/en/.

be needed). As the fractional order approaches 0, the sound source becomes more and more diffuse and difficult to localise (this is illustrated in Figure 4.20).

Each individual looper was connected to a separate ambisonic encoder/decoder pair. Between each encoder/decoder there is a widening function, through the object [hoa.2d.wider∼]. This widens the localisation of a diffused sound, at 0., the sound is omnidirectional (0th order) and at 1. the sound is localised at the ambisonic order specificed in the encoder, in this case at seventh order. The choice of seventh order stemmed from early experimentation with the real-time loopers conducted in the same 24 channel studio and through the same motivations as the case study in section 4.4, to examine spatial features of multiple superimposed recordings. The seventh order ambisonics under correct conditions provides a high degree of localisation accuracy of the spatialised source, yet in this instance that would not be so as the signal was decoded to 8 channels. Also, through the wider-function the fractional ambisonic orders would most often blur the sound image.

Going back to section 3.6.2, we looked at the problems of encoding mono sources into ambisonics. The same problem comes into play here, where the installation used mono, point-source microphones to capture sound. Here this was approached through three methods, 1) to rely on the fractional order transformation, where a sound would sometimes be "easily" localized in a specific direction

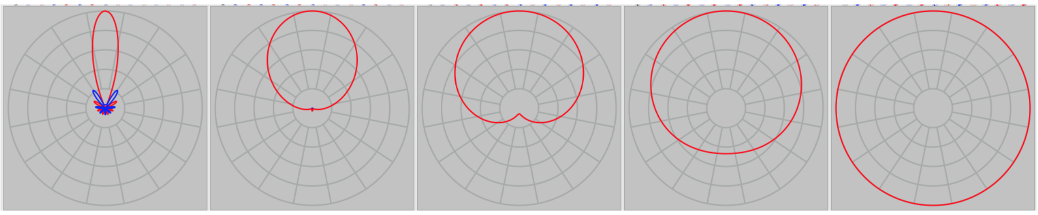


Figure 4.20: Diffusion of a localised sound between fractional orders 1. and 0.; where the graph on the left side displays a focused source, the graph on the right side displays a diffuse source. In this instance, a fractional value of 0. will create a 0th order encoded source and a fractional value of 1. will create a source to the order specified in the encoder. The variable values between 1. – 0. display linear changed logarithmically.

and at other times blurred; 2) the reflections from the room itself causes localization blur, due to the curved, steel ceiling, and concrete floor; and 3) several different sounds are layered at any given time. Specifically, the architecture and materials of the space creates a large amount of reflections, and given the sounds coming in through the door made it so that localisation of the individual sound sources were next to impossible.

4.6.2 Outcomes

As can be seen in Figure 4.17, most of the walls and ceiling in the exhibition space is made of corrugated steel while the rest is concrete. "Controlling" a sound diffusion in a room such as this is next to impossible as the surfaces will create unpredictable reflections. The nature of this space would prove it difficult to use other well-established point-source techniques both as the sweet-spot, in the conventional concert-setting, is not present and as the room itself would make the localisations difficult. The sweet spot mimics a concert situation (Lossius, 2008) and is generally not much suitable for situations where the audience is expected to move around the room. As the bunker opened up onto the North Sea, one could clearly hear the outside sounds influence the perceptions of the sound inside the bunker.

For anyone to precisely gauge the properties of direct-to-reflected sound, the proportion of inside-to-outside sound is next to impossible in this situation. Rather than attempting to control this nature of the sound perception, the installation used the fractional orders function to create layers of diffuse sound, where "precisely" localised ambisonic sources would be blended with diffuse sources, which would again be blended with the reflections from the room and the subsequent timbral colourations introduced by the reflections. More loudspeakers would also mean a higher degree of perception of the gait between the different ambisonic orders.

The approaches in this installation investigated the *dynamic* profiles of the sounds as they were arbitrarily sampled from the surrounding landscape and how this influences changes in *mass*. Owing to the nature of the space, the sound's *gait* oscillates between the foreground and the background, between the smeared and the focused.

Although there are many notable historical events and anecdotes that could form the basis of an exploration of a sound installation on the site, this installation focused solely on the sound of the place. In addition, the weather was an extension to this. Being extremely exposed to the changing weather patterns coming in from the North Sea, this is a fundamental experiential trait of the area.

Future work on this installation seeks to implement an extension to the sampling paradigm through machine learning. As each microphone records a sample and plays it back in the space through the process discussed previously, the sample is also stored to disk. As a database of sound is accumulated, a learning function searches through the database of existing sound and attempts to match the currently recorded sample with a sample from the past. This way the installation both explores a situatedness in the landscape and charts a "history" of the changing sounds in the landscape.

4.7 Case 4: *City Dwellers II*

This case study discusses a sound design and spatial programming project for the hybrid theatre performance/installation *City Dwellers II*, performed on Stage 2 at Kunsthøyskolen i Oslo (KHiO) as part of the PhD-project *1:100 - the performative hybrid text as a feedback loop* by Tale Næss Lysestøl.[28]

City Dwellers II is a city portrait, where the audience meets over a hundred different voices sounding from their individual houses and brought together in a city-space, the formation of which was heavily inspired by Wim Wenders' 1987 film *Wings of desire*. The "city" is undefined but rather exists as a structural metaphor to encapsulate the different voices. The individual houses were white, 3D printed objects that each housed a loudspeaker and a sound-activated light that drew the audience's visual attention. The houses were suspended vertically on two wires attached to the floor and a steel frame suspended from the ceiling (see Figures 4.21 and 4.22). The voices were in a multitude of languages, including Norwegian, English, Farsi, Danish, and French, which made up indi-

[28]The work done for *City Dwellers II* is to be continued as a sound-only installation. A residency to rework the material and to develop the installation was scheduled to happen as part of the Intonal Festival at Inter Arts Center, Malmø, Sweden, in April 2020, but has been postponed due to Covid-19.

vidual testimonies, fragments of conversations, warnings, laughter, confessions, and dreams.[29]

The voices are “disembodied” and acousmatic, but with specific spatial locations. The role of the sound design was to tie the room together as a surrounding layer around the sounding voices and the audience, as a background texture that moves through and around the space. In this sense the aim for the sound design was to create a *mass* with a slow *gait* through the room. This perspective afforded a thinking about the depth, change, and definitions of the space as the sound design should never mask the subtle voices from the houses. For the four performances we utilized an approximated dome 3D loudspeaker setup, with 12 azimuth and eight elevated speakers along with two subwoofers located at either end of the room.

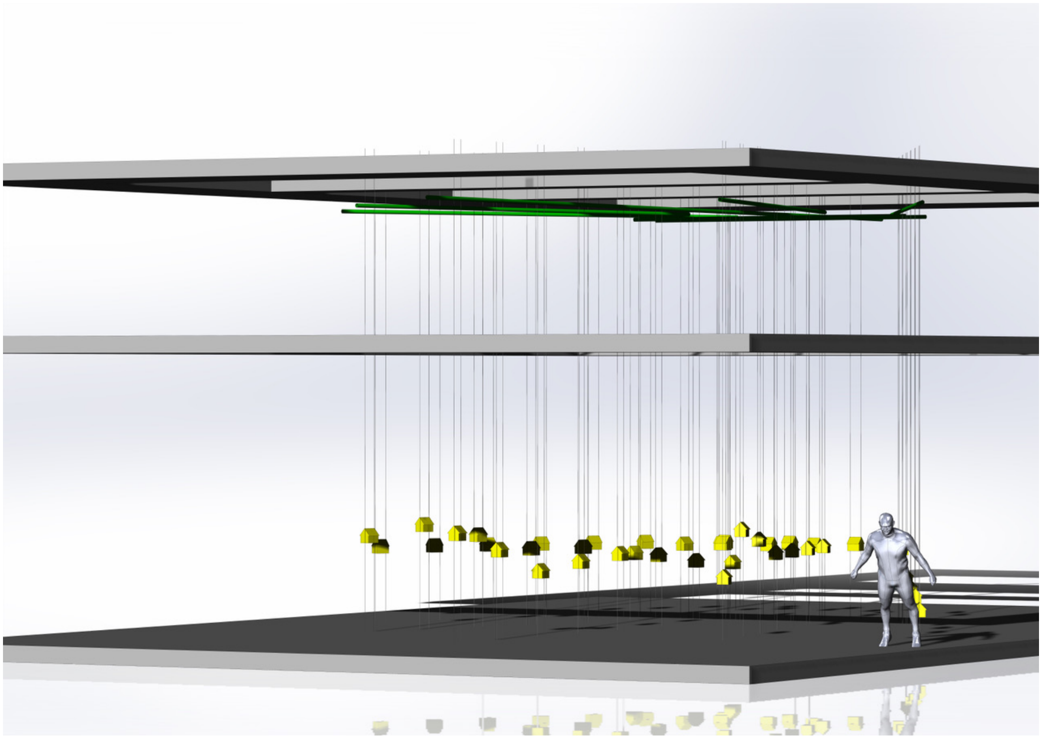


Figure 4.21: Scenographic design for *City Dwellers II*, illustrating the layout of the house-structure throughout the stage. Scenography by Lawrence Malstaf.

The first approach to the sound design experiment was anecdotal, through the use of processed and unprocessed field recordings, most of which were made in a city-space: passing cars during a rain and thunderstorm, ambient city sounds,

[29] The “pool” of texts that make up the voices can be found at: https://www.fromonetoahundred.com/city-dwellers-project/.

ventilation, birds, children playing, church bells, and public transport, along with synthesized textures. These were all mixed as a *virtual soundscape* (Roads, 2015), drawing on the work of Luc Ferrari, especially from the different voices and languages in *Presque rien No. 1 Le Lever du Jour au Bord de la Mer* (1970) (Ferrari, 2009), where the layering of different languages is set against looped background materials and processed sounds. The anecdotal sounds and the literal interpretation of a city space created a conflict with the sounding voices from the different houses and in many instances masked the voices that were the centre of attention. A complete re-working of the sound design was needed.

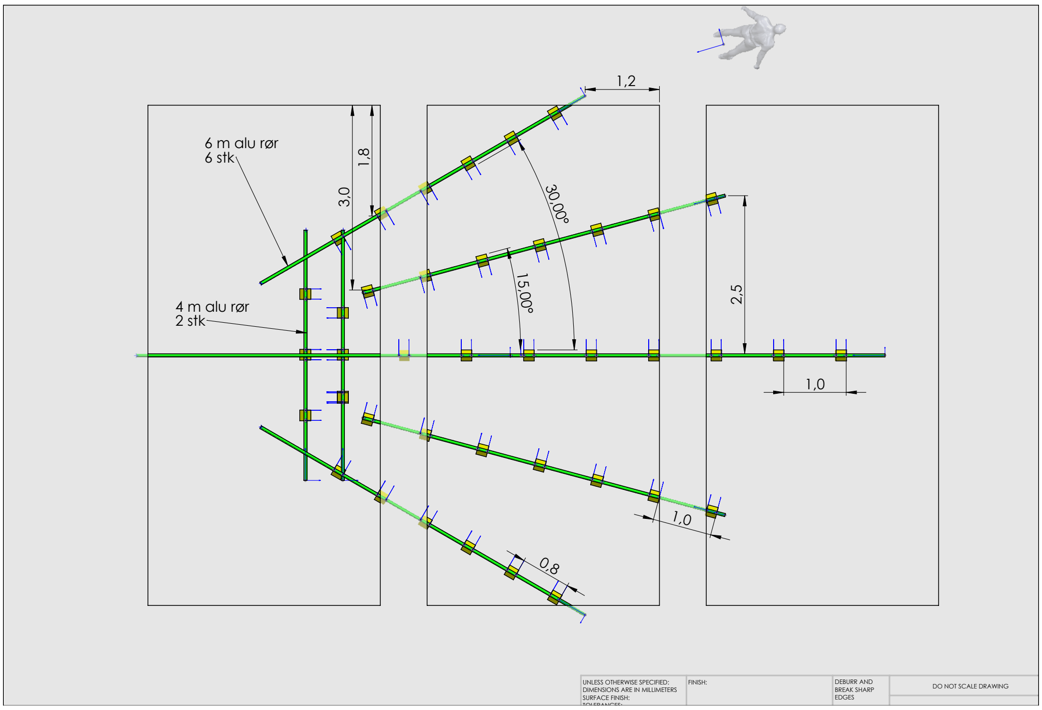


Figure 4.22: Top view of the scenographic design for *City Dwellers II*, illustrating the layout of the house-design throughout the stage. Scenography by Lawrence Malstaf.

Considering the installation/performance as a city space, its presence rests on a sense of place and how this is lived. I considered the *structure* of place to be of importance (Norberg-Schulz, 1971). We have seen, in section 3.1, that social practice is place-bound, place is the "moment" when the conceived, the perceived and the lived attain a structured coherence (Merrifield, 1993), that which we embody. This structure consists of several layers, the voices make up the primary focus points in the space and they are located "indoors", in individual and spatially distributed houses.

At the beginning of the performance, all the houses are placed close to the floor, at the bottom of the wires. When each house lights and sounds for the first time, one of three actors would move to that location and slide the house up the wires to place them at different heights. This activity is the primary focus. Secondly, the audience is conscious of their own spatial position in the space; seated or standing along the walls, the proximity to the other audience members, and the distance to the centre of the "city". Lastly, the audience is surrounded by sound, a background ambience, or ambience labelling information (Lennox et al., 2001).

When the original, anecdotal sound design failed, it became clear that an anecdotal background would be a disturbance for the audience's focus towards the voices sounding from their individual positions in the city. Instead, voice samples were processed to create different layers of non-representational drones to encapsulate and tie the room together. We found that this had an advantage of being "site-less", when the sounds did not represent any concrete, real-world event, it would be easier to blend the sounds into the background to create both a context from where we hear the foreground sounds and something which "fills" the space and create presence and a variable spatial impression. This approach was inspired by the thinking which went into Anders Vinjar's *Le camere invisibili (2016)*, discussed in section 1.3.1. Before discussing the technical implementation and spatial distribution in the space, I will briefly discuss a central notion underpinning the performance, the social space.

4.7.1 The thirdspace

The previous three case studies discuss space and place differently. In case study 1, space is a product of the morphological manipulations of simple sound materials through a *gait* between foreground and background and the *mass* of the soundfield. In case study 2, the stochastic *gait* of a sound through space creates differently perceived *mass* depending on where in the space the listener is located. In case study 3, the landscape is represented within the landscape itself and the audience can experience the interplay between the site and the representation of the site within the same space.

A city-space is a social place, and in section 3.1 we discussed Lefevbre and his concerns with space and place. In everyday life, place is synonymous with what is lived and social practice is place-bound. Drawing on Lefevbre, Edward Soja's concept of the *thirdspace* is a way of thinking about and interpreting social space (Borch, 2002), based on Lefebvre's triad of the perceived-conceived-lived (Lefebvre, 1991). Soja constructed a model of a "fully lived space", that he called "a 'thirdspace' perspective, and distinguished it from the traditional binary mode of looking at space from either a material/real perspective or a mental/imagined perspective, a sort of objective–subjective simple binary that I described as 'bicameral' way of looking at space" (Blake, 2002, p. 141). This notion of thirdspace indicates that we can represent space as a multidimensional feature *place*, where axes move through space and change the relationships among objects, in this instance as a structural metaphor for the synthesis of new places:

> First and second spaces are two different, and possibly conflicting, spatial groupings where people interact physically and socially: such as home (everyday knowledge) and school (academic knowledge). Third spaces are the in-between, or hybrid, spaces, where the first and second spaces work together to generate a new third space. (Mayhew, 2015)

Not unlike Soja's ideas of a fully lived space, Marc Augé introduced the concept of "non-place" (Augé, 2008) to refer to spaces where relations, histories, and identities are erased. He uses examples such as motorways, hotel rooms, airports, or supermarkets. These spaces are empty of any historical, social, or experiential memory. *City Dwellers II* started as an empty space, filled with anonymous, white houses. Gradually, the voiced fragments synthesize a new social space.

4.7.2 Implementation and outcomes

An approximation of a 3D dome was incorporated surrounding the installed houses and the audience with 20.2 channels, using a variety of Genelec and Meyer loudspeakers. All sounds were encoded to third order, higher order ambisonics using the IEM ambisonic plug-in suite. This plugin suite features an AllRAD decoder (All-Round Ambisonic Decoder) (Zotter & Frank, 2012, 2018), which is well suited for arbitrary loudspeaker arrangements. Stage 2 at KHiO is approximately $155m^2$, where the house-structure in figures 4.21 and 4.22 occupies most of the stage-space. The "walls" encapsulating the stage consisted of a black curtain and the loudspeakers forming the background layer lined by curtain all around the space.

A well-known problem with ambisonics is the need for a static, central listening area, or sweet-spot. However, in this case the audience was encouraged to physically move around the space during performance, whereby any central listening space would not exist. This causes problems for the rendering of the decoded ambisonic scenes, which became evident in strong phasing distortions when using concrete sounds and moving around the space. An inherent "problem" with using concrete, real-world sounds is that the audience has an internal memory of *how* these sounds would be "correctly" heard when experienced in the real world. Given this problem, there were few possibilities to convincingly re/create a type of fictional city soundscape.

However, when using drone-based or completely synthsesized sounds the phasing distortions were much less evident, most likely due to the lack of representational content. In attempts to reduce the distortions I used inPhase optimisations through the Directivity Shaper plugin[30] to reduce the distortions due to phasing. Sometimes considered a drawback, here it was an advantage as the inPhase optimizations in third order ambisonics contribute to reduced spatial clarity and localisation of sources and this caused more “smeared” sound images (Barrett, 2010). The Directivity Shaper plugin is a frequency dependent ambisonic encoder, which can split a mono signal into four separate directivity band signals, similar to beamforming (see section 3.6). The plugin allows each split band to be encoded to a different ambisonic order, similar to the fractional order function in section 4.6 and panned on the sphere. This also mirrors some of the intentions which were discussed in section 4.3.

These simple steps for spatially projecting the sound design in the space, both created a clear *presence* of the sound materials, and provided a large *spatial impression* surrounding the audience. This contributed to a changing *mass*, where the *mass profile* is influenced by a slow *gait* of the sound materials. The smeared sounds across the speaker array made the *grain* of the sound audible. All these perceptions were contingent upon slow a *gait* through the space and the difference in nature between the foreground and background sounds. Unlike the requirements in *Hot Pocket*, (see section 4.5), this sound design did not run in real-time and instead played back a fixed-file, decoded version of the sound design in the space.

The individual 3D printed houses were self-contained in terms of its playback technology. Each house had an iPod, a small speaker and a small circuit to power the flashing LED in sync with sound playback. This involved manual synchronisation of the playback units, four units at a time, and each exported sound file had to be offset to align with the synchronisation. There was much room for human error in the synchronisation and we did not have any way of synchronising the playback units with the computer for sound design playback.

The planned reworking of this project into a sound installation approaches the material in a different way. Where the performance of *City Dwellers II* was arranged in time and space through a mix, the installation will be generative. A selection of voice recordings from the original performance has been selected by Næss and these form the new “pool” for the installation. For the reworking of the project, the distribution of the voices is done by simple randomisation algorithms. The details are yet to be decided, but the current conceptual state is: first, the number voices is selected; second, the number of instances of these voices and the distributions in space are determined; and third, the number of loops are determined for each voice. The selection process is similar to those discussed in

[30] https://plugins.iem.at/docs/plugindescriptions/#directivityshaper

sections 4.5 and 4.6. As each voice is placed in space, the distances to the centre of the array is determined through the uses of a distance filter (see section 4.4) to create depth. Each voice will be stationary in space but the changing perspective of depth along with the shifting *mass* of the soundfield is thought to create a *gait* of the *mass*, which is an oscillation in the perceived spatial impression of the space.

In the next section, and last case study, we will look at spatial audio as integrated into a learning environment and explorations of the uses of room acoustics to create a place for instrumentalists to practice their performance skills through a virtual reality application.

4.8 Case 5: *Music in the interactive space*

The last case study in this thesis is slightly different from the four preceding cases, in that it is focused on the uses of technology in pedagogy through the use of virtual reality for teaching, performance, and practice preparation of music students. The project explores the possibilities of performance enhancing training for students through Virtual-Reality and multichannel audio, at the Centre for Excellence in Music Performance Education[31] at the Norwegian Academy of Music in collaboration with Johannes Lunde Hatfield.[32]

We set up a small lab at the Academy with a 12-channel sound playback system, as seen in Figure 4.23. The students could experience several different performance situations, from waiting in a "green room" to performing on a 360° stage modelled after the Elbphilarmonie in Hamburg, to a smaller Viennese concert hall, as well as audition situations from the two concert halls. The Virtual Reality animation and simulation was made in Unity by PointMedia.[33] Owing to various delays with the virtual reality simulation, the project has not been finished to a degree where perceptual experiments could be conducted.

The project's aim was to establish and develop a learning laboratory in which music students could improve on their instrumental practice and performance preparation. Music students will practice up to about 7800 hours through out a five-year study course (Hatfield, 2016), most of this time is spent alone in a practice room, which also can contribute to stress when faced with a concert or audition situation (Kenny et al., 2004). To overcome this, new technological affordances are starting to emerge for classroom and performance simulations (Bissonnette et al., 2015). This project has been inspired by the Performance Simulator at Royal

[31] https://nmh.no/en/research/cempe
[32] https://nmh.no/en/research/cempe/news/instrumental-practice-in-an-interactive-sphere
[33] https://www.pointmedia.no/?lang=en

Figure 4.23: View of the virtual reality lab at the Norwegian Academy of Music. One performer is seated surrounded by 12 loudspeakers but without the virtual reality headset.

College of Music,[34] which places the performer in a room surrounded by projection screens of the concert hall and audition situation (Williamon et al., 2014).

To accompany the visual animation/simulation of the green room, concert halls, and auditions (see Figure 4.24), the spatial audio integration was used to model the acoustic environment in the concert halls, the sound design of the concert halls as well as locating the performer on stage. Although the aim has been an acoustic modelling, the goal has not been to create an "accurate" modelling of specific halls, but rather to create perceptually believable sound environments. Much of the work has been tuned "by ear".

Unlike most other projects so far in this thesis, the target of this project is to represent spatial attributes "exactly" as they would be experienced at a concert venue (Laitinen et al., 2012) and the relationship between the foreground and background sounds needs be controlled. Unlike many room acoustics experiments, the spatial attributes that are to be represented in this project is as they are

[34] https://www.rcm.ac.uk/research/projects/performancesimulator/

perceived on stage and not in the hall where the audience would be seated. From discussions with several experienced performers and professors of instrumental practice, the consensus has been that the perception of the sound on stage where the performers are performing is perceived to be "dry", meaning little or no reverberant effect from the room. To model virtual acoustics, three criteria are needed for the acoustical communication: source modelling, transmission medium, and receiver modelling (Savioja et al., 1999). This source-medium-receiver is common to all communication systems.

Acousticians will often refer to objective parameters, as defined in ISO3381-1:2009 [35] as a guideline; however, this guideline does not discuss the subjective perception nor preference of listeners (Lokki, 2013). The subjective perceptions of concert halls are difficult to measure but this does highlight the need to go beyond the impulse response measurement and standard criteria (Halmrast, 2015), to focus on the perceptual consequences of frequency-dependent phenomena in musical instruments and human spatial hearing (Lokki, 2016).

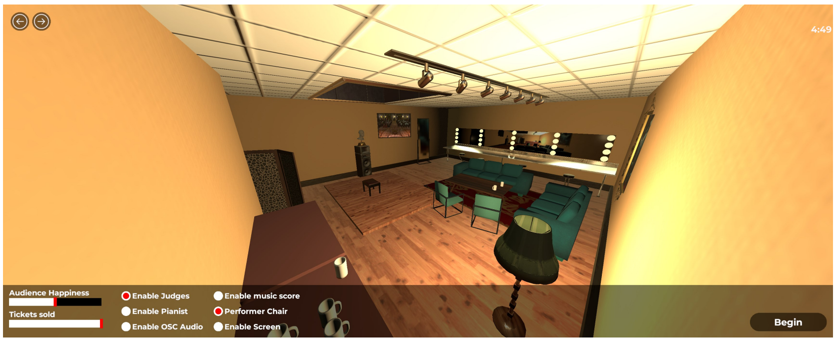
4.8.1 Implementation

The lab was setup in August 2018 and a first round of audio-only experiments with 13 participants was run in October 2018. The lab was set up so that each performer was seated (or standing, dependent on instrument) in the listening area surrounded by speakers and wearing a virtual reality headset. A microphone[36] was installed on the instrument and is used as the main interface from where the students played through the simulation. The performer generally did not wear headphones but a binaural stream of the "main mix" was sent to a dedicated set of headphones either for use by the student or a supervising teacher.

The hypothesis for this project is that through the establishment of a learning laboratory, students can improve their instrumental practice and performance preparation through the application of Virtual Reality technology and spatial audio. Rather than opting for a solution where the student would wear headphones and listen back to a binaural stream we have worked from the hypothesis that the students would feel a higher degree of immersion and realistic experience of playing on the stage in a concert hall over loudspeakers than over headphones. In section 3.4.1 we saw that sound source localisation over headphones localises a source inside the head, whereas over loudspeakers we localize externally to ourselves (Rumsey, 2001). The uses of speakers will also aid in changes of spatial impression, spaciousness, and presence between the different performance rooms.

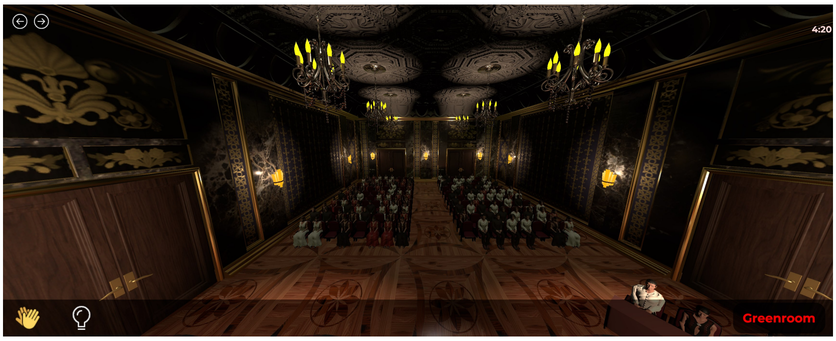
[35] https://www.iso.org/standard/40979.html

[36] We employed DPA 4099 (high SPL) microphones for all instruments https://www.dpamicrophones.com/instrument/4099-instrument-microphone.



(a) The green room.

(b) Stage in the Elbphilarmonie.



(c) Stage in the Rococo room.

Figure 4.24: Three different screenshots from the different simulation rooms in the virtual reality application: (a) the green room, where the performers will wait before entering the stage; (b) the stage in the Elbphilarmonie; (c) a rococo room. In the two concert hall images, a seated jury can be seen on stage.

The lab has been set up with an approximated dome of 12 speakers; eight azimuth and four elevated. The room is approximately 4.6 x 2.9 meters and the listening/playing area is a circle of approximately 1 metre diameter. The elevated speakers allows for modelling of sound sources from audience members seated

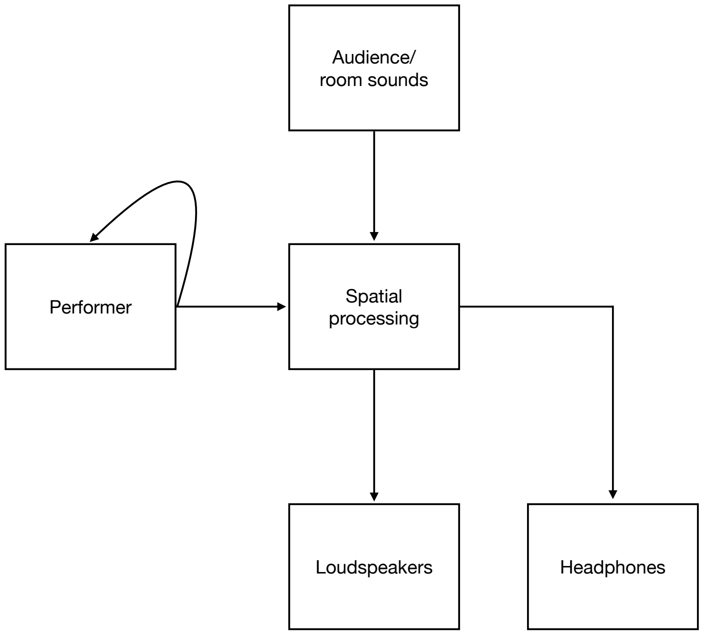


Figure 4.25: General overview of audio processing for virtual reality simulation. The performance is captured by a microphone attached to the instrument and processed spatially through acoustic modelling, along with room/audience sounds. This is played back to the performer over loudspeakers. The performer also listens back to the direct sound of the instrument.

in the upper balconies, but where as in the smaller hall and with the audition situation, there are only horizontal source locations. As stated previously, the project had developed two different-sized concert halls, along with a "green room" that functions as a waiting room for students before and after the performance. There is also an audition committee that can be present in both concert halls.

When the simulation starts, the student is seated in the centre of the circle of speakers. The animation shows the outside of the concert venue and the operator selects which concert hall to use for the session. The student is then placed in the green room where he/she has to wait for 5 minutes. There is a clock counting down and a voice announces the remaining time before the performance starts. Throughout the period in the green room, the performer can hear the sound of the audience filling up the concert hall, albeit muted. Then, when the performance starts, the student is lifted up onto the stage and is met with applause. The audience falls silent and the performance can begin. Once the performance is finished, the audience applauds, and the student is lifted back into the green room.

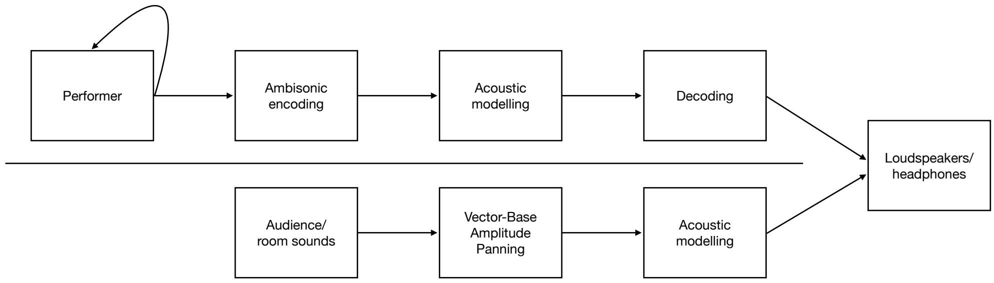


Figure 4.26: Overview of room processing of two different signals in the virtual reality concert hall. The performance is captured by a microphone, while the performer listens back to the direct sound. This signal is encoded in ambisonics, acoustically modelled and decoded back over loudspeakers. The audience sounds are placed in space using VBAP for the performer to as accurately as possible localise the sounds.

During the performance, each performer listens back to the direct sound of the instrument as well as to the simulated acoustic colouring of the sound that will be reflected back from the room. The instrument sound is generally quite dry and never over-accentuated. The goal for each performer is not only to experience a simulated concert hall setting but to rehearse in a setting where there are multiple disturbances from the surrounding room. The audience will cough, talk, have ringing phones, and the like, as well as applause, of course. In the Viennese halls, the sounds are frontal, while in the 360° hall the sounds can come from any direction.

The audio interaction software was implemented in MaxMSP with IRCAM's Spat 5 library. Spat is currently the most comprehensive library for spatial audio methods available next to all other available tools in the MaxMSP ecosystem. At the heart of the library, is the reverberation algorithm in [spat5.spat∼] which is a Feedback-Delay Network that can be used for room acoustics simulation. The Spat library offers a "perceptual control interface" through the [spat5.oper] object, which is a high-level control interface. It can be used to control the reproduced effect through a small set of controls, and the perceptual operation in [spat5.oper] allows for separating the processing into different "rooms". The advantage is that different room acoustical properties can be assigned to each room as to model differences in point of audition. The [spat5.oper] provides an interface to control aspects of an acoustical modelling such as room presence; source presence, brilliance and warmth; time-domain filters per source; air-absorption; localisation; along with a series of reverberation modelling functions such as reverberance, room size, and intervals between the different types of reflections.

In each room, the influence of the room reverb can be modelled separately. In Spat each section of the traditional reverb graph is divided into four sections, see Figure 3.3 where, next to the direct sound, the room response is divided into early and late (reverb) reflections. In Spat this had been extended to: Direct sound, early reflections, late reflections, and late reverb. Direct sound and the early reflections are precisely localized in space and the two last sections are spatially diffuse. This diffuseness is due to the decorrelation effects produced by the Feedback Delay Network-reverb module (discussed in a previous section), which is at the heart of the function [spat.spat∼].

During performance, the performer will experience disturbances from the audience. To localize this, the simulation will upon loading define the number of audience members present in the room and provide each source with an X,Y,Z coordinate for precise localization of the disturbance. Each source will be spatialized using Vector-base amplitude panning (Pulkki, 1997) and given "room presence" through the reverberation algorithm. Modelling of the instrument's reflections from the room back to the stage has not been finished. The challenge is creating different room modelling setups for each individual instrument group and sub-group.

4.8.2 Sound design

In order to recreate a "convincing" concert hall experience, we recorded a series of sounds to accompany the simulation to give the feel of a real concert hall. These sounds are first experienced in the green room, when the performer is waiting to go on stage. First, the performer can hear the sounds of the concert hall filling up with chatter and people moving as they are getting to their seats. A voice announces to the performer each minute over 5 minutes for how much time is remaining before the performer will be on stage. This waiting time is expected to provoke expectations in the performer and give them time to mentally zoom in on the upcoming event. For the different stages we have recorded room tones from two different concert halls, along with individual clapping, coughing, phones ringing, and "bravo"-shouts. For the audition situations we recorded male and female voices in Norwegian and English giving instructions to the performers, along with sounds of shuffling paper and writing. Some recordings were made where one of the judges would stop the student during the performance.

Room sounds were recorded in Lindemanssalen and Levinsalen at the Academy and at Oslo Concert Hall. These recordings afforded room features from different-sized concert halls and allowed us to map the room sounds to the different levels of simulation. The sounds from the concert halls that can be heard in the green room are muted using a distancefilter as discussed in a previous section. When the performer is transported from the green room to the concert hall, the distance filtering of the concert hall is slowly diminished through the transport before erupting in applause as the performer steps onto the stage.

Audience sounds were recorded in a dry space, so the distance modelling in the room can be as effective as possible. In the last version of the simulation, each audience member would trigger a clap individually so the distances and directions of each audience member could be modelled individually. For the 360° stage, the applause comes from all directions and from the Viennese hall, the applause is frontal. The simulation was set up to send audio triggers over Open Sound Control (Wright et al., 2003) to MaxMSP to trigger the different signal operations and sound design. At the start of each simulation, the number of audience members and their positions would be sent to Max to initialize the room modelling and placements of all sources. Next to this, an internal clock would trigger the green room sounds, along with the lifting of the performer up to and down from the stage. The applause would be triggered manually by the operator or teacher, as would the sounds from the audition panel.

4.8.3 Outcomes and problems

As noted in the introduction to this case study, the project has not been finished to such a degree that perceptual experiments could be conducted. However, the first part of a pilot project was completed in October 2018 with 13 participants. This study only used audio, as the virtual reality simulation was not yet complete. Delays in the development of the simulation led to the follow-up experiments being postponed and in the end cancelled. The final simulation was not finished until February 2020, and follow-ups were difficult due to Covid-19, which has led the project to being halted. The project is planned to be followed up in 2022, where improvements to the animated audience through the application of motion capture animation techniques will be of importance. This will aid in the creation of a smooth audiovisual solution to be used by music performance students and teachers. However, the timeline for a finished project where perceptual experiments can be conducted is still open.

The pilot project was planned in two parts, one with audio only and the second with the full simulation, using the same group of students. The pilot experiments were conducted in two parts, where the students would first play their instrument and receive the room feedback over headphones and the second over loudspeakers. The first round of experiments had 13 participants, consisting of three violins, two flutes, two cellos, two pianos, one singer, one horn, one trombone, and one viola.

Feedback from the participants highlighted some key points in the further development of the simulation:

1. Immersion into a virtual concert hall was not present due to lack of visual input, the reality-equivalent perception of the concert hall was gone, as the students could only hear the sounds but not see anything except the white walls of the lab.
2. The sounds are familiar and several students could respond to the sounds by sharing memories of experience from being in concert situations.
3. Several students reported that playing with eyes closed provided a more realistic experience, but also that the background sounds from the concert hall needed to be more diverse and varied.
4. Several students noted that having two people sit behind them as they were playing was very disturbing, especially as concentration would be broken by slight sounds and motions. This indicates that the operator/supervisor should be located in a separate control room.
5. The background sounds should have been more present and louder, to give the impression of an actual audience in the concert hall.
6. Different students noted that the frequency-dependent radiation from the instrument affected the perception of the concert hall and in some registers this gave an unrealistic room impression.
7. For most students the simulated acoustics gave the perception was of a large, church-like hall, which became unrealistic with the combined sounds of the performance and the background sounds, especially when the students were visually situated in a small space. Many responded that the acoustics sounded more realistic when they closed their eyes.
8. Several students responded that playing with headphones was unfamiliar and caused them to bump into the instrument.
9. The localisation of sound felt more realistic over loudspeakers than over headphones, where some students responded that they could localise coughs and other audience sounds better.

Although the feedback was anecdotal, it still has provided valuable comments for the continued development of the project. All students apart from one preferred the simulation over loudspeakers rather than headphones, which could also indicate that there might be a problem with performing with a virtual reality headset. It is also certain that the "feel" of performing with a virtual reality headset will take some time to get used to for the students, and could perhaps lead to some technical difficulties for performances. Another key concern also in the use of the virtual reality headset is both in its weight and the cables hanging down from its back, which could possibly be remedied by using wireless headsets. The simulation system developed at the Royal College of Music uses projected screens, rather than head-mounted Virtual Reality, along with auditory and visual cues (background sounds and spotlights) which would solve potential problems related to students finding a headset uncomfortable to perform with. Unfortunately the experiments never extended this far to uncover potential problems.

The room that houses the lab is very small for the task (approximately 13 m^2), although it has been acoustically treated there is not much space for all the equipment, and one performer and one operator/supervisor. Owing to the size restrictions it is difficult to accommodate some instruments; for example, the trombone and the French horn proved difficult due to the powerful nature of the instrument's sound. In addition, instruments such as the double bass and percussion could prove problematic, due to the small space within the circle of speakers.

A group of jazz students were invited into the lab for an unofficial test. These students reported that the "feel" of virtual rooms is very different to the types of concert stages they frequently experience. The concert venues for jazz musicians is very different from the classical musician and most times a jazz musician will experience talking among the audience, clattering of beer glasses, and activity at the bar, as well as clear sounds coming into the room from the outside - all of which are very unlike a traditional concert hall setting. An inherent drawback with the current spatial simulation is that it is tied explicitly to classical music performance and experience. This is a potential avenue of research within the same project.

Although the short-comings of the lab and the lack of (as of yet) rigorous testing with both Virtual Reality simulation and 3D audio, some of the anecdotal feedback from the students in the first round of experiments has indicated that there is a good potential for this type of tool among students. Specifically it allows for a development of subjectively experienced acoustics of different spaces and how this can afford a development of each student's performance skills. A study has found that students used the simulation at the Royal College of Music to enhance specific performance skills, rather than just use it to reduce anxiety related to performance (Aufegger et al., 2017). In using the simulation, the students uses their existing knowledge of instrumental practice and performance as a structuring principle in the mapping of their existing experiences to new experiences in the simulation. To further develop the simulation we will need a larger data collection on student feedback, and, based on existing work (Ziemer, 2017; Ziemer & Bader, 2015a, 2015b), a rigorous testing of different instruments and their frequency-dependent radiation patterns through the virtual acoustics is needed for the future development of the project.

4.9 Summary

The preceding chapters have discussed the *intrinsic* and *extrinsic* features of sound objects and the object's relationship to spatial audio, through the notions of *objects* and *structures*. The features of a sound object references shapes, places, or spaces through the concept of the *sound landscape*, the imagined place or location of the sound source. The externalized space and place that the object references and exists in, can be analysed and synthesized through extending the criteria of the typomorphological framework (Godøy, 2006, 2021; Schaeffer, 2017).

This chapter has given a short explanation of the framework, along with discussions on related topics such as mapping and timbre, and have applied these perspectives through discussions on five case studies. The typomorphology is a framework for understanding transitions in sound perceptions through a series of approximations, rather than categorical classifications. The tasks of the typomorphology are divided into identification, classification, and description (Chion, 2009) and affords us a rich feature space to discuss, analyse, transform, and synthesize spatial features.

These cases studies are the representations of approaches to practice as research, where they seek to represent aspects of the sound landscape through spatial audio as well as providing an interface for educational focus. Whatever intention is behind the exploration of spatial audio in an artistic context or through simulations, the mesh of interactions that tie sound and space together through references to people, histories, technologies and ideas are always present. Sonic experience is only possible through its exteriors (Sterne, 2003), and the sensations and experiences we have in the world belong to something that is outside of ourselves. This anamorphic perception of sound, is dependent on our listening position and surroundings, and unfold as non-linear, dynamic systems.

Chapter 5

Conclusions and further perspectives

The four preceding chapters have all sought to discuss various intersecting aspects and explorations of sound and spaces through a focus on sound objects and spatial audio. Through addressing theoretical, aesthetic, practical, and artistic concerns, the principal contribution of this thesis is through the conceptual and artistic exploration of the relationships between sound and space, into considering the spatial features afforded by place, site, and landscape as a structuring principle in spatial authoring.

Although there is much awareness within acousmatic music environments of the spatial relevance of sounds, this is still predominantly pursued through stereo diffusion as a preferred format for many composers. This thesis has investigated concepts that concern sound design, composed music, and sound installations and has adopted perspectives from different "genres" of experimental music, sound art, and the visual arts. From the perspectives of objects to structures, hopefully the discussions in this thesis will lead to further discussions and investigations into aspects of the site and place-bound, and its references in spatial audio. The following will summarise the preceding chapters and discuss some of their findings. Chapter 1 presented a conceptual and theoretical framework for sound and space, as different *models* for approximations and analysis–synthesis through various epistemological discussions around phenomenology, semiotics, metaphor, and cognition, loosely grouped around the heading *representations*. In this chapter I proposed a view of the sound object as somewhat "extended" from what Pierre Schaeffer originally proposed, where the sound object belongs to a complex mesh of interactions between people, histories, technologies, and ideas. This "extension" defines the sound object as consisting of three parts: a *shape*, a *site/location*, and a *model*. This represents the intrinsic and extrinsic features of the sound object. The shape refers to, among others, the sound's overall envelope, dynamic profile, and spectrum. An object references a site or location, either a real or an imaginary space or place that connects the sound object to *extrinsic* features. The model references how we analyse and synthesize the object and its subsequent iterations, the model guides our approximations between the object's shape and site, and most of the chapter is devoted to presenting these different perspectives. The relationship between the intrinsic and extrinsic is, as Smalley has stated, interactive.

Chapter 2 discusses the intrinsic and extrinsic features that make up our cognition of objects through the models of listening proposed by Schaeffer and Chion, and through the soundscape listening of Schafer, Truax, and Westerkamp, to the real-world listening of Norman and the ecological acoustics of Gaver. The perspectives on an object's feature, shape, and surface is discussed and identified as necessary for the contexts and relationships in spatial perception. These discussions centre around the semiotic and phenomenological aspects of sound perception and finds that it is with this spatiality of sound that description can begin (Ihde, 1976). This supports the idea that objects are distinguishable, three-dimensional wholes that exist through space and time. Events happen to these objects and sometimes change them. Objects are understood as *shapes* and *surfaces*, as an interface between medium and substance where energy is absorbed or reflected, vibrations are passed and where dissipation and diffusion occurs. These discussions reflect the notions that objects cannot be know outside of sensory experience, which finds that objects are conceptual structures and that most of this cognition happens in the unknown.

Further contextualised in chapter 3, these abstract, conceptual ideas of three-dimensional wholes are seen to reference a site/location through the concept of the *sound landscape*. This represents the *structures* that create the relationships between objects and their reference to an imagined source or location. Considering multichannel sound in terms of a landscape points to a perspective that seeks to move past the mere "panning" paradigm of moving sounds through the room, and instead consider the possible implications of spatial composition and the construction of complex spatial scenes. Extending the landscape as a metaphorical, *imagined* location of sources, the structural aspects of this chapter considers the place-bound and the site-specific as important features in spatial audio. Working with space in composition and sound art is never void of human presence and

> spaces produced by human activity I refer to as enacted spaces, and they can be divided into two primary types – utterance spaces, which are articulated by vocal sound, and agential spaces, where space is produced by human movement and (inter)action with objects, surfaces, substances, and built structures; we can also include human intervention in the landscape. (Smalley, 2007, p. 38)

Space is a boundless volume, where there are no clear edges. Rather, the move into the sound landscape posits an awareness of *place*: that which is lived (Merrifield, 1993) and that which we embody (Thrift, 2003). In this sense the use of spatial audio for composition, installations, simulations, experiments, and the like considers that sound and space are always *interactions* between themselves and that which they reference.

When writing about John Chowning's *Turenas* (1972), Smalley finds that the synthesised timbres bear no semblance to any aspect of the real world and "neither identity is source-bonded in the real-life sense (there is nothing 'realistic' about the timbres; there is no viable real cause), and any extrinsic link will relate to the velocity and spatial articulation of imagined sonic objects" (Smalley, 1993, p. 285). In this instance, source-bonding then refers to listening towards the cause of the sound and not to what the sound signifies. Perhaps the lack of metaphoric references in the work's title is what causes problems for Smalley's interpretation, which is different for the reading of Truax's *Riverrun* (see section 1.3.1). Despite this, *Turenas* demonstrates a high degree of sophistication in terms of its spatial composition through its use of Lissajous figures and Doppler shift, and creates an illusory *gait* though harmonic and inharmonic spectra among the four loudspeakers. The movements of the sound objects cannot be separated from the sound object's identity.

A source identity is connected to the perceived site it belongs to, however non-representational the source might be. The source is connected through its extrinsic features and "*external determinants* are related to influences existing outside the musical work. They are involved with the *pre-information* about the music which is bound to condition listeners' identifications" (Hoopen, 1994, p. 66). In this instance this refers to a programme note, score, or existing knowledge about the piece, however the external determinants of a given work are also the references to the imagined features of the sound landscape and the *acousmêtre*-like influences (see section 2.6.1) in loudspeaker-mediated sound.

In chapter 4, five case studies were presented as contextual discussions on the typomorphological framework. It was proposed how this feature space can be used to describe spatial features in sound and how the sound moves through space. The tasks of the typomorphology are identification, classification, and description, and each of the criteria are divided, subdivided, and sub-subdivided in a top-down, subjective exploration of feature categories. These feature categories affords an attention to timbre and mapping, where timbre is not considered as isolated musical values, but rather is described in terms of its dynamic morphologies that make up the sound colour, the sound characteristics, and our subjective perceptions of space. These dynamic morphologies also influence how we can translate the traditional term *mapping* into spatial audio, beyond the mere musical instrument discourse. This builds on the discussions from chapter 3 where we saw that the sounds we employ are place-bound. Mapping extends from the human body and its gestures to understandings of space and to a cartographic translation of physical and social forms. These series of interactions, can be composed and decomposed into new instances of interactions and can refer to what Schnell called an *action-action* relationship (Schnell, 2013). Space is a container for action, something involved in action and something that cannot be divorced from action (Tilley, 1994). The case studies presented in this thesis range from musical experimentation, installations, exhibitions, theatre, and pedagogical simulations through virtual reality and offer different perspectives on the application of the typomorphology and its feature space.

All of the processes involved in the analysis, description, classification, and synthesis of sound objects and space are dependent on one aspect, *listening*. Discussed in section 2.7, listening and sonic examination affords us a subjective judgement on what we hear that is different from, and a complement to, measurements and the visual study of spectrograms. The typomorphological framework affords us a feature space through reduced listening, that is, a mode of listening where we temporarily suspend our knowledge of the world and listen to a sound for *itself* to understand the sound and its building blocks. A listening schema affords us a way to listen *objectively* to *concrete* sounds, for indicators in external events - to *subjective* perceptions of the *concrete* sound to gain an idea of the object. Then, by selecting aspects of the sound, we make qualified *subjective* perceptions of the *abstracted* sound before we *objectively* perceive the sound content through signs and *abstract* values from them (Schaeffer, 2017, p. 83). Listening involves listening *for* something, this exists within the network of interactions we experience and listening allows for highlighting particular connections within this network.

The discussions throughout this thesis have hopefully demonstrated the relevance of Schaeffer's theories on the sound object to musical research today. It should be clear that the sound object is, through its existence as a multidimensional unit, a flexible and rigorous way to examine sound and space.

5.1 Proposals for further study

This thesis has only scratched the surface on a number of different topics that can all be explored further in future research. I have sought to draw on many different fields for the development of this project. At the outset this project set out to develop tools for spatialisation to afford flexibility for its users through algorithmic processes. However, developing tools is not enough, and we must also develop theories and methods, answer questions, and pose new questions. As such, this thesis has perhaps posed more questions than it has answered.

Specifically, Pierre Schaeffer's typomorphology should be further developed to afford a greater understanding of sound and space through their morphological dynamics. I will join Roads in his statement that "a formal theory of spatial relations remains to be developed" (Roads, 2015, p. 281) and through the combination of the rich current research on acoustics, psychoacoustics, signal processing, music information retrieval, and the cognitive sciences this can become a rich framework and feature space to describe space and the spatial occupancy along with the sound material, in a way where the subjective understandings between sound and space are preserved.

This thesis has sought to focus on ecologically valid research rather than controlled laboratory experiments. Furthermore, based on the theories developed in this thesis a series of perceptual experiments are planned. The psychoacoustic research presented in chapter 3 is often concerned with the evaluation of concert hall acoustics and from recorded classical music. There is still much to be studied from the perspective of spatial audio that goes beyond an evaluation of the "effectiveness" of an algorithm or continued studies of human sound localisation. Rather, listening tests for the evaluation of the criteria *spatial impression* and *spaciousness*, rather than just *immersion* and *envelopment*, can be studied through acousmatic music and complex timbral musical material, by drawing on the ideas presented as part of the sound landscape. However, it is in the combination with the criteria defined in the typomorphology that these experiments should be conducted, through a systematic exploration of this rich feature spaces. This can create room for a morphological spatial description, akin to a "Spatial Information Retrieval".[1]

These experiments will be pursued through both ambisonics and wave field synthesis. Along with listening tests to collect data on subjective perceptions around these psychoacoustic criteria, it would be of interest to track the listener's physical head movements during listening to acousmatic music through motion capture. A working hypothesis is that we will visually track or attempt to follow a heard sound source moving through space, despite there being no visual cues to follow.

In terms of software development, the projects and experiments contained in this thesis will be further developed through the Faust programming language[2] (Orlarey et al., 2009), a highly efficient language for functional DSP programming with a strong focus on sound synthesis and audio effects. The language features a wrapping system where objects written in Faust can be compiled to a number of different environments, including VST plugins, MaxMSP and Super-

[1] This perspective will need to draw as much on the research in Music Information Retrieval as from fields such as Geographic Information Retrieval.

[2] Functional audio stream: https://faust.grame.fr/

Collider, among many others. The different examples presented in section 4.3 will be implemented as standalone functions rather than having a continued existence as MaxMSP abstractions, to afford a functioning library for different platforms.

The stochastic approach discussed in section 4.5 has already been used by several other artists in a variety of contexts and should be further developed to form a stable library. This is planned to be extended to include several other applications of stochastic motion of sound to create a flexible tool for spatial audio, by including time and frequency domain splitting of the input sounds for individual spatial trajectories. This library will continue to be developed for amplitude-panned approaches. Future collaborations with the artist are being planned.

The continued development of the project discussed in section 4.8, is planned to be resumed in 2022 and although the main focus is on the rendering of "realistic" audience members and continued refining of the concert hall architecture, my main focus is a complete rewrite of the previously unfinished acoustic simulation. The reverberation algorithm in Spat5 has so far been used to create the simulated acoustics but when the project is resumed, this will be rewritten in Faust in order to avoid the closed source environments of MaxMSP and Spat. As previously discussed, in order to render an effective simulation of the acoustics, we will need rigorous testing of different musical instruments and their frequency-dependent radiation patterns to model how an instrument behaves on stage.

Future projects are envisaged from the ideas contained in this thesis. A new technological and artistic research project is underway, with the working title of *Transpositional Structures*,[3] this encapsulates a research project which investigates landscape and spatial synthesis through wave field synthesis (Berkhout, 1988) and procedural audio applications (Farnell, 2007, 2014) by drawing on interactive principles from game audio design (Collins, 2009; Zdanowicz & Bambrick, 2019). This project takes the site-specific and place-bound as a structural principle and seeks to develop computational methods for spatial synthesis by drawing on existing work on computational morphological description (Peeters & Deruty, 2008; Ricard & Herrera, 2004) for the interpretation of a landscape through spatial audio, with the development of a data-basing structure through Geographic Information Systems, descriptor-libraries to recreate surface, curve, and topological features of a given location, landscape, and site. Each dataset collected on site will form the basis for a point-cloud statistical analysis to extract spatial features from the site. This project will result in a series of exhibitions, concerts, software libraries, and publications.

It is my hope that the work contained in this thesis can contribute to a greater awareness of the "fabric of citations" contained within the uses of spatial audio

[3]The project has been generously supported by the Norwegian Arts Council and the Composers Renumeration Fund.

applications in composition, sound design and installations.

Bibliography

Adriaensen, F. (2006). Near field filters for higher order ambisonics. http://kokkinizita.linuxaudio.org/papers/hoafilt.pdf.

Ahrens, J. (2012). *Analytic methods of sound field synthesis*. Springer Science & Business Media.

Ando, Y. (1985). *Concert hall acoustics*. Springer series in Electrophysics 17.

Aufegger, L., Perkins, R., Wasley, D. & Williamon, A. (2017). Musicians' perceptions and experiences of using simulation training to develop performance skills. *Psychology of Music*, *45*(3), 417–431.

Augé, M. (2008). *Non-places: An introduction to supermodernity*. Verso.

Augoyard, J. F. & Torgue, H. (2014). *Sonic experience: A guide to everyday sounds*. McGill-Queen's Press-MQUP.

Baalman, M. A. (2003). Application of wave field synthesis in the composition of electronic music. *Proceedings of the International Computer Music Conference*.

Baalman, M. A. (2010). Spatial composition techniques and sound spatialisation technologies. *Organised Sound*, *15*(3), 209–218.

Barad, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press.

Barreiro, D. L. (2010). Sonic image and acousmatic listening. *Organised Sound*, *15*(1), 35–42.

Barrett, N. (2002). Spatio-musical composition strategies. *Organised Sound*, *7*(3), 313–323.

Barrett, N. (2007). *Trade Winds* [CD]. Aurora: ACD5056.

Barrett, N. (2010). Kernel Expansion: A three-dimensional spatial composition combining different ambisonics spatialisation techniques. *Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics, 6–7 May, Paris.*

Barrett, N. (2016). A musical journey towards permanent high-density loudspeaker arrays. *Computer Music Journal*, *40*(4), 35–46.

Barrett, N. (2019). Composing images in space: Schaeffer's allure projected in higher-order ambisonics. *Proceedings of the International Computer Music Conference*.

Barron, M. & Marshall, A. H. (1981). Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure. *Journal of sound and Vibration*, *77*(2), 211–232.

Barron, M. (1971). The subjective effects of first reflections in concert halls—the need for lateral reflections. *Journal of sound and vibration*, *15*(4), 475–494.

Barthes, R. (1977). *Image-music-text* (S. Heath, Ed. & Trans.). Fontana Press.

Bayle, F. (1993). *Musique acousmatique*. INA, Institut national de l'audiovisuel.

Bayle, F. (2007). Space, and more. *Organised Sound*, *12*(3), 241–249.

Becker, H. S. (2008). *Art worlds: Updated and expanded*. University of California Press.

Benjamin, E., Heller, A. & Lee, R. (2010). Design of ambisonic decoders for irregular arrays of loudspeakers by non-linear optimization. *Audio Engineering Society Convention 129*.

Berg, J. (2009). The contrasting and conflicting definitions of envelopment. *Audio Engineering Society Convention 126*.

Berg, J. & Rumsey, F. (2003). Systematic evaluation of perceived spatial quality. *AES 24th International Conference on Multichannel, Audio26-28 June 2003, Banff, Alberta, Canada*.

Berkhout, A. J. (1988). A holographic approach to acoustic control. *Journal of the audio engineering society*, *36*(12), 977–995.

Berkhout, A. J., de Vries, D. & Vogel, P. (1993). Acoustic control by wave field synthesis. *The Journal of the Acoustical Society of America*, *93*(5), 2764–2778.

Bijsterveld, K. (2019). *Sonic skills: Listening for knowledge in science, medicine and engineering (1920s-present)*. Springer Nature.

Bishop, C. (2005). *Installation art*. Tate.

Bissonnette, J., Dubé, F., Provencher, M. D. & Moreno Sala, M. T. (2015). Virtual reality exposure training for musicians: Its effect on performance anxiety and quality. *Medical problems of performing artists*, *30*(3), 169–177.

Black, J. (1997). *Maps and politics*. Reaktion Books.

Blake, E. (2002). Spatiality past and present: An interview with Edward Soja, Los Angeles, 12 April 2001. *Journal of Social Archaeology*, *2*(2), 139–158.

Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization*. MIT press.

Blesser, B. & Salter, L.-R. (2009). *Spaces speak, are you listening?: Experiencing aural architecture*. MIT Press.

Blumlein, A. D. (1931). UK patent specification 394,325 (improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems). *Journal of the Audio Engineering Society*, *6*(2), 91–130.

Bogost, I. (2008). *Unit operations: An approach to videogame criticism*. MIT press.

Bogost, I. (2012). *Alien phenomenology, or, what it's like to be a thing*. University of Minnesota Press.

Borch, C. (2002). Interview with Edward W. Soja: Thirdspace, postmetropolis, and social theory. *Distinktion: Scandinavian Journal of Social Theory*, *3*(1), 113–120.

Borgdorff, H. (2006). *The debate on research in the arts* (Vol. 2). Kunsthøgskolen i Bergen.

Borgdorff, H. (2012). *The conflict of the faculties: Perspectives on artistic research and academia*. Leiden University Press.

Born, G. (Ed.). (2013). *Music, sound and space: Transformations of public and private experience*. Cambridge University Press.

Boulanger, R. & Lazzarini, V. (Eds.). (2010). *The audio programming book*. MIT Press.

Bown, O., Eldridge, A. & McCormack, J. (2009). Understanding interaction in contemporary digital music: From instruments to behavioural objects. *Organised Sound*, *14*(2), 188.

Bradley, J., Reich, R. & Norcross, S. (2000). On the combined effects of early-and late-arriving sound on spatial impression in concert halls. *The Journal of the Acoustical Society of America*, *108*(2), 651–661.

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT Press.

Brümmer, L. (2017). Composition and perception in spatial audio. *Computer Music Journal*, *41*(1), 46–60.

Cage, J. (2012). *Silence: Lectures and writings*. Wesleyan University Press.

Carpentier, T. (2017). Ambisonic spatial blur. *Audio Engineering Society Convention 142*.

Carpentier, T., Barrett, N., Gottfried, R. & Noisternig, M. (2016). Holophonic sound in IRCAM's concert hall: Technological and aesthetic practices. *Computer Music Journal*, *40*(4), 14–34.

Carpentier, T., Noisternig, M. & Warusfel, O. (2015). Twenty years of IRCAM Spat: looking back, looking forward. *In Proceedings of the International Computer Music Conference*, 270–277.

Carter, P. (2004). Ambiguous traces, mishearing, and auditory space. In V. Erlmann (Ed.), *Hearing cultures: Essays on sound, listening and modernity* (pp. 43–63). Berg.

Cascone, K. (2000). The aesthetics of failure: "Post-digital" tendencies in contemporary computer music. *Computer Music Journal*, *24*(4), 12–18.

Caux, J. & Ferrari, L. (2012). *Almost nothing with Luc Ferrari: Interviews*. Errant Bodies Press.

Chadabe, J. (1997). *Electric sound: The past and promise of electronic music*. Prentice Hall.

Chadabe, J. (2002). The limitations of mapping as a structural descriptive in electronic instruments. *Proceedings of the 2002 conference on New interfaces for musical expression*, 1–5.

Cheng, E. (2015). *Cakes, custard, and category theory*. Profile Books.

Chion, M. (1994). *Audio-vision: Sound on screen*. Columbia University Press.

Chion, M. (2009). *Guide to sound objects: Pierre Schaeffer and musical research* (J. Dack & C. North, Trans.). http://www.ears.dmu.ac.uk.

Chion, M. (2011). Dissolution of the notion of timbre. *differences*, *22*(2-3), 235–239.

Chion, M. (2016). *Sound: An acoulogical treatise*. Duke University Press.

Chowning, J. M. (1971). The simulation of moving sound sources. *Journal of the Audio Engineering Society*, *19*(1), 2–6.

Clozier, C. (2001). The Gmebaphone concept and the Cybernéphone instrument. *Computer Music Journal*, *25*(4), 81–90.

Collins, K. (2009). An introduction to procedural music in video games. *Contemporary Music Review*, *28*(1), 5–15.

Cook, N. (2013). Classical music and the politics of space. In G. Born (Ed.), *Music, sound and space: Transformations of public and private experience* (pp. 224–238). Cambridge University Press.

Cosgrove, D. (1998). *Social formation and symbolic landscape*. University of Wisconsin Press.

Cosgrove, D. (2005). Maps, mapping, modernity: Art and cartography in the twentieth century. *Imago Mundi*, *57*(1), 35–54.

Cox, C. (2011). Beyond representation and signification: Toward a sonic materialism. *Journal of visual culture*, *10*(2), 145–161.

Cox, C. (2013). Sonic philosophy. *Artpulse: http://artpulsemagazine.com/sonic-philosophy (first accessed 09.10.20)*, *4*, 15.

Cox, C. (2017). Sonic thought. In B. Herzogenrath (Ed.), *Sonic thinking: A media philosophical approach* (pp. 99–109). Bloomsbury Academic.

Cox, C. & Warner, D. (2004). *Audio culture: Readings in modern music*. Bloomsbury Publishing USA.

Cox, T. (2014). *Sonic wonderland: A scientific odyssey of sound*. Random House.

Craven, P. G. & Gerzon, M. A. (1977). Coincident microphone simulation covering three dimensional space and yielding various directional outputs.

Cvejic, B. (2015). *Choreographing problems: Expressive concepts in contemporary dance and performance*. Palgrave Macmillan.

Daniel, J. (2003). Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format. *Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction*.

Daniel, J., Moreau, S. & Nicol, R. (2003). Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging. *Audio Engineering Society Convention 114*.

Dattorro, J. (1997a). Effect design, part 1: Reverberator and other filters. *Journal of the Audio Engineering Society*, *45*(9), 660–684.

Dattorro, J. (1997b). Effect design, part 2: Delay line modulation and chorus. *Journal of the Audio engineering Society*, *45*(10), 764–788.

Dattorro, J. (2002). Effect design, part 3: Oscillators: Sinusoidal and pseudonoise. *Journal of the Audio Engineering Society*, *50*(3), 115–146.

Davis, T. & Rebelo, P. (2005). Hearing emergence: Towards sound-based self-organisation. *http://eprints.bournemouth.ac.uk/17274/1/Hearing_Emergence.pdf*.

De Certeau, M. (1988). *The practice of everyday life*. University of California Press.

Deleuze, G. & Guattari, F. (1988). *A thousand plateaus: Capitalism and schizophrenia*. Bloomsbury Publishing.

Demers, J. (2010). *Listening through the noise: The aesthetics of experimental electronic music*. Oxford University Press.

Desantos, S., Roads, C. & Bayle, F. (1997). Acousmatic morphology: An interview with François Bayle. *Computer Music Journal*, 11–19.

de Saussure, F. (1959). *Course in general linguistics*. McGraw-Hill Books Library.

Di Scipio, A. (1994). Micro-time sonic design and timbre formation. *Contemporary Music Review*, *10*(2), 135–148.

Di Scipio, A. (2003). Sound is the interface: From interactive to ecosystemic signal processing. *Organised Sound*, *8*(3), 269–277.

DiBiase, D. (1990). Visualization in the earth sciences. *Earth and Mineral Sciences*, *59*(2), 13–18.

Döbereiner, L. (2011). Models of constructed sound: Nonstandard synthesis as an aesthetic perspective. *Computer Music Journal*, *35*(3), 28–39.

Dodge, C. & Jerse, T. A. (1997). *Computer music: Synthesis, composition and performance*. Schirmer Books.

Eco, U. (1979). *A theory of semiotics*. Indiana University Press.

Ekeberg, F. (2013). Manipulating space, changing realities: Space as primary carrier of meaning in sonic arts.

Ellison, S. (2013). SpaceMap: 20 years of audio origami. *Lighting and Sound America*, 80–88.

Emmerson, S. (1986). *The language of electroacoustic music*. Ashgate Publishing Limited.

Emmerson, S. (2007). *Living electronic music*. Ashgate Publishing Limited.

Eno, B. (1996). *A year with swollen appendices*. Faber & Faber.

Entrikin, J. N. (1991). *The betweenness of place: Towards a geography of modernity*. The Johns Hopkins University Press.

Essl, K. (1992). Lexikon Sonate. *http://www.essl.at/sounds.html#lexson-porgy*.

Farnell, A. (2007). An introduction to procedural audio and its application in computer games. *Audio mostly conference*, *23*.

Farnell, A. (2010). *Designing sound*. MIT Press.

Farnell, A. (2014). Procedural audio theory and practice. In K. Collins, B. Kapralos & T. H (Eds.), *The Oxford handbook of interactive audio* (pp. 532–540). Oxford University Press.

Fels, P. (1996). 20 years of the Delta Stereophony System – High-quality sound design. *Audio Engineering Society Convention 100*.

Ferrari, L. (2009). *L'œuvre Electronique* [CD]. INA-GRM: Ina G 6017/6026.

Filimowicz, M. & Stockholm, J. (2010). Towards a phenomenology of the acoustic image. *Organised Sound*, *15*(1), 5.

Forsythe, W. & Kaiser, P. (1999). Dance geometry. *Performance Research*, *4*(2), 64–71.

Frayling, C. (1993). *Research in art and design*. Royal College of Art.

Gabor, D. (1947). Acoustical quanta and the theory of hearing. *159*(4044), 591–594.

Garnett, G. E. (1991). Music, signals, and representations: A survey. In A. P. Giovanni De Poli & C. Roads (Eds.), *Representations of musical signals* (pp. 325–369). MIT Press.

Gaver, W. W. (1993a). How do we hear in the world? Explorations in ecological acoustics. *Ecological psychology*, *5*(4), 285–313.

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

Geier, M. & Spors, S. (2012). Spatial audio with the SoundScape Renderer. *27th Tonmeistertagung*.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive science*, *7*(2), 155–170.

Gentner, D. & Bowdle, B. (2008). Metaphor as structure-mapping. In R. Gibbs (Ed.), *The cambridge handbook of metaphor and thought* (pp. 109–128). Cambridge University Press.

Gerzon, M. A. (1974). Surround-sound psychoacoustics. *Wireless World*, *80*(1468), 483–486.

Gerzon, M. A. (1992). General metatheory of auditory localisation. *Audio Engineering Society Convention 92*.

Gibson, J. J. (1982). What is involved in surface perception? In J. Beck (Ed.), *Organization and representation in perception* (pp. 151–157). Lawrence Erlbaum.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Lawrence Erlbaum.

Godøy, R. I. (1997). *Formalization and epistemology*. Scandinavian University Press.

Godøy, R. I. (2006). Gestural-sonorous objects: Embodied extensions of Schaeffer's conceptual apparatus. *Organised sound*, *11*(2), 149–157.

Godøy, R. I. (2018). Sonic object cognition. In R. Bader (Ed.), *Springer handbook of systematic musicology* (pp. 761–777). Springer Verlag.

Godøy, R. I. (2019). Thinking sound-motion objects. In M. Filimowicz (Ed.), *Foundations in sound design for interactive media: A multidisciplinary approach*. Routledge.

Godøy, R. I. (2021). Perceiving sound objects in the musique concrète. *Frontiers in Psychology*, *12*, 1702.

Goguen, J. (1999). An introduction to algebraic semiotics, with application to user interface design. In L. Nehavic (Ed.), *Computation for metaphors, analogy, and agents* (pp. 242–291). Springer.

Gottfried, R. (2016). Studies on the compositional use of space. *Technical report. Institut de Recherche et Coordination Acoustique/Musique (IRCAM)*.

Green, O., Tremblay, P. A. & Roma, G. (2019). Interdisciplinary research as musical experimentation: A case study in musicianly approaches to sound corpora. *Electroacoustic Music Studies Network Conference: Electroacoustic Music: Is it Still a Form of Experimental Music?*

Grey, J. M. & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *The Journal of the Acoustical Society of America*, *63*(5), 1493–1500.

Griesinger, D. (1996). Spaciousness and envelopment in musical acoustics. *Audio Engineering Society Convention 101*.

Griesinger, D. (1997). The psychoacoustics of apparent source width, spaciousness and envelopment in performance spaces. *Acta Acustica united with Acustica*, *83*(4), 721–731.

Griesinger, D. (1999). Objective measures of spaciousness and envelopment. *Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction*.

Griffiths, T. D. & Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*, *5*(11), 887–892.

Grond, F. & Hermann, T. (2014). Interactive sonification for data exploration: How listening modes and display purposes define design guidelines. *Organised Sound*, *19*(1), 41–51.

Haas, H. (1972). The influence of a single echo on the audibility of speech. *Journal of the Audio Engineering Society*, *20*(2), 146–159.

Hagan, K. L. (2017). Textural composition: Aesthetics, techniques, and spatialization for high-density loudspeaker arrays. *Computer Music Journal*, *41*(1), 34–45.

Halmrast, T. (2000). Orchestral timbre: Comb-filter coloration from reflections. *Journal of Sound and Vibration*, *232*(1), 53–69.

Halmrast, T. (2011). More combs. *Proceedings of the Institute of Acoustics (UK) Vol. 33, Dublin*, *33*.

Halmrast, T. (2015). Acoustics in between: Perception of sound in rooms beyond standard criteria. *Psychomusicology: Music, Mind, and Brain*, *25*(3), 256.

Halmrast, T. (2018). Attack the attack: Reverberation influences attack and timbre. *Proceedings of the Institute of Acoustics, Hamburg*, *40*, 335–345.

Hannula, M., Suoranta, J. & Vadén, T. (2014). *Artistic research: Methodology: Narrative, power and the public*. Peter Lang.

Harley, M. A. (1998). Music of sound and light: Xenakis's polytopes. *Leonardo*, *31*(1), 55–65.

Harman, G. (2002). *Tool-being: Heidegger and the metaphysics of objects*. Open Court.

Harman, G. (2010). *Prince of networks: Bruno Latour and metaphysics*. re. press.

Harman, G. (2011a). *The quadruple object*. Zero books.

Harman, G. (2011b). The road to objects. *continent*, *3*(1), 171–179.

Harrison, J., Vaggione, H. & Wishart, T. (2004). *ETC* [CD]. Electronic Music Foundation – EMF CD 053.

Harvey, J. (1981). "Mortuos Plango, Vivos Voco": A realization at IRCAM. *Computer Music Journal*, *5*(4), 22–24.

Harvey, J. (1999). *Tombeau De Messiaen* [CD]. Sargasso: SCD 28029.

Hatfield, J. (2016). Performing at the top of one's music game: The mental edge of musicianship. *Frontiers in Psychology*, *7*, 1–16.

Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). Blackwell Publishing.

Heller, A. J., Benjamin, E. & Lee, R. (2012). A toolkit for the design of ambisonic decoders. *Linux Audio Conference, CCRMA Stanford University*, 1–12.

Helmholtz, H. L. (1885/1954). *On the sensations of tone as a physiological basis for the theory of music*. Dover Publications.

Holbrook, U. (2018). An approach to stochastic spatialization - a case of Hot Pocket. In T. M. Luke Dahl Douglas Bowman (Ed.), *Proceedings of the* *international conference on new interfaces for musical expression* (pp. 31–32). Virginia Tech.

Holbrook, U. (2019). Sound objects and spatial morphologies. *Organised Sound*, *24*(1), 20–29.

Hollerweger, F. (2005). An introduction to Higher Order Ambisonics. *http://flo.mur.at/writings/HOA-intro.pdf*.

Holmes, T. (2012). *Electronic and experimental music: Technology, music, and culture*. Routledge.

Hoopen, C. T. (1994). Issues in timbre and perception. *Contemporary Music Review*, *10*(2), 61–71.

Housefield, J. (2002). Marcel Duchamp's art and the geography of modern Paris. *Geographical Review*, *92*(4), 477–502.

Howard, D. M. & Angus, J. (2009). *Acoustics and psychoacoustics*. Taylor & Francis.

Hunt, A. & Wanderley, M. (2002). Mapping performance parameters to synthesis engine. *Organised Sound*, *7*(2), 103–114.

Hunt, A., Wanderley, M. M. & Paradis, M. (2003). The importance of parameter mapping in electronic instrument design. *Journal of New Music Research*, *32*(4), 429–440.

Husserl, E. (2012). *Ideas: General introduction to pure phenomenology*. Routledge.

Huygens, C. (1690). *Traité de la lumière*. P. van der Aa.

Ihde, D. (1976). *Listening and voice: A phenomenology of sound*. Ohio University Press.

Ingold, T. (2005). The eye of the storm: Visual perception and the weather. *Visual studies*, *20*(2), 97–104.

Irwin, R. (1996). Being and circumstance: Notes toward a confidential art. In K. Stiles & P. Selz (Eds.), *Theories and documents of contemporary art* (pp. 572–574). University of California Press.

James, S. (2015). Spectromorphology and spatiomorphology of sound shapes: Audio-rate aep and dbap panning of spectra.

Jot, J.-M. (1999). Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces. *Multimedia systems*, *7*(1), 55–69.

Kahn, D. (2001). *Noise, water, meat: A history of sound in the arts*. MIT press.

Kane, B. (2007). "L'Objet Sonore Maintenant": Pierre Schaeffer, Sound Objects and the Phenomenological Reduction. *Organised sound*, *12*(1), 15–24.

Kane, B. (2014). *Sound unseen: Acousmatic sound in theory and practice*. Oxford University Press.

Kane, B. (2015). Sound studies without auditory culture: A critique of the ontological turn. *Sound Studies*, *1*(1), 2–21.

Kendall, G. S. (1995a). A 3-D sound primer: Directional hearing and stereo reproduction. *Computer music journal*, *19*(4), 23–46.

Kendall, G. S. (1995b). The decorrelation of audio signals and its impact on spatial imagery. *Computer Music Journal*, *19*(4), 71–87.

Kendall, G. S. (2008). What is an event? the EVENT schema, circumstances, metaphor and gist. *Proceedings of the International Computer Music Conference*.

Kendall, G. S. (2010). Meaning in electroacoustic music and the everyday mind. *Organised Sound*, *15*(1), 63–74.

Kendall, G. S. & Ardila, M. (2008). The artistic play of spatial organization: Spatial attributes, scene analysis and auditory spatial schemata. *International Symposium on Computer Music Modeling and Retrieval*, 125–138.

Kenny, D. T., Davis, P. & Oates, J. (2004). Music performance anxiety and occupational stress amongst opera chorus artists and their relationship with state and trait anxiety and perfectionism. *Journal of anxiety disorders*, *18*(6), 757–777.

Kim, S.-J. (2010). Imaginal listening: A quaternary framework for listening to electroacoustic music and phenomena of sound-images. *Organised Sound*, *15*(1), 43–53.

Kim-Boyle, D. (2005). Sound spatialization with particle systems. *Proceedings of the 8th international conference on digital audio effects (DAFX-05), Madrid, Spain*, 65–68.

Kim-Boyle, D. (2008). Spectral spatialization - An overview. *Proceedings of the International Computer Music Conference*.

Kim-Cohen, S. (2009). *In the blink of an ear: Towards a non-cochlear sonic art*. Bloomsbury Publishing.

Kim-Cohen, S. (2016a). Against ambience. *Against ambience and other essays* (pp. 3–81). Bloomsbury Publishing.

Kim-Cohen, S. (2016b). No depth: A call for shallow listening. *Against ambience and other essays* (pp. 131–143). Bloomsbury Publishing.

Kirkegaard, J. (2006). *4 Rooms* [CD]. Touch - Tone: 26.

Kittler, F. A. (1999). *Gramophone, film, typewriter*. Stanford University Press.

Klapholz, J. (1991). Fantasia: Innovations in sound. *Journal of the Audio Engineering Society*, *39*(1/2), 66–70.

Krauss, R. (1979). Sculpture in the expanded field. *October*, *8*, 31–44.

Kubovy, M. & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, *80*(1-2), 97–126.

Kwon, M. (2002). *One place after another: Site-specific art and locational identity*. MIT Press.

Laban, R. (1963). *Modern educational dance*. MacDonald; Evans Ltd.

LaBelle, B. (2015). *Background noise: Perspectives on sound art*. Bloomsbury Publishing.

LaBelle, B. (2019). Sonic site-specificities. In P. Weibel (Ed.), *Sound art: Sound as a medium of art* (pp. 518–527). MIT Press.

Laitinen, M.-V., Pihlajamäki, T., Erkut, C. & Pulkki, V. (2012). Parametric time-frequency representation of spatial sound in virtual worlds. *ACM Transactions on Applied Perception (TAP)*, *9*(2), 8.

Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh*. Basic books.

Lakoff, G. & Johnson, M. (2008). *Metaphors we live by*. University of Chicago Press.

Landau, B. & Jackendoff, R. (1993). "What"and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, *16*, 217–265.

Landy, L. (1991). *What's the matter with today's experimental music?: Organized sound too rarely heard*. Harwood Academic Publishers.

Landy, L. (1994). The "something to hold on to factor" in timbral composition. *Contemporary Music Review*, *10*(2), 49–60.

Landy, L. (2007). *Understanding the art of sound organization*. MIT Press.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.

Latour, B. (1996). On actor-network theory: A few clarifications. *Soziale welt*, 369–381.

Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.

Lefebvre, H. (1991). *The production of space* (D. Nicholson-Smith, Trans.). Blackwell.

Lennox, P., Myatt, T. & Vaughan, J. (1999). From surround to true 3-D. *Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction*.

Lennox, P., Myatt, T. & Vaughan, J. (2001). 3D audio as an information environment. *Audio Engineering Society Conference: 19th International Conference: Surround Sound-Techniques, Technology, and Perception*.

Leonard, B. (2017). Applications of extended multichannel techniques. In A. Roginska & P. Geluso (Eds.), *Immersive sound:The art and science of binaural and multi-channel audio* (pp. 333–355). Routledge.

Leonelli, S. (2019). What distinguishes data from models? *European journal for philosophy of science*, *9*(2), 22.

Lesne, A. & Bourgine, P. (2010). Introduction. In P. Bourgine & A. Lesne (Eds.), *Morphogenesis: Origins of patterns and shapes* (pp. 1–14). Springer Science & Business Media.

Lewis, G. E. (2000). Too many notes: Computers, complexity and culture in Voyager. *Leonardo Music Journal*, 33–39.

Licht, A. (2007). *Sound art*. Rizzoli.

Licht, A. (2009). Sound art: Origins, development and ambiguities. *Organised Sound*, *14*(1), 3–10.

Lockwood, A. (1989). *A sound map of the Hudson River* [CD]. Lovely Music: LCD 2081.

Lokki, T. (2013). Throw away that standard and listen: Your two ears work better. *Building Acoustics*, *20*(4), 283–293.

Lokki, T. (2014). Tasting music like wine: Sensory evaluation of concert halls. *Physics Today*, *67*(1), 27. https://doi.org/10.1063/PT.3.2242

Lokki, T. (2016). Why is it so hard to design a concert hall with excellent acoustics? *Dim*, *1*, 73–13.

Lokki, T. & Pätynen, J. (2020). Auditory spatial impression in concert halls. In J. Blauert & J. Braasch (Eds.), *The technology of binaural understanding* (pp. 173–202). Springer.

Long, M. (2005). *Architectural acoustics*. Elsevier.

Lossius, T. (2008). Controlling spatial sound within an installation art context. *Proceedings of the International Computer Music Conference*.

Lossius, T. & Anderson, J. (2014). ATK Reaper: The Ambisonic Toolkit as JSFX plugins. *Proceedings of the International Computer Music Conference*.

Lossius, T., Baltazar, P. & de la Hogue, T. (2009). DBAP – Distance-Based Amplitude Panning. *Proceedings of the International Computer Music Conference*.

Loy, G. (2006). *Musimathics, volume 1: The mathematical foundations of music*. MIT Press.

Loy, G. (2007). *Musimathics, volume 2: The mathematical foundations of music*. MIT Press.

Lucier, A. (1981). *I am sitting in a room* [CD]. Lovely Music: 7-4529-51013-2-1.

Lucier, A. (1995). *Reflections: Interviews, scores, writings* (Vol. 3). MusikTexte.

Luening, O. (1964). Some random remarks about electronic music. *Journal of Music Theory*, *8*(1), 89–98.

Luque, S. (2009). The stochastic synthesis of Iannis Xenakis. *Leonardo Music Journal*, 77–84.

Lütticken, S. (2010). Art and thingness, part 1: Breton's ball and Duchamp's carrot. *e-flux journal*, *13*(2), 16.

Lyon, E. (2014). The future of spatial computer music. *In Proceedings of the International Computer Music Conference*.

MacCallum, J., Gottfried, R., Rostovtsev, I., Bresson, J. & Freed, A. (2015). Dynamic message-oriented middleware with Open Sound Control and O.dot. *Proceedings of the International Computer Music Conference*.

MacEachren, A. (1994). Visualization in modern cartography: Setting the agenda. In A. MacEachren & F. Taylor (Eds.), *Visualization in modern cartography* (pp. 1–12). Pergamon.

Magnusson, T. (2009). Of epistemic tools: Musical instruments as cognitive extensions. *Organised Sound*, *14*(2), 168–176.

Maletic, V. (1987). *Body-space-expression: The development of Rudolf Laban's movement and dance concepts*. Mouton de Gruyter.

Malham, D. (1998). Spatial hearing mechanisms and sound reproduction. *University of York*.

Malham, D. (2001a). Spherical harmonic coding of sound objects – the ambisonic 'O' format. *Audio Engineering Society Conference: 19th International Conference: Surround Sound-Techniques, Technology, and Perception*.

Malham, D. (2001b). Toward reality equivalence in spatial sound diffusion. *Computer Music Journal*, *25*(4), 31–38.

Malham, D. & Myatt, A. (1995). 3-D sound spatialization using ambisonic techniques. *Computer music journal*, *19*(4), 58–70.

Manning, P. (1993). *Electronic and computer music*. Oxford University Press.

Marcolli, M. (2020). *Lumen naturae: Visions of the abstract in art and mathematics*. MIT Press.

Mattern, S. (2016). Cloud and field. *Places Journal*. *https://placesjournal.org/article/cloud-and-field/*.

Matthews, K. (1997). *cd Ann* [CD]. Annette Works.

Matthews, K. (1998). *cd Bea* [CD]. Annette Works.

Mayhew, S. (2015). *A dictionary of geography*. Oxford Paperback. https://www.oxfordreference.com/view/10.1093/acref/9780199231805.001.0001/acref-9780199231805.

Mayo, B. (1961). Objects, events, and complementarity. *The Philosophical Review*, *70*(3), 340–361.

Mazzola, G. (1997). *Semiotics of music: A handbook on the sign-theoretic foundations of nature and culture*. Walter de Gruyter.

McCormack, L., Politis, A. & Pulkki, V. (2021). Parametric spatial audio effects based on the multi-directional decomposition of ambisonic sound scenes. *Proceedings of the 24th International Conference on Digital Audio Effects (DAFx20i2n02)*.

Mera, M. (2016). Towards 3-D sound: Spatial presence and the space vacuum. In L. Greene & D. Kulezic-Wilson (Eds.), *The palgrave handbook of sound design and music in screen media* (pp. 91–111). Springer.

Merleau-Ponty, M. (2012). *Phenomenology of perception*. Routledge.

Merrifield, A. (1993). Place and space: A Lefebvrian reconciliation. *Transactions of the institute of British geographers*, *18*(4), 516–531.

Mershon, D. H. & King, L. E. (1975). Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics*, *18*(6), 409–415.

Miranda, E. R. (2001). *Composing music with computers*. CRC Press.

Miranda, E. R. & Wanderley, M. M. (2006). *New digital musical instruments: Control and interaction beyond the keyboard* (Vol. 21). AR Editions, Inc.

Moore, B. (2003). *An introduction to the psychology of hearing*. Academic Press.

Moorer, A. J. (1985). Signal processsing aspects of computer music. In J. Straw (Ed.), *Digital audio signal processing* (pp. 149–220). A-R Editions Inc.

Nagel, T. (1974). What is it like to be a bat? *The philosophical review*, *83*(4), 435–450.

Nattiez, J.-J. (1990). *Music and discourse: Toward a semiology of music*. Princeton University Press.

Nelson, R. (2013). *Practice as research in the arts: Principles, protocols, pedagogies, resistances*. Palgrave Macmillan.

Neuhaus, M. (2000). Sound art? *Introduction to the exhibition "Volume: Bed of Sound," at the PS.1 Contemporary Art Center, New York, July*.

Nicol, R. (2017). Sound field. In A. Roginska & P. Geluso (Eds.), *Immersive sound* (pp. 290–324). Focal Press.

Noisternig, M., Zotter, F. & Katz, B. F. (2011). Reconstructing sound source directivity in virtual acoustic environments. In Y. Suzuki & D. Brungart (Eds.), *Principles and applications of spatial hearing* (pp. 357–372). World Scientific.

Norberg-Schulz, C. (1971). *Existence, space & architecture*. Studio Vista.

Norman, K. (1996). Real-world music as composed listening. *Contemporary Music Review*, *15*(1-2), 1–27.

Normandeau, R. (2009). Timbre spatialisation: The medium is the space. *Organised Sound*, *14*(3), 277–285.

November, V., Camacho-Hübner, E. & Latour, B. (2010). Entering a risky territory: Space in the age of digital navigation. *Environment and planning D: Society and space*, *28*(4), 581–599.

Nyström, E. (2018). Topographic synthesis: Parameter distribution in spatial texture. *Proceedings of the 2018 International Computer Music Conference*, 117–122.

O'Callaghan, C. (2007). *Sounds: A philosophical theory*. Oxford University Press.

Okamoto, T., Katz, B. F., Noisternig, M., Iwaya, Y. & Suzuki, Y. (2011). Implementation of real-time room auralization using a surrounding 157 loudspeaker array. In Y. Suzuki & D. Brungart (Eds.), *Principles and applications of spatial hearing* (pp. 373–382). World Scientific.

Oliveros, P. (1995). Acoustic and virtual space as a dynamic element of music. *Leonardo Music Journal*, *5*(1), 19–22.

Olsen, B. (2010). *In defense of things: Archaeology and the ontology of objects*. AltaMira Press.

O'Modhrain, S. (2011). A framework for the evaluation of digital musical instruments. *Computer Music Journal*, *35*(1), 28–42.

Opie, T. & Brown, A. (2006). An introduction to eco-structuralism. *Proceedings of the International Computer Music Conference 2006*, 9–12.

Opie, T. & Brown, A. (2009). Eco-structural data forms. *improvise*, 123.

Orlarey, Y., Fober, D. & Letz, S. (2009). FAUST: An efficient functional approach to DSP programming. *Editions DELATOUR FRANCE. NEW COMPUTATIONAL PARADIGMS FOR COMPUTER MUSIC*, 65–96.

Otondo, F. (2007). Creating sonic spaces: An interview with Natasha Barrett. *Computer Music Journal*, *31*(2), 10–19.

Palombini, C. (1993). Machine songs V: Pierre Schaeffer: From research into noises to experimental music. *Computer Music Journal*, *17*(3), 14–19.

Parmegiani, B. (2010). *L'œuvre Musicale En 12 CD* [CD]. INA/GRM: INAG6000/11.

Parry, N. (2014). Exploded sounds: Spatialised partials in two recent multi-channel installations. *Divergence Press*, *3*(1). http://divergencepress.net/2014/12/01/2014-12-1-exploded-sounds-spatialised-partials-in-two-recent-multi-channel-installations/

Peeters, G. G. & Deruty, E. (2008). Automatic morphological description of sounds. *Journal of the Acoustical Society of America*, *123*(5), 3801.

Peirce, C. S. (1960). *Collected papers of Charles Sanders Peirce: Principles of philosophy and elements of logic* (C. Hartshorne & P. Weiss, Eds.; Vol. 1). Harvard University Press.

Penha, R. & Oliveira, J. (2013). Spatium, tools for sound spatialization. *Proceedings of the Sound and Music Computing Conference*, 660–667.

Peters, N. (2011). *Sweet [re]-production: Developing sound spatialization tools for musical applications with emphasis on sweet spot and off-center perception*. PhD-thesis.

Peters, N., Lossius, T., Schacher, J., Baltazar, P., Bascou, C. & Place, T. (2009). A stratified approach for sound spatialization. *Proceedings of the Sound and Music Computing Conference, Porto, Portugal*.

Peters, N., Lossius, T. & Schacher, J. C. (2012). SpatDIF: Principles, specification, and examples. *Proceedings of the 9th Sound and Music Computing Conference*, 500–505.

Peters, N., Lossius, T. & Schacher, J. C. (2013). The spatial sound description interchange format: Principles, specification, and examples. *Computer Music Journal*, *37*(1), 11–22.

Peters, N., Marentakis, G. & McAdams, S. (2011). Current technologies and compositional practices for spatialization: A qualitative and quantitative analysis. *Computer Music Journal*, *35*(1), 10–27.

Petitot, J. (1989). Perception, cognition and morphological objectivity. *Contemporary Music Review*, *4*(1), 171–180.

Petitot, J. (1999). Morphological eidetics for a phenomenology of perception. In J. Petitot, F. Varela, B. Pachoud & J.-M. Roy (Eds.), *Naturalizing phenomenology: Issues in contemporary phenomenology and cognitive science* (pp. 330–371). Stanford University Press.

Petitot, J. (2011). The morphodynamical turn of cognitive linguistics. *Signata. Annales des sémiotiques/Annals of Semiotics*, (2), 61–80.

Plenge, G. (1974). On the differences between localization and lateralization. *The Journal of the Acoustical Society of America*, *56*(3), 944–951.

Politis, A., Tervo, S. & Pulkki, V. (2018). Compass: Coding and multidirectional parameterization of ambisonic sound scenes. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6802–6806.

Porcello, T. (2004). Speaking of sound: Language and the professionalization of sound-recording engineers. *Social Studies of Science*, *34*(5), 733–758.

Potard, G. & Burnett, I. (2004). Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays. *Proceedings of the International Conference on Digital Audio Effects (DAFx'04)*.

Pulkki, V. (1997). Virtual sound source positioning using Vector Base Amplitude Panning. *Journal of the audio engineering society*, *45*(6), 456–466.

Pulkki, V. (2000). Generic panning tools for Max/MSP. *Proceedings of the International Computer Music Conference*.

Pulkki, V., Lokki, T. & Rocchesso, D. (2011). Spatial effects. In U. Zölzer (Ed.), *DAFx: Digital audio effects* (pp. 139–184). Wiley Online Library.

Raaijmakers, D. (2000). *Cahier "M" – A brief morphology of electric sound*. Leuven University Press.

Rayleigh, L. (1909). On the perception of the direction of sound. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *83*(559), 61–64.

Reiss, J. D. & McPherson, A. (2014). *Audio effects: Theory, implementation and application*. CRC Press.

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH computer graphics*, *21*(4), 25–34.

Rheinberger, H.-J. (2018). Transpositions: From traces through data to models and simulations. In M. Schwab (Ed.), *Transpositons - aesthetico-epistemic operators in artistic research* (pp. 215–224). Leuven University Press.

Ricard, J. & Herrera, P. (2004). Morphological sound description: Computational model and usability evaluation. *Audio Engineering Society Convention 116*.

Risset, J.-C. (1988). *Songes* [CD]. WERGO: WER 2013-50.

Risset, J.-C. (1991). Timbre analysis by synthesis: Representations, limitations, and variants for musical composition. In A. P. Giovanni De Poli & C. Roads (Eds.), *Representations of musical signals* (pp. 7–43).

Risset, J.-C. & Wessel, D. L. (1999). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The psychology of music* (pp. 113–169). Academic Press.

Roads, C. (1996). *The computer music tutorial*. MIT Press.

Roads, C. (2004). *Microsound*. MIT Press.

Roads, C. (2005). The art of articulation: The electroacoustic music of Horacio Vaggione. *Contemporary Music Review*, *24*(4-5), 295–309.

Roads, C. (2015). *Composing electronic music: A new aesthetic*. Oxford University Press.

Rocchesso, D. & Fontana, F. (2003). *The sounding object*. Mondo estremo.

Rolfe, C. & Keller, D. (2000). Decorrelation as a by-product of granular synthesis. *Proc. of the XIII Colloquium on Musical Informatics*.

Rosenblum, L. D. (2004). Perceiving articulatory events: Lessons for an ecological psychoacoustics. In J. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 219–248). Elsevier.

Rowe, R. (1992). *Interactive music systems: Machine listening and composing*. MIT Press.

Rudi, J. (2018). Unpacking the musical and technical innovation of Knut Wiggen. *Organised Sound*, *23*(2), 195–207.

Rudi, J. (2019). *Elektrisk lyd i Norge fra 1930 til 2005*. Novus Forlag.

Rumsey, F. (2001). *Spatial Audio*. Focal Press.

Rumsey, F. (2002). Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm. *Journal of the Audio Engineering Society*, *50*(9), 651–666.

Russolo, L. (1913). *The art of noise* (R. Filliou, Trans.). Something else press.

Samkopf, K. (1994). *Mårådalen Walk* [CD]. Grappa: HCD2907.

Sauer, C. (1925/2008). The morphology of landscape. In T. Oakes & P. P (Eds.), *The cultural geography reader* (pp. 108–116). Routledge.

Savioja, L., Huopaniemi, J., Lokki, T. & Väänänen, R. (1999). Creating interactive virtual acoustic environments. *Journal of the Audio Engineering Society*, *47*(9), 675–705.

Sazdov, R., Paine, G. & Stevens, K. (2007). Perceptual investigation into envelopment, spatial clarity, and engulfment in reproduced multi-channel audio. *AES 31st International Conference*.

Scaletti, C. (2018). Sonification$\neq$ music. In A. McLean & R. Dean (Eds.), *The oxford handbook of algorithmic music* (pp. 363–385). Oxford University Press.

Schacher, J. C. (2010). Seven years of ICST ambisonics tools for MaxMSP – a brief report. *Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics*.

Schacher, J. C. & Kocher, P. (2006). Ambisonics spatialization tools for Max/MSP. *In Proceedings of the International Computer Music Conference*.

Schaeffer, P. (2012). *In search of a concrete music*. University of California Press.

Schaeffer, P. (2015). *Le Trièdre Fertile* [LP]. Recollection GRM: REGRM 001.

Schaeffer, P. (2017). *Treatise on musical objects: An essay across disciplines*. University of California Press.

Schafer, R. M. (1994). *The soundscape: Our sonic environment and the tuning of the world*. Destiny Books.

Schnell, N. (2013). *Playing (with) sound: Of the animation of digitized sounds and their reenactment by playful scenarios in the design of interactive audio applications*. PhD-thesis.

Schnupp, J., Nelken, I. & King, A. (2012). *Auditory neuroscience: Making sense of sound*. MIT press.

Schrimshaw, W. (2017). *Immanence and immersion: On the acoustic condition in contemporary art*. Bloomsbury Publishing.

Schroeder, M. R. (1962). Natural sounding artificial reverberation. *Journal of the Audio Engineering Society*, *10*(3), 219–223.

Schwab, M. (2018). Introduction. In M. Schwab (Ed.), *Transpositons - aesthetico-epistemic operators in artistic research* (pp. 7–21). Leuven University Press.

Sèdes, A., Guillot, P. & Paris, E. (2014). The HOA library, review and prospects. *International Computer Music Conference— Sound and Music Computing*, 855–860.

Seiffarth, C. (Ed.). (2017). Maia Urstad, Urban Sound art XIII [Exhibition catalog].

Seldess, Z. (2014). MIAP: Manifold-Interface Amplitude Panning in Max/MSP and pure data. *Audio Engineering Society Convention 137*.

Shaw, M. L. & Gaines, B. R. (1989). Comparing conceptual structures: Consensus, conflict, correspondence and contrast. *Knowledge Acquisition*, *1*(4), 341–363.

Siedenburg, K. & McAdams, S. (2017). Four distinctions for the auditory "wastebasket" of timbre. *Frontiers in psychology*, *8*, 1–4.

Smalley, D. (1993). Defining transformations. *Journal of New Music Research*, *22*(4), 279–300.

Smalley, D. (1997). Spectromorphology: Explaining sound-shapes. *Organised Sound*, *2*(2), 107–126.

Smalley, D. (2007). Space-form and the acousmatic image. *Organised sound*, *12*(1), 35–58.

Smith, A. D. (2003). *Routledge philosophy guidebook to Husserl and the Cartesian meditations*. Routledge.

Smith, J. (1985). An introduction to digital filter theory. In J. Straw (Ed.), *Digital audio signal processing* (pp. 69–135). A-R Editions Inc.

Smith, J. (1991). Viewpoints on the history of digital synthesis. *Proceedings of the International Computer Music Conference*, 1–10.

Smith, J. O. (2007). *Introduction to digital filters: With audio applications*. W3K Publishing.

Smith, S. W. (1999). *The scientist and engineer's guide to digital signal processing*. California Technical Publishing.

Smithson, R. (1996). *Robert Smithson: The collected writings* (J. Flam, Ed.). University of California Press.

Solnit, R. (2006). *A field guide to getting lost*. Canongate Books.

Sommer Eide, E. (2019). *The Waves* [CD]. Sofa Music.

Stautner, J. & Puckette, M. (1982). Designing multi-channel reverberators. *Computer Music Journal*, *6*(1), 52–65.

Steinberg, J. C. & Snow, W. B. (1934). Auditory perspective — physical factors. *Electrical Engineering*, *53*(1), 12–17.

Steintrager, J. A. (2018). Listening with Adorno, Again. In J. A. Steintrager & R. Chow (Eds.), *Sound objects* (pp. 73–93). Duke University Press.

Sterne, J. (2003). *The audible past*. Duke University Press.

Sterne, J. (2011). The theology of sound: A critique of orality. *Canadian Journal of Communication*, *36*(2), 207–225.

Sterne, J. (2012). Sonic imaginations. In J. Sterne (Ed.), *The sound studies reader* (pp. 1–17). Routledge.

Strogatz, S. H. (2015). *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. Westview Press.

Stroustrup, B. (2014). *Programming: Principles and practice using C++*. Addison-Wesley.

Subotnick, M. (1967). *Silver Apples Of The Moon* [LP]. Nonesuch: H-71174.

Surel Lange, T. (2021). *Works for Listening 1-10* [CD]. Sofa Music.

Thompson, D. W. (1992). *On growth and form: The complete revised edition*. Dover Publications, Inc.

Thompson, E. A. (2004). *The soundscape of modernity: Architectural acoustics and the culture of listening in america, 1900-1933*. MIT press.

Thompson, W. F., Graham, P. & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, *2005*(156), 203–227.

Thompson, W. F., Russo, F. A. & Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cognition and Emotion*, *22*(8), 1457–1470.

Thoresen, L. (2015). *Emergent musical forms: Aural explorations*. Studies in Music, Department of Music Research; Composition, Don Wright Faculty of Music, University of Western Ontario.

Thoresen, L. & Hedman, A. (2007). Spectromorphological analysis of sound objects: An adaptation of Pierre Schaeffer's typomorphology. *Organised Sound*, *12*(2), 129–141.

Thrift, N. (2003). Space: The fundamental stuff of geography. In S. R. Nicholas Clifford Sarah Holloway & G. Valentine (Eds.), *Key concepts in geography* (pp. 95–107). Sage Thousand Oaks.

Tilley, C. (1994). *A phenomenology of landscape: Places, paths, and monuments*. Berg Publishers.

Toop, D. (1995). *Ocean of sound: Aether talk, ambient sound and imaginary worlds*. Serpent's Tail.

Toop, D. (2004). *Haunted weather: Music, silence and memory*. Serpent's Tail.

Topper, D. (2000). On anamorphosis: Setting some things straight. *Leonardo*, *33*(2), 115–124.

Topper, D., Burtner, M. & Serafin, S. (2003). Spatio-operational spectral (SOS) synthesis. *Proceedings of the International Computer Music Conference*.

Torchia, R. H. & Lippe, C. (2004). Techniques for multi-channel real-time spatial distribution using frequency-domain processing. *Proceedings of the 2004 conference on New interfaces for Musical Expression*, 116–119.

Traer, J. & McDermott, J. H. (2016). Statistics of natural reverberation enable perceptual separation of sound and space. *Proceedings of the National Academy of Sciences*, *113*(48), E7856–E7865.

Trochimczyk, M. (2001). From circles to nets: On the signification of spatial sound imagery in new music. *Computer music journal*, *25*(4), 39–56.

Truax, B. (1988). *Digital soundscapes* [CD]. WERGO: WER 2017-50.

Truax, B. (1998). Composition and diffusion: Space in sound in space. *Organised Sound*, *3*(2), 141–146.

Truax, B. (2001). *Acoustic communication*. Ablex Publishing Corporation.

Tuuri, K. & Eerola, T. (2012). Formulating a revised taxonomy for modes of listening. *Journal of New Music Research*, *41*(2), 137–152.

Vaggione, H. (2017). *Fluides* [CD]. empreintes DIGITALes: IMED 17145 / 2017.

Välimäki, V., Parker, J., Savioja, L., Smith, J. O. & Abel, J. (2016). More than 50 years of artificial reverberation. *Audio Engineering Society conference: 60th international conference: Dreams (Dereverberation and reverberation of audio, music, and speech)*.

Van Nort, D., Wanderley, M. M. & Depalle, P. (2014). Mapping control structures for sound synthesis: Functional and topological perspectives. *Computer Music Journal*, *38*(3), 6–22.

Van Valkenburg, D. & Kubovy, M. (2004). From Gibson's fire to gestalts: A bridge–building theory of perceptual objecthood. In J. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 113–147). Elsevier.

Van Veen, B. D. & Buckley, K. M. (1988). Beamforming: A versatile approach to spatial filtering. *IEEE assp magazine*, *5*(2), 4–24.

Vilkamo, J., Lokki, T. & Pulkki, V. (2009). Directional audio coding: Virtual microphone-based synthesis and subjective evaluation. *Journal of the Audio Engineering Society*, *57*(9), 709–724.

Warren, W. H. & Verbrugge, R. R. (1984). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. *Journal of Experimental Psychology: Human perception and performance*, *10*(5), 704.

Waters, S. (2000). Beyond the acousmatic: Hybrid tendencies in electroacoustic music. In S. Emmerson (Ed.), *Music, electronic media and culture* (pp. 56–83). Ashgate Publishing Limited.

Watson, C. (2013). *In St Cuthbert's Time* [CD]. Touch: TO89.

Weibel, P. (2019). *Sound art: Sound as a medium of art*. MIT Press.

Wessel, D. L. (1979). Timbre space as a musical control structure. *Computer music journal*, 45–52.

Westerkamp, H. (2002). Linking soundscape composition and acoustic ecology. *Organised Sound*, *7*(1), 51–56.

Westerkamp, H. (2010). *Transformations* [CD]. empreintes DIGITALes: IMED 1031 / 2010.

Williamon, A., Aufegger, L. & Eiholzer, H. (2014). Simulating and stimulating performance: Introducing distributed simulation to enhance musical learning and performance. *Frontiers in Psychology*, *5*, 25.

Wilson, M. & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological bulletin*, *131*(3), 460.

Wilson, S. (2008). Spatial swarm granulation. *In Proceedings of the International Computer Music Conference*.

Wilson, S. & Harrison, J. (2010). Rethinking the BEAST: Recent developments in multichannel composition at Birmingham ElectroAcoustic Sound Theatre. *Organised Sound*, *15*(3), 239–250.

Winderen, J. (2018). *Spring bloom in the marginal ice zone* [CD]. Touch: Tone 65.

Windsor, L. (2000). Through and around the acousmatic: The interpretation of electroacoustic sounds. In S. Emmerson (Ed.), *Music, electronic media and culture*. Ashgate.

Winkler, T. (2001). *Composing interactive music: Techniques and ideas using Max*. MIT Press.

Wishart, T. (1994). *Audible design: A plain and easy introduction to practical sound composition*. Orpheus the Pantomime.

Wishart, T. (1996). *On sonic art*. Routledge.

Wood, D. & Fels, J. (2008). The natures of maps: Cartographic constructions of the natural world. *Cartographica: The International journal for geographic information and geovisualization*, *43*(3), 189–202.

Worrall, D. (1998). Space in sound: Sound of space. *Organised Sound*, *3*(2), 93–99.

Wright, M. (2005). Open Sound Control: An enabling technology for musical networking. *Organised Sound*, *10*(3), 193–200.

Wright, M., Beauchamp, J., Fitz, K., Rodet, X., Röbel, A., Serra, X. & Wakefield, G. (2000). Analysis/synthesis comparison. *Organised sound*, *5*(3), 173–189.

Wright, M., Freed, A. & Momeni, A. (2003). Open Sound Control: State of the art 2003. *Proceedings of the 2003 conference on New interfaces for musical expression*, 153–160.

Xenakis, I. (1992). *Formalized music: Thought and mathematics in composition*. Pendragon Press.

Zdanowicz, G. & Bambrick, S. (2019). *The game audio strategy guide: A practical course*. Routledge.

Ziemer, T. (2017). Source width in music production. methods in stereo, ambisonics, and wave field synthesis. In A. Schneider (Ed.), *Studies in musical acoustics and psychoacoustics* (pp. 299–340). Springer.

Ziemer, T. & Bader, R. (2015a). Complex point source model to calculate the sound field radiated from musical instruments. *Proceedings of Meetings on Acoustics 170ASA*, *25*(1), 035001.

Ziemer, T. & Bader, R. (2015b). Implementing the radiation characteristics of musical instruments in a psychoacoustic sound field synthesis system. *Audio Engineering Society Convention 139*.

Zotter, F. & Frank, M. (2012). All-round ambisonic panning and decoding. *Journal of the audio engineering society*, *60*(10), 807–820.

Zotter, F. & Frank, M. (2018). Ambisonic decoding with panning-invariant loudness on small layouts (AllRAD2). *Audio Engineering Society Convention 144*.

Zotter, F. & Frank, M. (2019). *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality*. Springer Nature.