

Exploring 3D Audio as a New Musical Language

Timothy Schmele

Submitted for the award of a
Master of Science in Sound and Music Computing.

Music Technology Group
Department of Information and Communication Technologies
Universitat Pompeu Fabra

Supervisors: Pau Arumí, Josep Comacunjosas
Submitted to the Pompeu Fabra University, September, 2011

Declaration

Full Title: Exploring 3D Audio as a New Musical Language

Supervisors: Pau Arumí

Josep Comacunjosas

This thesis is presented in partial fulfillment of the requirements for the degree of Master of Science in Sound and Music Computing. It is entirely my own work and has not been submitted to any other university or higher education institution, or for any other academic award in this university. Where use has been made of the work of other people it has been fully acknowledged and fully referenced

Signature:

Timothy Schmele

September 1, 2011 – Barcelona, Catalonia, Spain

Acknowledgements

I would like to thank both my supervisors Pau Arumí and Josep Comacunjosas for all their help, advice and support. Many thanks go out to Toni Mateos, as well as the whole crew at Barcelona Media for their help and for giving me the opportunity to carry out my work in their facilities. Without the technology they provided and the trust Barcelona Media showed in giving me full access to the studio, this thesis would be unthinkable!

Thank you also to the Sound and Music Computing Master program coordinator, Xavier Serra, and all the staff at the Music Technology Group, UPF and ESMUC for sharing their knowledge and providing their competence with the upmost respect.

Furthermore, I want to stress my gratitude to a few people who explicitly helped by providing their expertise: Alexandros Katsakos, for an inspiring discussion about the aural perception of blind people, Felix Tutzer & Jan van Balen, for the many five-in-the-morning brainstorming sessions, Imanol Gomez, for his great collaboration audifying brainstorming sessions, Katrin Neue, for her expertise on postmodern dance, Kerry Hagan, for her valuable insights into textural composition, Arthur Sauer, for his help on Wave Field Synthesis, as well as Robert Sazdov and Natasha Barrett, who shared their research without hesitation and especially Maja Trochimczyk, for scanning in and sharing her complete PhD thesis on her own initiative!

I am also grateful to the German Academic Exchange Service (DAAD) for selecting and giving me the graduate scholarship for German students studying abroad, without which my whole studies at the Pompeu Fabra University and my life here in Barcelona would not be as fruitful as it was!

Special thanks go to all my class mates, the chungos and the mijas, at the Sound and Music Computing Master program for all their support, help and especially for their friendship, with whom I shared some great, memorable times during the creation of this work. It was highly inspiring to work in such an international group of bright and creative minds. Let's have one more after this is finished... or five...

Very special thanks and love to my parents and family for all the support and love one could possibly wish for. Although I copy and paste this sentence every time, its meaning is as strong as ever!

Xenakis!

Abstract

This thesis explores the artistic potential to make use of three dimensional, physical space as a new musical language. It does so in three ways: It first investigates the greater history of spatialization and traces its epitaxy as an emancipated musical parameter up to recent developments since the new millennium. The second part then brings these historic ideas together into a contemporary context, in the effort to further establish space a prime vessel able to carry a musical message of expression. Several approaches to spatial composition are elaborated, keeping previously mentioned psychoacoustic limitations in mind. Finally, in the practical part of the thesis, two different three-dimensional spatial sound synthesis methods are presented, followed by a qualitative evaluation of their use. Based on the insights – and especially the complications encountered relating to the success of this approach, the thesis concludes by arguing the necessary conditions under which space as a musical parameter might hopefully be established.

Σύνοψη¹

Η διατριβή αυτή εξερευνά την καλλιτεχνική δυνατότητα χρησιμοποίησης του τρισδιάστατου φυσικού χώρου σαν μια καινούρια μουσική γλώσσα. Το επιτυγχάνει με τρεις τρόπους: Αρχικά, διερευνά την μακρά ιστορία της χωροδιάστασης και εξετάζει την ανέλιξη της σαν μια ανεξάρτητη μουσική παραμετρος μέχρι και τις τελευταίες επιτεύξεις της νέας χιλιετίας. Το δεύτερο μέρος ενσωματώνει αυτές τις ιστορικές ιδέες σε ένα σύγχρονο πλαίσιο, στην προσπάθεια κατοχύρωσης του χώρου σαν κύριο μέσο μεταφοράς στην έκφραση ενός μουσικού μηνύματος. Παρατίθεται ανάλυση των διαφόρων μεθόδων μουσικής σύνθεσης στον χώρο, έχοντας πρωτίστως υπόψη τους περιορισμούς της ψυχοακουστικής. Τέλος, στο πρακτικό μέρος της διατριβής, παρουσιάζονται δύο διαφορετικές συνθέσεις ήχου σε τρισδιάστατο χώρο, η χρήση των οποίων ακολουθείται από ποιοτική αξιολόγηση. Με βάση τις γνώσεις, και κυρίως τις πολυπλοκότητες που συνδέονται με την επιτυχία της μεθόδου αυτής, η διατριβή ολοκληρώνεται επιχειρηματολογώντας τις αναγκαίες συνθήκες κάτω από τις οποίες ο χώρος θα μπορέσει επιτέλους να κατοχυρωθεί σαν μια μουσική παράμετρος.

¹..για την θεία προσωποπρέπεια του!

Contents

Declaration	i
Acknowledgements	ii
Abstract	iv
1 Introduction	1
2 Historical Backgrounds	9
2.1 From Antiphony to Theater: Space in Acoustic Ensembles	10
2.1.1 Polychorality in Antiphonic Music	11
2.1.2 Space as a Means of Dramaturgy	13
2.1.3 Symbolical Spaces	17
2.2 Technological Developments in Multichannel Speaker Arrays	19
2.3 Spatial Parameters in the Electro-acoustic Tradition	25
2.3.1 A Space of Moving Objects	26
2.3.2 Serialization of the Spatial Parameter	29
2.3.3 Musical Space in the Computer Age	34
3 Spatial Composition Methodology	41
3.1 Psychoacoustic Implications	43
3.1.1 Localization on the Spherical Periphery	43
3.1.2 Distance Perception	48
Experiment I: Assessing the impact of linear signal expansion	53
Experiment II: Panning audio into the center of the listening space	55
3.1.3 Ecological Psychoacoustics	58
3.2 On the Notion of Space	62
3.3 Composing the Space	71
3.3.1 Point Source and Trajectorial Approaches	75
Decorrelation effects	83
<i>3dSpamlePlayer</i> : a Max/MSP spatialization tool	85
3.3.2 Architectural Pieces and Soundscapes	89
3.3.3 Virtual Spaces and Spatial Textures	92

4	Spatial Sound Synthesis	99
4.1	Timbral Spatialization	102
4.1.1	Implementation of an extensible Spectral Splitter	103
4.1.2	Conclusions on the Spectral Splitter	107
4.2	Rapid Panning Modulation Synthesis	109
4.2.1	Implementation of the Sphere Panning Spatial Synthesis Module	111
4.2.2	Using Spynth	114
4.2.3	Gestural Synthesis	117
5	Conclusions	119
5.1	Redefining the Musical Conception of a Whole Culture	122
	Bibliography	128

Chapter 1

Introduction

”The space parameter is no more an effect in pitch music, but pitch is only an effect in space music. Space as a finality in music expression.” –Leo Kupper¹

Establishing the physical space² as a musical language might alienate those who think of music in terms of pitches and melodies, but if one looks at the history of musical development, spatial aspects and considerations have always been part of the respective periods own respective aesthetics. Space is present in most musical vocabulary, as well as projected into many other musical characteristics and parameters. Traces can be found starting from the deliberate separation of ensemble parts to articulate antiphonal compositions in biblical times, and continuing with architecturally motivated compositions over symbolical spaces up to virtual soundscapes. This is possibly due to the recent developments in audio technology in the past 100 years. As pitch disappears, new building blocks such as sine tones, concrete sounds and noises are used. Since melodies and harmonies do not hold anymore, there is a need for a new parameter which brings more depth into the musical expression. Space and the concept of spatialization in electronic music today is a substantiated aspect of the music. The question, asked throughout this thesis, is the effectiveness in how far the spatial considerations in any example given herein constitutes as a true musical language.

The way space is used in the electronic music domain is unique and radically new compared to any previous acoustic effort is an accepted fact (Harley 1994, Smalley 2007, Normandeau 2009). As Normandeau (2009) points out, the development of the loudspeaker had a fundamental impact on the way composers see space. The loudspeaker is the ultimate instrument, not only being able to emulate any timbre, but it has the unique ability to separate and exchange a timbre between several speakers. The psychoacoustic effect is that the timbre will be spread between the space in-between the loudspeakers. If this exchange is done via amplitude panning techniques between two speakers, for example, the sound will give the impression to continuously wander from one speaker to the next. Hence, the space becomes continuous. A speaker and its position in space is suddenly much more closely linked compared to an acoustic

¹As quoted in Normandeau (2009)

²Note aside: The reader should be advised that in the thesis, when referring to space or, most specifically, musical space, it is the physical space surrounding the listener that is meant, not timbral space or pitch space or any other space that could be considered musical.

instrument. Its relationship with its position is more apparent than in instrumental music. The impact of this awareness was immense: suddenly sounds could fly around the audience without having any mass connected. The sound could move away in all directions thinkable, fully detached from its source and origin:

”Die Bewegung der Klänge ist also nicht mehr notwendig an den Körper gebunden. Das ist etwas sehr Wichtiges. [...] Die Tatsache, daß ein Klang oben erscheint und sich mit einer Geschwindigkeit, die schneller ist, als je ein Mensch sich bewegen könnte, von oben links nach halbrunden rechts mit einem bestimmten Winkel und einem bestimmten Grad innerhalb meines Wahrnehmungsfeldes bewegt, kann genauso eine wesentliche Information sein wie eine Quarte in einer Tonhöhenmelodie”³ (Stockhausen 1964)

Although use of space is a common factor in electronic music, its use can often be belittled as an effect to enhance the listening experience. Multichannel works are squashed into a stereo image without hesitation. If the message can be delivered in a stereo mix-down of the piece, then, speaking in the true sense of this thesis, where is the intention of the spatialization through more than two channels other than a garnishing effect? Stereo mix-downs are always required for medial distribution (over a CD or stereo file), broadcast over the radio or as an application to a festival or something similar for peer-reviewers to examine. This means that all musical information must be contained in the reduced two channel stereo version. Otherwise, the composer might run the risk of not being understood and rejected. But if that is the case, then the act of spatialization over more channels during the actual concert either means: (a) the composer is either able to compensate the loss of spatiality via live diffusion or the piece can in fact carry itself removing a complex, intricate and meaningful spatialization, or (b) the spatialization elevates the listening experience, but brings no new message across, hence it solely embellishes the piece. While the second point is valid in its own account, this is exactly what this thesis is trying to avoid.

Composing space becomes more complex as we move from the one dimensional melodic lines to the three dimensional space. This is similar to timbral composition which even lacks any sort of dimensional thinking as each timbre is a category in itself. Similar to how timbral composition has not yet fully established itself in the regular musical thinking yet, space is far away from being grasped by the average general music listener. Also, just like a specific orchestration can sound right on an unconscious level (a rock band for the rock music fan), a specific spatialization may make one composition or production sound better than another. What I am seeking is a true use of space as a musical parameter, not just an effect, nor a dramaturgic additive. Neither do I want to use space as a tool or a means to clarify musical complexity. While the other approaches have their validity in their respective context. Using

³”The movement of the sounds is therefore no longer necessarily attached to the body. That is something very important. [...] The fact that a tone appears from above and moves from the upper left to half-below on the right with a certain angle and a defined level within my field of perception at a speed which is faster than any person has ever been able to move, can be exactly an essential piece of information as a fourth in a pitch melody.”

physical space as a musical language goes beyond the above mentioned examples of its use. The following quote by Stockhausen could simply have the word *timbre* replaced with *space* to state my desire to define this new language as timbral composition is now relatively established:

”Dasselbe trifft für die Klangfarben zu, die noch bis Ende des vorigen und Anfang dieses Jahrhunderts unemanzipiert komponiert wurden, das heißt, sie dienten einfach zur Klärung von Verhältnissen, die sowieso schon geschaffen waren in Harmonik und Melodik. Selbst die Rhythmik diente im wesentlichen der Klärung der harmonischen Verhältnisse. [...] Schönbergs Konzeption, daß man mit Klangfarben genauso Musik machen könnte wie vorher mit Veränderungen der Tonhöhen, ist ja bis heute noch nicht ins Bewußtsein der meisten Musiker gedrungen; daß man zum Beispiel ein Stück machen könnte, in dem die Tonhöhe völlig konstant ist, für eine halbe Stunde, und dieselbe Information musikalisch komponiert und wahrgenommen sein könnte - nur durch Veränderungen der Klangfarben -, wie früher in einer melodischen Komposition. Wenn jetzt eine Systematik für die Ordnung der Klangfarben einsetzte ähnlich wie für die Intervalle, so daß wir auch Intervalle der Farbigkeit komponieren und wahrnehmen könnten, mit denen wir genauso sinnvoll Musik machten, indem wir die Tonhöhen einfach einmal neutralisierten, so hätten wir eine Gleichberechtigung von Tonhöhen- und Klangfarbenkomposition erreicht.”⁴(Stockhausen 1964)

Instead, composers of spatial music will face the circumstance that the exploitation of their music is more difficult due to the fact of limited facilities to perform such music, especially considering three dimensional spatialization (Lynch & Sazdov 2011b). But, even more importantly – and more restrictively – the music will hardly ever be heard in any setting other than a concert situation. This means, that situations of showing the music to colleges and simply interested people (i.e. exploitation through the internet) is almost impossible – being able to prove oneself to reviewers and evaluators is just as much, without having concerts to which the composer, though, would have to apply in the first place. (Peters 2010) shows a survey, which demonstrates that the majority of composers (58%) only use the space as a means to enhance the listening experience, i.e. create a sense of engulfment. The mere sensation of envelopment is reason enough for most composers to engage in the spatialization of their works. Lynch & Sazdov (2011b) states, for example, that based on ”studies [...] and observations made by composers of electroacoustic music [...] the most expressed term can be concluded to be enveloped, or surrounded by sound.” While envelopment is a prime aspect of spatial music, it

⁴”The same is true for timbres, which were still composed in an unemancipated way until the end of the last century and beginning of this one. That means they were only used to clarify the relationships which were already there for harmony and melody. Even rhythm was used to clarify the harmonic relationships [...] Schönbergs concept that music could also be composed with timbres as before with changing tone pitches has not reached the consciousness of the composers today, i.e. that a piece could be composed in which one pitch stays constant for half an hour and the same information is musically composed and perceived only by changing the timbres, just as in a melodic composition before. If we were to establish a taxonomy to order timbres, similarly as it exists for intervals, so that we can compose with intervals of colorations, with which we can also create meaningful music by simply neutralizing pitches once, and perceive these, then we would achieve an equal status between pitched and timbral composition.”

is a sole embellishing sensation. The research undergone in this thesis went to investigate three dimensional sound as a complete new way for composers to express themselves, which calls for much more sophisticated methods. Coming back to the study conducted in (Peters 2010), 40% of the composers questioned stated they organize and structure their sounds spatially, which can be seen as a listening enhancement on the one side, but perhaps is meant as an artistic expression as well. Interestingly, 44% feel exactly similar about space as will be discussed in this thesis, namely that their efforts to use space is meant as an expressive musical act.

Space is only then a musical parameter, if the meaning and the intent behind the piece suffers greatly or, in the truest sense, is removed completely. In all other cases, space is just an effect or a quality of a recording that heightens the listening pleasure. The lines are blurry and not well defined, though, of course. Although this comparison might be a bit fatuitous, take, for example, industry style recordings done with a Mid/Side configuration (M/S). When looked at from a technological point of view, within the range of the stereo image all spatiality is present. But what would constitute this spatiality in a musical sense? No matter how much thought the artist might have put into the stereo image, the fact that the recording was done in M/S indicates that there was a certain drive for mono compatibility, be it for the intent of radio airplay or whichever motivation. The piece should have similar qualities when played in mono just as it did in stereo. Therefore, the sole fact that M/S was used to record the music does not make it spatial. If the music can live with its spatial properties, then how would you argue it spatial musicality? I never heard of a Mozart being played on only one pitch, just because the performer brought a didgeridoo. Or, would Scelsi's famous one-note timbral compositions still be the same if we changed and removed the instrumentation? Composers of spatial music are often asked to provide "stereo mixes" of their works. What does it mean to the spatial composer if his pieces are reduced in spatial dimensionality, spatial freedom? "[Normandeau] feels frustration from the fact that his multichannel works are often produced in stereo format." (Zelli 2010) There is not just a need for a spatial awareness at the audience, but also at the organizers and concert hall owners. I am somewhat aware of the technical and social problems connected with this request, but at the same time, this thesis argues and discusses problems that are noticeable throughout the complete academic domain (Weale 1999). These problems are most likely affecting spatial music itself probably more than any other current musical movement.

Discussing this topic is important, since not enough has been brought together in this context. As recent studies show (Marentakis et al. 2008, Kendall & Cabrera 2011) composers still fall into the same traps, because there exists no complete framework for spatial composition. I myself had to experience the disappointment when first entering the three dimensional VBAP studio at Barcelona Media and notice in an instant that all my previously prepared spatial designs were infantile and, basically, worthless. I had to revise my whole approach to spatialization, and this is where the line ends for most. I believe that due to the limitations noticed with spatialization at first, many practitioners perhaps do not feel the need to investigate further into the matter. This is why spatialization is usually a transportation means for the message, but never truly the message itself. This might be the reason, why the ini-

tial efforts of the likes of Varèse, Stockhausen and Xenakis were left aside as experiments and why their spatial interpretation came so much later, compared to other analytical approaches. Furthermore, spatialization in a sophisticated way is a large logistic effort in both equipment and technique. Many sounds have to be controlled through space while they also have to be generated, of course.

The topicality of this issue can further be demonstrated in the many PhDs referenced in this work, most of which are written no more than about five years ago⁵. See the early exception in [Harley \(1994\)](#), the acoustic spatial analysis in [Solomon \(2007\)](#) or the analysis Stockhausen's spatial efforts undergone in [Miller \(2009\)](#) (and the PhDs referenced therein), that discuss fundamentals on compositional methodology on spatialization and the interpretation and analysis. On the other hand, fundamental writings by composers on the subject, such as Stockhausen himself are over 50 years old now – Ives almost 100 years, and even Mozart writes of echo effects in his scores for several orchestras! For one, this shows how new the concept is, but also demonstrates for how long it has been ignored! Spatialization as a musical language is just beginning to be noticed, recognized and perceived in its full potential. An ever radical idea needs time to be soaked up, and one big argument is time we need to give spatial works to be accepted by the general society. That said, it also shows that the ideas of the hundred year old thinkers in spatial music finally bear fruit. Needless to say, that today one can find many more composers that actively engage with the theoretical dispute of this aesthetic, such as Barry Traux, Dennis Smalley and Robert Normandeau. As [Smalley \(2007\)](#) himself also writes: "Although there has been much of value written about spatial attributes and the role of space, mainly by composers, the thinking is somewhat scattered, and as yet there is no substantial, unified text on the topic, nor any solid framework which might provide a reasonably secure basis for investigating space."

Another, larger topic is the perception of space and how we map it out from our early development on. One perspicuous reason why humans would rather rely on vision to guide their hearing is given in [Handel \(2006\)](#): "Visual objects usually are opaque, so that the illumination of each spatial point comes from one object. But the acoustic signal is transparent, so that the air pressure intensity at each temporal point may come from more than one sound-producing object." Although people might perceive spatial gestures – and they do so at least unconsciously – they are just not accustomed to listening to them in a musical way. At the same time, be it either the technology not working correctly (considering psychoacoustic principles) or our perception of audible space being too varying and personal, composers need to be aware that their efforts and intentions might be terribly mistakingly interpreted as others will simply not hear complex gestures or even listen to wrong, sometimes even non existing or intended spatial patterns. Due to the sensibility of many spatialization technologies (see section 2.2) for an introduction on the matter) and the partly still misunderstood psychoacoustic principles,

⁵Or max. fifteen years, as can be seen a major work on the topic given in [Harley \(1994\)](#). In her introduction, she also mentions over 400 pieces of literature on spatial aspects in music, but criticizes the fact there is no whole methodology. By what I can judge, her work is probably the first to bring this notion to a level of importance that the minds of Brandt, Varèse or Stockhausen had strived for.

physical space as a compositional framework needs to accustom to the fragility of the mimetic spatial image.

”Da gewinnt die Musik nun eine ungeheure Bedeutung. Sie kann zum Beispiel durch die Erweiterung des räumlichen Parameters dazu beitragen, daß eine Wahrnehmungsphänomenologie, die ebenso auf dem Akustischen wie auf dem Visuellen basiert, allmählich wieder entwickelt wird.”⁶ (Stockhausen 1964)

The practical work and valuable insights gained due to the same were acquired by kindly being able to use the VBAP⁷ studio at Barcelona Media in Barcelona, Spain. The studio consists of 24 speakers in an approximate dome-like setup with a server running the VBAP algorithm. The connection to the server is done via a single ethernet cable (see figure 1.1). Originally, the audio was planned to be sent over an audio interface, but before the work in the studio could start, an ethernet solution to send audio buffers from Jack⁸ over the network was implemented. Each channel in Jack corresponded to the source’s identification number. The source management, including the addition and deletion of sources, was done over TCP, requiring a secure answering system for the client to maintain a individual source management system itself. In this case, the source management at client side was realized using Max/MSP⁹ and a [coll]-object, but the actual TCP communication had to be done via PureData¹⁰ using a combination of the [packOSCstream]- and [tcpclient]-objects, because Max/MSP did not provide any way of maintaining a persistent TCP connection. The spatial information was sent in UDP packets, containing the source identification number, the two angles θ and ϕ , as well as the distance d (see figure 1.2). A special modification of the VBAP algorithm, Barcelona Media calls PBAP would allow a fourth parameter specifying the source width w . The spatial information, tied to a specific channel in Jack, would be translated into amplitude variations by the server, sending the source out to the respective loudspeakers responsible for the virtual source. This approach allowed for the necessary abstraction from the low level algorithms and the high level usage to be applied.

The thesis is structured the following way:

- Chapter 2 will try to look at the history of spatial music to a greater extent. It is grouped into different general identified trends in spatial music, starting from antiphony in the acoustic domain, over the dramatic use of space and musical theater to symbolic spaces, reaching into spiritual spaces and symbols spatially laid out by a specific localized seating plan of the ensemble, for example. The chapter then focuses more on the electronic use of space in section 2.2, first introducing a short history of the most common spatialization

⁶”The music takes on a enormous relevance. For example, by expanding the spatial parameters it can contribute the perceptual phenomenology, which is both acoustically as well as the visually based, being gradually developed again.”

⁷VBAP stands for *vector based amplitude panning*. For more information on this other spatialization technologies, see chapter 2.2.

⁸<http://jackaudio.org/>

⁹<http://cycling74.com>

¹⁰<http://puredata.info>

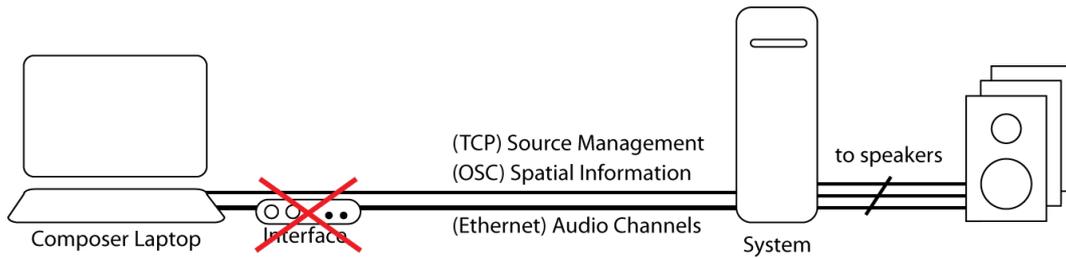


Figure 1.1: Depicting the audio setup in the studio at Barcelona Media.

technologies and then going over to more contemporary ideas of spatial use in section 2.3. Here the space shows a clear development with advancements of technology and increase of computational power. As the methods of dispersion become more advanced, so do the thoughts on the possibilities of using space in a musical context.

- The theoretical part of the thesis, chapter 3, is divided into two parts. The first part encompasses the psychoacoustic discussion in section 3.1 as well as the arrival at the metaphysical discussion of the notion of space in section 3.2. This discussion serves as a ground for the following ideas on different techniques that use space musically. They are divided into three general groups: the first being the trajectorial approaches of section 3.3.1, which are mainly concerned with the rather conventional movement of sounds through space. Nevertheless, contemporary trends and new ideas are sought out as this approach is found to be fundamental to most other approaches in spatial music. Shortly introducing the matter of site-specific installations and soundscape compositions in section 3.3.2, the text goes over to a rather new and radical notion of spatial sounds and virtual space created by enveloping synthesis techniques in section 3.3.3.
- These virtual synthesis techniques are then presented in their own chapter 4, due to being the main practical contribution of this thesis. It mainly looks into two possibilities: the first is an extension and improvement to previously existing timbral spatialization algorithms in section 4.1. The main focus on this new module lies in its adaptability and extensibility, giving room for experimentation with user built synthesis and decorrelation techniques. The second method in section 4.2, uses a known but rather unexploited

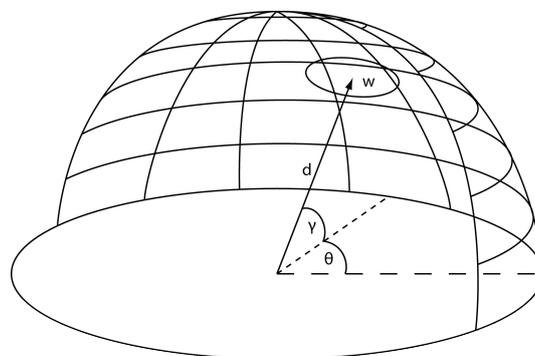


Figure 1.2: Illustrating the parameters available for each source in the VBAP/PBAP system.

technique, panning an audio signal in several simultaneous circles in kHz speeds. This creates an odd but spatialized synthesis technique similar to that of an amplitude modulator. Probably the most important insight and point for future work is the identification of gestural synthesis. By interlocking these different panning frequencies and the source sound frequency, one can achieve specific fluctuations in the overall enveloping sound that are grouped together by our senses, seemingly forming a trajectory. This synthesis method in particular is exclusive for spatial music and could serve as a tool to progress its development.

”Seit der Mitte des Jahrhunderts findet dann allmählich eine Emanzipation der räumlichen Bewegung statt, indem also auch der Punkt, an dem in einem gegebenen Raum, im Freien oder im Saal, ein Ton erklingt, und die Richtung, aus der im ihn höre, genauso maßgebend sein könnten, wie (in der übertragenen Vertikale der Tonhöhen) Töne verschieden hoch klingen können. Das ist ganz neu und wird in größerem historischem Abstand einmal als eine Revolution bezeichnet werden, vergleichbar der revolutionierenden Emanzipation der Dynamik oder der Klangfarben. Richtung und Geschwindigkeit eines Klanges in einem gegebenen Raum könnten also genauso relevant werden, wie die Frequenz eines Tones.”¹¹ (Stockhausen 1964)

¹¹”Since the middle of the century, the spatial movement then gradually reaches a level of emancipation, to which a spot where a tone sounds in a given room, outside or in a hall, and the direction from which I hear it, can be just as decisive as (in the figurative vertical of pitches) tones can sound at different pitches. This is completely new and will, someday, be marked as a revolution, in the distant future, comparable with the revolutionary emancipation of dynamics or timbres. Direction and speed of a sound in a given space can just as well be as relevant as the frequency of a tone.”

Chapter 2

Historical Backgrounds

As of the 20th century, the spatialization of music has received much focus since the dawn of the modernist period, especially with technological advances in sound reproduction techniques and electro-acoustic music on the rise since around the 1950's (Harley 1994, Zelli 2001). But the notion of space in musical composition goes back farther than one might suspect at first. Acoustic ensembles always demonstrate an inherent notion of space as two separate instruments simply can not physically sound from one and the same location (Solomon 2007). Certainly, one might argue that with a concert hall big enough for reverberation to accumulate all sounds into one mass, exact aural localization of each instrument can be very difficult to accomplish. On the other hand, this enveloping cloud of reverberated sound is also just another spatialized aspect in music. For example, especially in today's electroacoustic music culture immersiveness is a highly desired aspect (Adair et al. 2008, Stefani & Lauke 2010, Peters 2010, Lynch & Sazdov 2011b).

Initially, acoustic spatial composition might be thought of by discrete points, spatially, dramatically separated instrument (groups) and trajectories are quickly thought of by having musicians move as they play, if physically possible. But spatial composition in acoustic ensembles can also be accomplished through careful blending between similar, stationary instrumental groupings. That is, by merely, compositionally spreading melodies or timbres across the ensemble, the impression of a sound moving in space may arise without physically moving any performer or instrument. A popular technique in pre-baroque music is the use of stereo choirs, most prominently in the music composed for St. Marks Basilica in Venice (Stockhausen 1959). These virtual panning techniques are comparable with techniques employed in today's electronic music. As Solomon (2007) shows, spatial panning in acoustic music can be done in ensembles as small as string quartets. Of course, he also states that the reception of spatial cues work better the more distant performers are spaced out from one another.

In the following sections, this chapter, as the title suggests, will deal with respectively different aspects of the history of space in music. It will display synthesis of the literature research done by the author followed by a discussion of the like. Even though the practical part of this thesis deals with mere acousmatic composition – a term coined by Pierre Schaeffer

(Demers 2010)¹ – the author will also devote a section to the developments of musical space within acoustic ensembles. Not only is this important for a complete picture of usage of space in music from its beginnings, the author believes that much can be learned from the early usage of space, the limitations of certain approaches and the unique elements of both acoustic and electronic music, considering the spatial parameter. Many great thoughts on the usage of space come from great names that composed for acoustic ensembles, either for technological reasons (the technology or electronics itself simply did not yet exist in their time), or they engaged in both acoustic and electroacoustic composition, as it is more prominently the case with Karlheinz Stockhausen. Therefore, and for the sake of a linear progression through time, the following section 2.1 will deal with exactly the early approaches to space in music and the usage of space in contemporary acoustic and electroacoustic ensembles. Then, after a brief discussion of the advances in multi channel speaker technologies in 2.2, the focus will be turned to the acousmatic approach to spatiality in music in section 2.3.

Some examples might be unconsciously spatial, but their spatial features are undeniable. The creative minds might not have thought of the usage of space in a musical manner, but their concerns with the physical space did affect the composition to a degree. What remains is the discussion to which degree this can be seen as a musical parameter or simply just an effect. As I will discuss later (see chapter 3) my own view of physical space as an equally important musical parameter is similar to the ideas of Normandeau (2009), where space can only then be considered a musical concept, if, when left away, the piece loses its intentions. In all other cases, I shall, for my part, neglect the attempts as mere effects to reinforce other artistic intentions, which would still be present without the spatialization, with the exception of perhaps being weakened. I differ here from Normandeau, as he uses the word 'interest' instead of 'intentions'. Although he uses the word as vague and subjective as such, I do believe we mean similar things if *all* interest is erased with the disappearance of spatial cues. Nevertheless, strictly speaking, an effect, if applied correctly, is usually used to increase attractiveness, without which a musical work might fail to attract the listeners interest. Such might be the case especially with earlier attempts to be discussed below.

2.1 From Antiphony to Theater: Space in Acoustic Ensembles

If left wide open for interpretation, space seems to have always been part of mankind's musical activities. Naturally, through their dance, tribal musicians practiced spatialization in their dance rituals.² By bringing movement into their music, the sound source follows the trajectory that the body moves along. Pointillistic spatio-musical gestures were – and still are – practiced in the

¹Acousmatic music describes a purely electronic style of music, which separates what is heard from its source (Chion 1994). This includes even the speaker from which the sound is played. In his effort to define a new style of music, Pierre Schaeffer tried to detach the sound completely from any connotation or resemblance and make the listener concentrate only on the sound itself. If successful, the listener would be in a listening mode Schaeffer called *reduced listening*. See chapter 3.2 where his ideas are elaborated on in more detail and worked into the concept of spatial music.

²See the *Nea-Rim Gut* circle of musical work for Wave Field Synthesis by Ji Youn Kang, for example, representing Korean shamanic rituals (<http://www.kingeve.com>)

Christian tradition where church bells would be heard resonating through the streets or even from several different directions in a larger city, creating a rhythmic interlocking, articulated in space (Beyst 2003b) – but also the Muslim culture has their prayers sung simultaneously from minarets spotted throughout space. Other examples include those, in which music is used as a means of communication over long distances, as it is the case with talking drums (Zvonar 2006). In an other example, a listener standing in the middle of an alpenhorn conversation would not just listen to the different positions of the performers but also be engulfed by the reverberation of the complete mountain range.

2.1.1 Polychorality in Antiphonic Music

The principle of communication leads to the first prime form of the usage of space in music, as Zvonar (2006) identifies it: "Antiphonal performance ("call-and-response") is extremely ancient, having been practiced in the chanting of psalms by Middle Eastern Jews in biblical times, and there is further evidence of the practice in the Roman Catholic church as early as the fourth century." Antiphonal psalms would usually be sung by a choir split in half on either side of the altar, alternating between each half. The *antiphon* would then refer to a refrain-like intermissions between the verses. Taking the analysis methodology of Solomon (2007) into account, antiphonal performances are more than simply spatially separated performers. Among other factors, the spatial separation helps emphasize ball bouncing of harmony and melody between the caller and the responder, enriching the musical experience. Although, as Solomon (2007) himself mentions, call-and-response does not necessarily introduce spatialization per sé.

Early, conscious separation of performers is believed to already have been practiced in Ancient Greece, where multi-chorus performances took place consisting of separate groups of actors on stage (Yildizeli 2010). To amplify the voices on stage, resonating urns were placed around the stage and it is believed that the placement was done according to "coherent spatial" plans based on their respective resonant frequency (Miller 2009)³. That said, it is relatively well known that philosophers and mathematicians in Ancient Greece believed in a strong tie between music and geometry, discovering relationships between musical harmony and proportions in architecture (de Nooijer 2009). Although debatable, but, if true, Ancient Greek culture would have produced one of the first site-specific musical works. Being written for a specific architectural structure, the music *becomes* (just like the architecture itself *is*) a spatial piece of work. In fact, site-specific composition, the composition *for* specific spaces, is one of the prime inspirations for spatialized composition.

Such is the case with the Basilica San Marco in Venice, around the 16th century, representing a milestone in antiphonal psalmody and a major influence in the popularization of the polychoral style. What made composers be inspired to compose spatially for this particular structure was its unique design containing two facing church organs on either side of the cathedral (Zvonar 2006). Additionally, the cathedral featured the two spatially separated choir lofts

³This belief is based on estimates done of the holes and niches that held the urns, still to be found in remaining ruins. Similar urn techniques can also be found and compared "[...] in Medieval churches from the British Isles to the far east of Russia." (Miller 2009) See quotations given in Miller (2009) for further reading.

(Bates 2009), as the *cori battente* or *cori spezzati* ("broken" or "split" choir) were a standard in antiphonal music at the time already. Flemish composer Adrian Willaert, who became the cathedral's musical director in 1527, is usually credited with being the first to dramatically make use of the spatial separation by scoring echo effects in his music. This was probably done for the first time in his composition *Vespers* (1550). In fact, this is probably the first example in history in which space itself is represented in the score and effectively part of the music. The real drama and enhancement, though, was done by creating stronger ties within- and scoring a *tutti* across the whole choir to mark the *finale* of the piece. This technique was surprisingly new and broke off from the mere jumping back and forth between the choir halves (Solomon 2007), as was common practice before Willaert in antiphonal psalmody. Instead of spatially *and* musically separating the ensemble, Willaert brought the choir back together at times for an impressively enveloping effect that would fill out the complete space.

After Willaert, Giovanni Gabrieli⁴ would go out to extend and sophisticate Willaert's ideas and usage of the *cori spezzati* by also integrating instrumental choirs in the spatial arrangement. For example, Gabrieli's most acclaimed work *In ecclesiis* (1615), as part of his *Symphoniae Sacre* (1597) – called for two vocal choirs, one instrumental choir and the cathedrals organs, and most famously features the echo effects already mentioned in Willaert's work. The spatial arrangement and stereophonic degree of the individual choir parts is not fully clear, as different statements are made between Scholars. In Solomon (2007), and the references given therein, it is assumed that the choir parts were often separated by register, creating spatial locations for either high and low voices around the altar. Bryant (1981), on the other hand, talks of complete choirs being opposed against each other. Not only were they facing each other on the horizontal plane, but they were also separated *in elevation*, as the main choir would be situated close to the altar while other vocal and instrumental choirs were placed away and up in the organ lofts. What can be said, though, is that due to the disagreements encountered, exact spatial locations were probably of no concern to the composers. Exact spatial arrangements were usually not indicated in the score (Harley 1994). It is merely sufficient to separate the individual groups spatially. This also leads to a point of critique as it brings up the question as to how far the space was used truly musical as opposed to merely as an effect.

Spatial separation in antiphonal works became the trademark style of Venice at the time, but did not remain exclusive to the north-east of Italy. By the time of Gabrieli's death in 1612, the polychoral style had gained popularity throughout Europe. Gabrieli's pupil Heinrich Schütz went out to spread the technique in Germany and composer Thomas Tallis pushed the number of spatialized choirs to eight in his composition *Spem in Alium* in England in 1573. Miller (2009) analyzed Tallis' *Spem in Alium* with the same methodology developed for the spatial analysis of Stockhausen's work and concluded remarkable spatial movement if one were to arrange the eight choirs in a circle. Orazio Benevoli, the musical director at the vatican at the time, probably climaxed antiphonal polychorality with his 48-voice mass in 1650 for twelve choirs, two organs and basso continuo, with a whirlwind of sound flying over the heads of the

⁴Giovanni Gabrieli and Willaert are only indirectly linked through Gabrieli's uncle and teacher Andrea Gabrieli, who was in turn a student of Willaert.

congregation in alteration between voices and instruments, high and low registers (Pierce 1959). Antiphony and the polychoral style continued to be popular in sacred choral music throughout the Renaissance into the Baroque period, at which point it started to fade. The style had aged and was rejected by avant-gardists at the time, to which effect the compositional focus steered towards homophony and unified ensembles (Zvonar 2006). To celebrate the past idols, a few scarce examples still make use of the technique, for example Carissimi in his oratorios and Bach in his motets and, most famously, in the *St. Matthews Passion* (1729) (Solomon 2007).

2.1.2 Space as a Means of Dramaturgy

During the Classical period, composers showed little to no interest in any spatial antiphony or general thought on musical space, as composers gave no consideration of its usage in their scores and were rather concerned with the overall homophonic blend of the ensemble. In rare cases where the orchestra seating plans were of considerable importance to allow for antiphonic effects to happen (Miller 2009). Classical composers did, however, write musical works that synthesized with theater, most prominently in large scale works, such as operas (*Don Giovanni* by Mozart, for example). Musical theater demonstrates spatialization in music as the actors in different points on stage sing and talk from different locations and may even cause the heads of those in the audience to turn. Conversely, composer A. N. Scriabin stated that spatial movement in music was like dance in sound, as his vision of the future of music had the individual players of an orchestra repositioning themselves within the performance space (Yildizeli 2010). Therefore, the usage of space in his case, voluntarily or not, introduced a theatrical element as the performers were directed like actors. Such stage action would focus the attention to the actual, visual movement, pushing the auditory localization into the background of attention (Austin & Smalley 2000) (see chapter 3.1 for perceptual details).

In symphonic music of the time, one can find few examples in which instrumentation was connected to specific spatial locations. A technique to become popular in late-Romantic and early-Modern times was the use of offstage ensembles which would serve as dramaturgic effects. One of the first to make use of this was classical composer Christoph Willibald von Glück in his *Orfée et Euridice* (1762). Von Glück calls for a second orchestra to be placed "derrière le théâtre" and uses it to separate the underworld, the main action, represented by the main orchestra, with *Orfée* distantly calling from earth through the offstage orchestra (Solomon 2007). The work also represents a transition to what will become a Romantic and Modern use of orchestrated space as it makes use of spatialized call-and-response while the distance effect's main goals are of programmatic nature.

An exception is Wolfgang Amadeu Mozart's *Serenada Notturna* (1776) for two small orchestras and *Notturmo* (1777)⁵ for four Orchestras, in which he demonstrates a tight interweaving of physical space with the music through motivic segmentation and dynamic interplay. It seems that his intents are very musical instead of programmatically driven. Three orchestras were to

⁵The reader may be advised that Mozart's quadrophonic orchestra piece may sometimes be (strictly speaking, incorrectly) labeled *Serenade*, such as it is the case in Solomon (2007) and Bates (2009).

be positioned around the audience, while the fourth was placed offstage. More astonishingly, he creates echo effects by not simply repeating phrases with each respective orchestra delayed in time, but considers dynamics, masking effects and gradually adds mutes to more instruments in each repetition to denote a gradual darkening at each reflection (Solomon 2007). The use of echos is not just an interpretation done by scholars but is actually notated in the score by the composer himself at each respective position. Commenting on this approach, Stockhausen (1959) seems to dismiss this fact and waves the effort aside by labeling it a mere amplification of the already established form principal of classical 'echo'-repetitions.

In the core chapter of his thesis, Solomon (2007) spatially analyzes Ludwig von Beethoven's *Quartet in C# Minor*. Spatial movement that sweeps across the quartet from left to right or circulates through the ensemble as well any spatial interplay between individual performers is identified. The analysis in the case of Beethoven seems to mainly be tied to melodic similarities that are repeated between respective members of the quartet. In this, Solomon identifies motivic copies, mostly modified at each "echo" coming from another location. It is Solomon's main intent to show that spatial gestures have always existed and provides a newly developed analysis methodology based on contour theory and analysis⁶. He defends his approach with quotes and references supporting the statement that gestures of any kind may emerge even without the composers intent. But, since Beethoven never specified a preferred seating plan, the spatial interpretations quickly become irrelevant if some performers might choose to sit in a different position than assumed when analyzed⁷. Solomon's approach is more interested in the perception of spatial gestures than their actual composition. Considering space as a compositional parameter, it can be said that Beethoven most likely never considered this aspect in his compositional process and the gestures are not an effect but arise by mere accident.

Increased musical interest in conscious spatiality does not reoccur until the mid-Romantic times, in which theatrical motivations led to spatial considerations. This inspiration hit Hector Berlioz, who wrote his *Requiem* in 1837 which features the often cited *Tuba Mirum* section. Serving as a purely programmatic purpose, the four trumpet ensembles surrounding the audience represent the apocalyptic trumpets of the Last Judgement as they successively join each other from all points on the compass (Trochimczyk 2001). While his music is intentionally spatial, it always has programmatic intentions, as, for example, in his *Symphonie Fantastique* (1830), where he creates the illusion of distance using offstage ensembles representing shepherds answering each other across the valley (Zainea n.d.), or his return to antiphonal spatial textures in *Te Deum* (1849) where the opposing organ and orchestra should "function as Pope and Emperor, speaking from across opposite ends of the nave."⁸

Gustav Mahler was influenced by Berlioz and often made use of offstage performers in his

⁶His ideas will be investigated in more detail in chapter ?? as Gestalt theory and auditory streams form a core concept on the perception of spatial gestures and constructs.

⁷Standardized seating plans for ensembles as we know them were not established until the 20th century. Solomon is fully aware of this fact and is more concerned with the notion that spatial gestures *do* exist in every acoustic multi-performer music instead of giving a definite analysis of this specifically analyzed work

⁸David Cairns, linear notes to *Berlioz: Te Deum*, London Symphony Orchestra & Chorus, Colin Davis, Conductor – as quoted in Solomon (2007)

own works (Bates 2009) to create musical theater. Most famously, the *finale* in his *Symphony No. 2* (1893 – 1894) includes a dramatic offstage ensemble of trumpets, horns and percussion. Although explicit seating plans were not customary among composers, he does advise the conductor to have the trumpets sounding from four different directions, most likely to create echo effects of approaching and receding, juxtaposed against the onstage flutes. All instruments that are part of the spatial interplay have symbolic meaning. According to the program, Mahler compares the surrounding trumpets to those otherworldly of the apocalypse while the flutes represent the singing of birds on earth.

As a contemporary of Mahler and coming from a quite different musical tradition, American composer Charles Ives had a similar take on spatialization (Harley 1994) (Bates 2009), but reaches a greater level of sophistication on this field. His general compositional technique and great (posthumously acclaimed) contribution to music, is the use of collage and juxtaposition of musical parameters, ranging from melody to tempo. Harley (1994) describes his music as a complex mix of layers and his music is an effort to represent real-life experiences of street music scenes, parades and other simultaneous events. For him, spatial separation, especially offstage placement, was a means to detangle these interwoven structures and push layers into the background. Similar to the thoughts on ambient music by Erik Satie, he saw music as being something "[...]too often all foreground,[...]", but his solution was to create a complex environment of both foreground and background music, instead of only the latter, and let the audience choose what to tap into and orient themselves on (Ives 1933). This, as Ives states, required a much more active part on the listener himself. More importantly, he recognizes that spatial separation and offstage placement "are possibilities [...] not only from the standpoint of clarifying the harmonic, rhythmic, thematic material, etc. but of bringing the inner contents to a deeper realization." (Ives 1933). Although, in the opinion of the author of this thesis, this deeper realization did not happen until sound dispersion technologies and electronic music reached the level of sophistication it has today.

On a side note, an extravagant composition by the Russian composer Arsenij Avraamov named *Simfoniya Gudkov* (1922), also called "Symphony of the Factory Sirens" should be mentioned. It was written for the fifth anniversary of the Soviet Republic and performed in and with the Caspian port of Baku, shortly after it was invaded half a year before. Written for a large number of choirs and bands, the work also made use of rather unusual instruments, many of which were military equipment, including the foghorns of the entire Soviet flotilla in the Caspian Sea, navy ship sirens and whistles, artillery guns, machine guns, cannons, hydro-airplanes, bus and car horns and, of course, factory sirens (Callanan 2010). As one can imagine, the instruments were all scattered throughout the port with several conductors on specially built towers controlling the spectacle with colored flags and gun shots. "A central *steam-whistle machine* pounded out *The Internationale* and *La Marseillaise* as noisy *autotransports* (half-tracks) raced across Baku for a gigantic sound finale in the festival square." (Callanan 2010) The Symphony was performed once more in Moscow a year later, but failed to live up to its expectations in this different space.

Having composed more than half his catalogue of work with deliberate spatial intentions,

Henry Brant took Ives' ideas to the extreme and was one of the first to base his compositional methodology around the musical potential of space. Being strongly influenced by Ives, especially having conducted *The Unanswered Question* (1906) in his early years, his main concern was the clarification of dense textures through spatial separation (Brant 1955). His general approach and use of spatialization was to cluster instruments timbrally in designated places in space. Such is the case with his first spatial composition *Antiphony I* (1953), which splits the orchestra into five instrumental groups (strings, percussion, woodwinds, horns and brass), which should "on no account [...] be placed together on the stage, or near the stage! This would go directly counter to the specific spatial-polyphonic concept of the music."⁹ His insights were acquired mainly through trial and error, experimenting with positions and movement of musicians (Harley 1997). He differentiates on the use of distance, distinguishing horizontal distance, dependent on room size, from vertical distance, which he relates to higher pitches as the distance increases. As we will see later in section 3.1, humans tend to associate higher pitched sounds as coming from above and lower ones from below, as one would naturally differentiate a bird singing from the rumble of an earthquake.

Other effects he investigated in his pieces include what he calls 'travel and filling-up', a technique used in *Millenium II* (1954). As indicated in a seating plan for the performers contained in the score, ten trumpets and ten trombones are setup facing each other on either side of the audience, so that both the trumpets and the trombones form a line along the side walls of the performance hall. Brant would compose trajectories that would travel down one line of instruments and sometimes go around the audience continuing back up the other instrument group's path and onto the stage. This would contrast with gradually introducing each instrument, slowly filling up the hall with sound "[...]principally along the walls, but also with some feeling of the center being progressively saturated, [...]"¹⁰. *Millenium II* is also a good example of why he associates each point in space with one exclusive timbral group and vice versa. Brant tries to overcome another effect he titles *spill*, which is a confusion in spatial localization when two similar timbres play from different points in space. Basically, this is the principle amplitude panning between speakers one usually relies on. Whenever *spill* happens, Brant sees his vision of musical space confused and going against the clarity he strives for and for which space is his real tool.

The end of the 1950's sees the first scored examples of physical movement by performers. Predating Luciano Berio's *Circles* (1960), Brant scored *Hieroglyphics I* in 1957 and later *Windjammer* (1969), which contains specific walking routes. Although, Ives' father, having a great musical impact on his son, is said to have already experimented with directing two marching bands moving through the town square from different directions (Zvonar 2006). Similarly, Brant's *Bran(d)t aan de Amstel* (1984) makes use of the whole city of Amsterdam, having four boats with 25 flutists and one percussionist each floating through the canals, mixing with other ensembles dispersed around town (Harley 1997). Part site-specific installation, part concert piece, Brant's *Bran(d)t aan de Amstel* and other examples in his catalogue mark a blurry tran-

⁹Quote taken from the introduction to the score of *Antiphony I*.

¹⁰Brant, as quoted in Harley (1997)

sition from the dramaturgy of musical theater to spatial composition. While the placement and movement of specific performers in specific places denotes a theatrical spectacle in itself, most of his music seems to resist being reduced to a version excluding the spatial sensations. These pieces do not just become more complex as dissonant textures agglomerate into one point (as it would be in the case of Ives), one actually removes a fundamental, almost personal part of the composition, most likely removing its core intention.

As already mentioned above, Luciano Berio thought of space musically in his composition *Circles*, but even predated Stockhausen with his spatial considerations in *Allelujah I* (1955). For *Allelujah I* he placed the five ensembles in different spots on stage, but later decided that the spatial effect was not perceivable and revised his work with *Allelujah II* (1957-58), placing the ensembles around the performance hall instead (Harley 1994). His composition *Circles* for female voice, harp and two percussionists experimented with distance effects, having the singer move backwards into the ensemble. In his later works, such as *Coro* (1975-76) for 40 voices and instruments, he takes a similar approach to Brant, segmenting space into timbral groups, or ordered by register.

Additionally, several electronic musicians in North America developed pieces influenced by Ives' and Brant's approach: John Cage's *Williams Mix* (1952), Earle Brown's *Octet* (1952) and Morton Feldman's *Intersection* (1953), are some examples. All pieces are done within the scope of the same project, and are composed of eight aleatoric, unsynchronized tapes positioned equidistantly around the audience (Zvonar 2006). As Cage once stated in a lecture in 1957: "[...]this new music, whether for tape or for instruments, is more clearly heard when the several loudspeakers or performers are separated in space rather than grouped closely together. However, this music is not concerned with harmoniousness as is generally understood, where the quality of harmony results from a blending of several elements. "Here we are concerned with the coexistence of dissimilars, and the central points where fusion occurs are many: the ears of the listeners wherever they are."¹¹ As Bates (2009) states, while this music, just like that of Brant before, holds spatialization as a central concept, it does not acknowledge its full potential. It is always used in a way to organize elements conceptually, but never to use it musically.

2.1.3 Symbolical Spaces

For Raymond Murray Schafer musical space is somewhat of a symbolic tool. For one, he relates soundscapes to social contexts and how the perception of soundscapes alters the state of consciousness of the listener. In his essay, (2007) he speaks of religious soundscapes and how they affected the close bond between the religious concept and its followers. As an example, he writes about how in Christendom the sound of bells in every town marks the area of the church's reach. You were either inside a commune or in the wild. Eventually, Christianity built bigger bells to increase God's reach, just like Islam adopted the loudspeaker to further deliver their prayers (Schafer 2007). Schafer is most known for his enormous cycle of works *Patria*¹²,

¹¹Cage, as quoted in Bates (2009)

¹²As of May 2006, the most recent source found, *Patria* is still uncompleted and not much critical analysis could be collected

which consists of 12 musical theatric *Gesamtkunstwerke* that are closely linked with the location of their setting (Smith & Smith 2006). In the prologue of *The Princess of the Stars* (1981) he requires the piece to be set on a lake, while *Patria 5: The Crown of Ariadne* must be set on a beach large enough to host a labyrinth that should be burnt down at the end of the show surrounded by a huge dance performance. Besides his approach to spatial composition as a ritual, a soundscape for the "radical change in the [theatre customer's] existential status"¹³, his immediate spatial compositions seem to involve rather conventional techniques, like "offstage" performers in the unseen soprano princess in his prologue or echo effects across the large space he requires.

Perhaps the most direct symbolic use of space is the actual creation of symbols, e.g. the mystical star or cross in *Apocalypse Part Two: Credo* (1977), similar to how stars in the night sky form zodiac signs. The piece is full of other symbolic connotations as the twelve spatialized choirs used allude to the twelve gates of New Jerusalem in St. John's Revelation (Harley 1994). Next to his affinity to religion, his preoccupation with his homeland Canada is not only present in *Patria* (Latin for homeland) but is further exemplified in *North/White* (1979), a form of protest of the continually growing man-made noise pollution of the Canadian North, symbolized through distance and offstage performers. *Music for Wilderness Lake* takes an extreme position, ignoring any audience and requiring the twelve trombonists to move to a remote lake. Schafer writes: "the location, the climate and time of day are as essential as the musical notes. [...] *Music for Wilderness Lake* return to a more remote era [...] when music took its bearing from the natural environment [...] This interplay requires a spiritual attitude [...]"¹⁴

Other symbolic allusions can be found in Charles Hoag's *Trombonehenge* (1980) placing thirty trombones based on Stonehenge around the audience. Brian Ferneyhough's *Transit* (1975) disperses the six solo voices and chamber orchestra in four semicircles, each representing the four spheres of the Universe, based on Renaissance belief. But probably one apogee of spatial (religious) symbolism is John Tavener's *Ultimos Ritmos* (1972). Dedicated to St. John of the Cross, it is full of religious connotations and rites in all aspects of the music. Considering spatialization, geometrically Tavener combines the cross with the circle, seating the instruments in the general cross shaped architectural design in churches and around the altar. Furthermore, he places instruments on elevated positions, in balconies, with which Tavener creates a vertical movement symbolizing the "the descent of the Eucharist." (Solomon 2007) Concluding, one can say that any seating plan for acoustic ensembles forms some kind of shape. As Harley (1994) remarks: "[...] floor plans and performer placement diagrams are integral, though inaudible, elements of the musical structure." This might not be accurate, as the inaudibility depends on the placement in relation to the audience, and, even more so, the composer of contemporary acoustic – as well as electronic music – should possibly be aware of the allusion he might unintentionally create.

¹³Schafer, as quoted in Harley (1994)

¹⁴Schafer, as quoted in Harley (1994)

2.2 Technological Developments in Multichannel Speaker Arrays

”Representations within a medium are always media-specific, and in the best case, the medium becomes intuitive. In that case, the people creating the multichannel tape, the interactive soundscape, the sound-rendering and its spatial parameter curves become unheard [...] If that is achieved, we have reached our goal.” (Goßmann & Dombois 2003)

Before we continue with the historical investigations of artistic use of space in a musical context – especially before touching upon the domain of electronic music – we may first need to have a general idea about the technological developments that shaped this musical thinking. As stated in the quote by Goßmann & Dombois (2003) above, the technology used to reproduce the composition will affect its sound and perception – and even possibly change a piece in its entirety. In contrast to most artistic uses of technology, making use of and misusing technology to create fantastic sound worlds or interesting artifacts, the audio technicians are mostly concerned with the precise reproduction of a sound or sound field. With this in mind, it is interesting that audio reproduction technology, from its beginnings, always displayed the importance of the spatial impression for a realistic impression, be it binaural, stereo or today’s sophisticated technologies, such as *vector based amplitude panning* (VBAP), *Ambisonics* and *Wave Field Synthesis* (WFS). But, no matter how many dimensions in which quality they exhibit, they all come with their specific flavor. In that, the spatial composer should know with which system he is dealing with. While this notion that technology (and room acoustics alike) is a prime factor for the coloring of the musical piece, it is especially prevalent with practitioners of sound diffusion. Nevertheless, every composer of electronic music is similarly affected by these unavoidable impacts on their music, and, hence, should never forget that the medium is your instrument: ”The choice of which technology to use to achieve the compositional concept is based on several factors, such as the desired accuracy of the spatial image, as well as the performance context.” (Baalman 2010)

How advances in technology impacted the notion of space and a drive for accurate spatial reproduction was already apparent in the World Fair in Paris in 1881. Clément Ader displayed a new audio transmitting technology bringing the performance of the Paris Opera House to the Fair using two telephone receivers, binaurally, one for each ear (Zelli 2001). Multichannel microphone technologies and stereo reproduction standards started to be standardized after Alan Blumlein defined his coincident microphone technique in a patent from 1931. In 1940 Walt Disney developed one of the first discrete multichannel surround multimedia technologies, called *Fantasound* (Peters 2010). The composer Leopold Stokowski had worked at Bell Labs and knew about Steinberg and Snow’s experiments to reproduce spatial cues (Robjohns 2001). He recorded the Philadelphia Orchestra using 33 microphones mixed down to 8 recording tapes. Developed for their new film *Fantasia* (1940), three main channels (left, center and right) delivered the principle stage action (Peters 2010), while the remaining effect and music channels

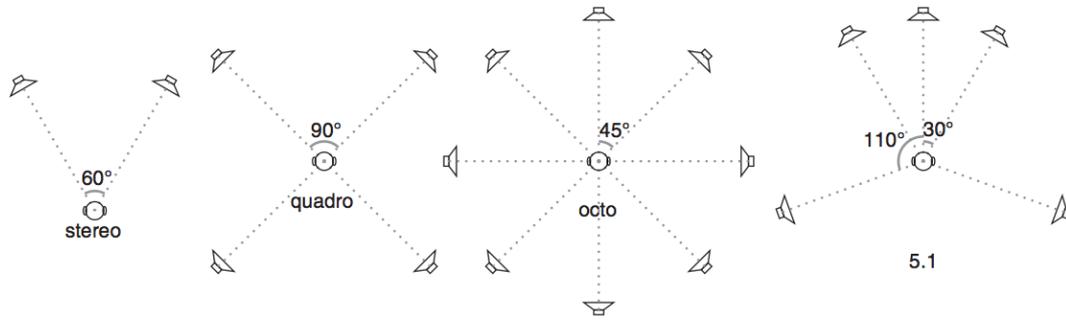


Figure 2.1: Some examples of stereophonic configurations (Baalman 2010).

were upmixed to a maximum of 65 speakers for an immersive listening experience. The diffusion was first realized live and was later automated using notches on the edge of the film (Garity & Hawkins 1998). Because no cinema could provide the technology needed, *Fantasound* premiered as a roadshow, carrying the diffusion technology to each venue. Unfortunately, World War 2 prevented the exposure of *Fantasia* to the European market and the show turned into a financial disaster. Rumors are, Disney planned on bringing *Fantasound* to Europe, but a German submarine sank the ship during the war, wiping the technology from the face of the earth (Robjohns 2001). Surround speaker technologies were not to be seen again until the end of the 1950's, when contemporary artists took creative interest in 3D audio.

Stereophony went out to become the de facto standard for the record industry, with a "stereo system" in every home. Today, the most common surround sound technologies are still based on the original experiments by Blumlein, Steinberg and Snow, with the 5.1 and 7.1 standards being well established. One should keep in mind that these standards were developed specifically for the cinema, accustoming to a large audience, which is mostly frontal oriented towards the screen. While for musical applications these standards are interesting because of the ease of distribution, other configurations, such as quadrophony and octophony are usually the weapon of choice in the performance hall (Baalman 2010), most likely due to their lack of preference to any one spatial direction (see figure 2.1). But, especially from an artistic point of view, one has to distinguish between stereophony and polyphony. Monophony and Polyphony represent the notion that one speaker is assigned to one sound. As Curtis Roads supposedly once put it: "If you want the sound to come from a specific place, put a loudspeaker there."¹⁵ This approach to spatialization is especially useful for installations and site-specific works, where the artist has to accustom to the locality, its architecture and its acoustics (Stefani & Lauke 2010). Many spaces do not allow standardized reproduction systems to be constructed, and custom solutions have to be sought out, enabling specific sounds coming from specific areas to enhance the acoustics and overall impression of the space. The difference with stereophony is, that it makes heavy use of *virtual sources*, meaning psychoacoustic illusions, having the sound appear between the speakers by means of amplitude panning or temporal differences.¹⁶ Stereophony

¹⁵As quoted in (Baalman 2010)

¹⁶Time panning techniques are very dependent on frequency, though, a behavior which amplitude panned signals do not exhibit that much (Pulkki 2001). Being considered more stable in that case, amplitude panning is widely popular and by far the most widely used panning method in music production.

battles with the problem of the sweet spot, a point within the speaker array in which the spatial impression (or illusion) is best heard. For the stereo effect to work, it needs carefully weighted amplitude differences and/or tiny time differences. It also assumes that the listener is fixed, as these differences change subjectively based on his movement and change in location. If the time difference is too big, or the level difference too drastic, the audio will collapse into one of the speakers. This problem arises due to an effect known as the precedence effect (see section 3.1.1). If an audience member is too close to one speaker, the time difference and especially the amplitude to the closer speaker is significantly increased, so that his spatial impression will suggest the audio to be coming from this one position only. For this, composers such as G. M. König, therefore, engaged in polyphonic composition, and refrained from using stereophony, as he feels dissatisfied with the weaknesses of the technology (Baalman 2010).

Spherical standards based on stereophony that defined the third dimension did not emerge until the later end of the 20th century. A relatively robust and wide-spread technology is the VBAP algorithm (Pulkki 2001). It employs stereophonic techniques triangulated between up to the three speakers along the surface area of the listening sphere and closest to the virtual source. The amplitude weighting is done by viewing the distances of the loudspeakers to the listener and to the virtual source as vectorial distances. The resulting gain factors for each speaker are then simply determined via a linear combination of these vectors. VBAP is usually applied in dome-like concert halls or studios and is reported to give a fairly constant spatial impression for large listening areas with at least eight speakers (Baalman 2010). Two drawbacks of pure VBAP are the small sweet spot, as it triangulates based on the assumption that the listener is exactly in the middle of the speaker array, and the absence of any distance considerations. RVBAP (*reverberated* VBAP) tries to fill this deficit by introducing a reverberation algorithm into the standard (Peters 2010). Another extension to VBAP is *distance based amplitude panning*

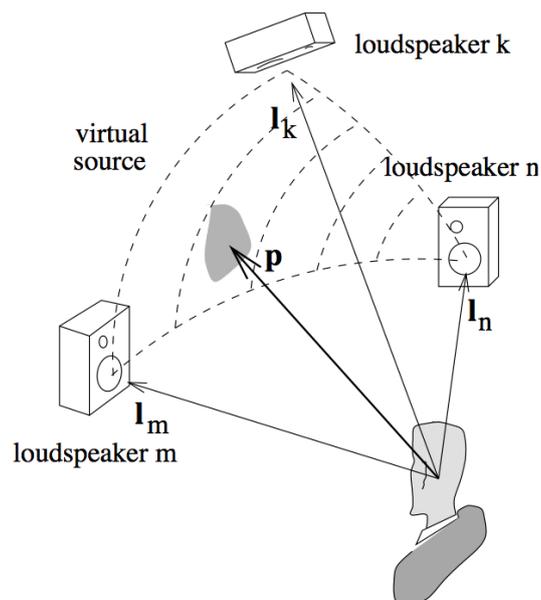


Figure 2.2: An illustration of the general VBAP algorithm (Pulkki 2001).

(DBAP), a technique avoiding any prior knowledge of the position of the listener, widening the listening sweet area by diffusing the source over all speakers (Lossius & Baltazar 2009). Adding this slight spatial blur, the method avoids the collapsing of the source into one speaker. Furthermore, it incorporates an amplitude reduction based on the distance of the virtual source to the nearest speaker.

Besides standardized technologies, there is an emerging multitude of "standardized" spaces for which a composer can compose for specifically.¹⁷ In terms of spatialization domes, early development was done by Belgian composer Leo Küpper in the 1980s, with spatialization halls in Rome, Avignon and the biggest one in Linz, made up of 104 speakers (Ramakrishnan 2009, Normandeau 2009). His intent was to separate the technology from the music and establish a set of standard spaces for spatial music to be performed. This philosophy was ported to contemporary dome-like systems, such as the *Klangdom* at the ZKM, Karlsruhe, Germany (Ramakrishnan 2009) and the dome at the Université de Montréal (Normandeau 2009). These type systems distinguish themselves by striving for a most transparent sound, using a uniform brand of loudspeakers in rooms that exhibit little coloration. Technology used with the *Klangdom*, for example, is mainly VBAP, with their own extension called *sound surface panning* (SSP), and with the possibility to use *Ambisonics*, and all incorporated into their in-house software named *Zirkonium* (Ramakrishnan 2009, Peters 2010).

The other trend in spatial dispersion systems are non-homogenous speaker arrays. Influenced by Pierre Henry's thoughts on stereo channel tape diffusion, François Bayle developed the *Acousmonium* in the beginning of the 1970 (Bates 2009); an *orchestra of speakers* consisting of around 60 speakers, usually grouped in pairs. It is stage oriented, similar to traditional acoustic concerts and accords to the knowledge that "[...] spatial hearing in the forward direction encourages the most regular and sophisticated spatial perception [...]" (Zelli 2010) (see chapter 3.1). "Bayle establishes the need for the Acousmonium on two grounds. First, he assumes that the dynamic character of the music in general requires a symmetrical distribution of technically identical speakers. Second, he believes that narrativity - a character-building factor in musique acousmatique - needs a spatially dispersed loudspeaker orchestra upon the stage which is capable of expressing and representing the inner structures of the music space." (Zelli 2010) Other systems that build on the grounds laid out by the *Acousmonium* are the *GMEBaphone* (Peters 2010) and the *Birmingham ElectroAcoustic Sound Theatre* (BEAST) (Wilson & Harrison 2010). The BEAST in particular is specifically designed for portability, "[...] adaptability and pragmatism." (Wilson & Harrison 2010) It employs speakers mainly in the front of the audience, but also around and above. Similar to the *Acousmonium*, the current playing speakers are chosen by the performer in real time for their characteristics – which also governs their placement, also keeping human psychoacoustics in mind: Tweeters, for example, are exclusively placed above the audience, since higher frequencies are associated with sounds coming from above (see chap-

¹⁷Although some of the homogenous system might seek to become a standard, the multitude of these systems itself, each with a different interface and technology, shows that each "standard" almost only exists for itself – and might not even be desired. The inhomogeneous systems further below, for example, dodge this problem by presenting a variable collection of different speaker types and a diffusion philosophy leaning towards live diffusion and adaptability.

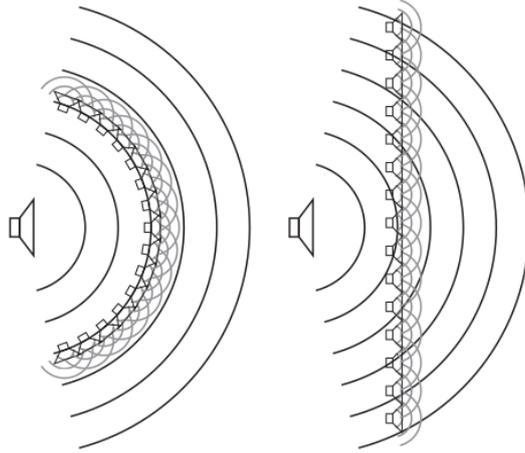


Figure 2.3: The original Huygen’s principle (left) and its general adaptation in WFS (right) in comparison (Baalman 2010).

ter 3.1). BEAST is a mobile system, touring all over the world, employing a variable amount out of over 100 speakers of different types, based on the composers intentions and the spatial demands.

Concerning the accessibility and availability of spatial music today, the concept of mobile sound dispersion systems and technology forms an alternative to the establishment of standardized spaces around the world. Instead of installing multiple systems in fixed places, these systems, like BEAST, can reach the listener instead of requiring the audience to travel to them. The Game of Life Foundation has established another of these mobile systems, but employing the *Wave Field Synthesis* (WFS) technology instead. Their 192 speakers are custom built to facilitate transportation and reconfiguration, being controlled by an in-house, extensible software written in Supercollider (Valiquet 2009). The large amount of speakers is a necessity and one major drawback of WFS. The underlying principle behind WFS is the *Huygens’ Principle*, which can be mentally derived as a wave traveling through a discrete set of equally spaced openings in a wall (or simply points) behind which the original wave itself is recreated by adding the fronts of each smaller wave emerging from these openings together (cf. figure 2.3). The advantages of the physical reconstruction in WFS are the large sweet area (as large as the speaker array itself), a sense motion parallax by keeping the source in the same virtual position from any listening point, the production of plane waves (used to simulate very far sources, for example), the a realistic sense of presence through a natural increase of amplitude upon approaching the virtual space and, most interestingly, the ability to comparatively stably project virtual sources *inside* the listening space (Theile 2004, Bates 2009, Baalman 2010) (see figure 2.4). Next to the high demand in hardware and processing power, already mentioned above, WFS exhibits aliasing effects above a certain frequency range, dependent on speaker size and spacing. The smaller and closer together the speakers are, the higher this aliasing threshold becomes, and, hence, the audio quality increases. The high demand in hardware has also prevented any three-dimensional WFS system to be built, so far. Furthermore, the WFS reconstruction is very sensitive to room reflections of the performance space, and especially the distance perception

can greatly suffer from strong, additional early reflections (Theile 2004, Bates 2009). Although WFS has not seen a large artistic exploitation, probably due to its high demands in hardware, it is very promising and appealing for composers (Bates 2009), and efforts, such as the Game of Life system, the fixed systems of the TU Berlin and the Fraunhofer Institute of Ilmenau, will hopefully encourage a larger availability and use of WFS in spatio-musical contexts. Currently, WFS does not provide solutions for recording sound fields and is only applicable to artificial soundscape synthesis.

A popular alternative to WFS is *Ambisonics*, also a sound field reconstruction method, but driven by psychoacoustic amplitude panning techniques, similar to stereophony (Pulkki 2001). Bates (2009), though, points to several points in which Ambisonics differ from stereophony: "Ambisonics is consistently preferred to amplitude panning for dynamically moving sources as it produces a more uniform phantom image and hence disguises the loudspeaker position. However, amplitude panning was also consistently preferred for static sources as this method uses fewer loudspeakers and so reduces the localization blur." For one, Ambisonics is fully three dimensional from its ground design by default.¹⁸ Furthermore, it separates the encoding (recording) stage from the decoding (playback), giving the possibility to reconstruct the recorded sound field over an arbitrary set of loudspeakers, usually allowing the most practical setup for the respective performance space. This is done by creating an intermediate stage called the *B-Format*, that encodes the three dimensional audio in spherical coordinates. First order Ambisonics, for example are encoded in three orthogonal figures of eight and an omnidirectional signal, and require a minimum of 8 loudspeakers to function properly. Higher order Ambisonics (HOA) then increase the accuracy of the sound field and the size of the sweet area, but also require a higher minimum number of loudspeakers spread around the room. The higher the order of Ambisonics, the more it matches WFS (Baalman 2010). Being considered mainly a microphoning technique (Pulkki 2001), Barrett (2004), for example, prefers Ambisonics to WFS due to its practicality of recording sound fields, though. In this, she seems to strive for

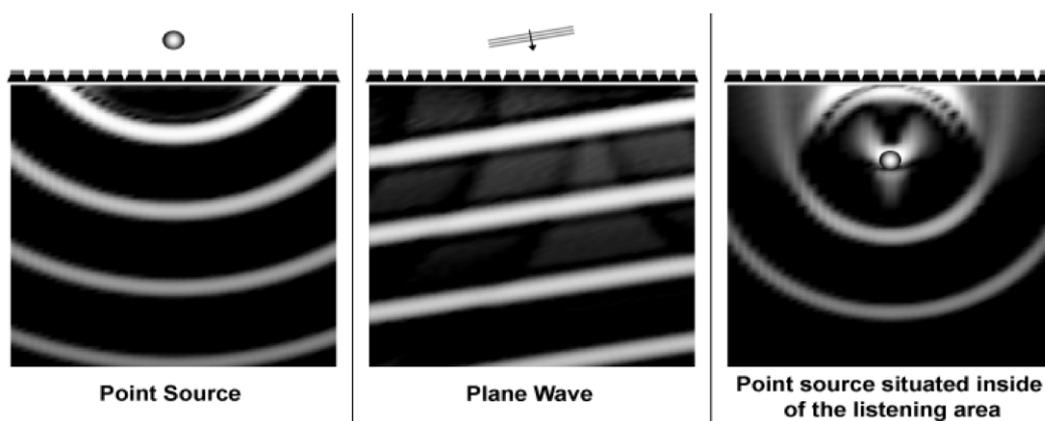


Figure 2.4: The three general source types in WFS, point source, plane source and a source inside the listening space (Bates 2009).

¹⁸Mathematically, this is a matter of generalization, but also technically speaking, Ambisonic microphones, such as the SoundField SPS200, naturally record and encode in three dimensions.

the generally desired feeling of being enveloped, a quality for which Ambisonics is renowned for. In her paper [2004](#), she also points to some artistic limitations of HOA and suggests a hybrid solution between first order and HOA. The problem lies in the high resolution of HOA, which assumes every source to be a point source. While source positioning and trajectories are more clear, source width and diffuse fields become a difficult task to overcome. This, furthermore and prominently shows that each spatialization system in its detail has its specific characteristics, which the composer needs to be aware of – hence, it shows that the technology clearly is the acousmatic composer’s instrument.

2.3 Spatial Parameters in the Electro-acoustic Tradition

In the author’s opinion, all examples in the previous section [2.1](#) that somehow include physical space in their music show remarkable and valuable efforts, but fail to treat the spatial dimension in music as an equal parameter to melody, harmony, rhythm, etc. Space was always an effect, a tool, an additive to enhance the musical experience – and considering the immense limitations composers face when composing with acoustic sound sources their way of dealing with space is more astonishing above anything else. But, it wasn’t until the invention of the loudspeaker and electronic reproduction technologies, such as phonographs and magnetic tapes, that made space continuous and, by perceptually detaching a sound from its source speaker, mass-less. Furthermore, the 20th century greatly challenged the concept of music detaching itself from pitched music and, instead, using everything from sines tones, over concrete sound to noise as the basic building blocks. Without the other traditional parameters, there was a need for another dimension in musical expression. These two facts brought forth a completely new understanding of the audible space and its compositional value.

Few composers in pre-electronic times have had a notion on how the spatial parameter could make a purely musical contribution. One of them is Erik Satie with his *musique d’ameublement*¹⁹. A description of the ambient sound installations of Erik Satie are given in [de la Motte-Haber \(1986\)](#). Being thought of as ambient interludes during breaks between acts, the music was to be played from all corners in the room and from certain spots on balconies. His intention was to fill the room with sound like the walls were covered with wallpaper ([Yildizeli 2010](#)). While the actual, musical output might have been similar to efforts described in section [2.1](#) – static, dispersed-pointillistic, depending on room reverberation and to which degree the performers were hidden from a direct line of sight – his thoughts as to how the space was to be composed makes a valuable contribution to what this thesis will try to elaborate on spatial composition in later chapters. Unfortunately for him, the audience did not understand his ambient intent, rushed back to their seats to listen... and a new, radical aesthetic was born.²⁰

¹⁹*Musique d’ameublement* translates into ”furniture music” and describes a musical work that is worth of no more nor no less importance than a chair, for example. It exists, in space, in the background, and is there for a purpose, mostly unnoticed. Considering Satie’s effort as an ambient sound installation, I hold his work in the context of this thesis higher than that of Brant, because sound installations (being mainly concerned with the space they exist *in*) are of lesser dramaturgy and primarily, almost abstractly, concered about space.

²⁰As the audience was coming back into the theater, Satie quickly got on stage and screamed:”Walk around,

The French *musique concrète* movement from 1948 is probably not only the first to define a completely electronic style composed of radically new musical material, its fathers Pierre Schaeffer and Pierre Henry quickly came to the notion of spatialization as an integral part of electronic music. With the absence of melody, harmony or meter, the music was in need of other parameters to add depth in expressivity. During their collaboration on *Symphonie pour un Homme Seul* (1950) they later, in 1951 together with Jaques Poullin, developed an interface with the sole function to spatialize source sounds (Zvonar 2006). They called it *potentiomètre d'espace* and it consisted of four induction coils operated by large arm movements of the performer, each driving one of the dispersed loudspeakers. In his early writings, Poullin describes two types of expression in electronic spatialization: *statiches Relief*, which describes the simultaneous projection of distinct sound from distinct loudspeakers, and *kinematisches Relief* referring to the trajectories sounds that move along during the performance of a *musique concrète* concert and for which the *potentiomètre d'espace* was used for (Harley 1994). In Schaeffer's vocabulary, a sound's movement through space was called *trajectoires sonores*. While being highly theatric, the *potentiomètre d'espace* was somewhat of an electronic instrument, as it made the spatialization process – isolated by itself – an expressive act. It played the space just like any pitched instrument would play a melody.

Vortex, by Jordan Belson and Henry Jacobs, was an audiovisual multimedia experience created at the Morrison Planetarium in San Francisco in 1957 and ran for 3 years during which thirteen programs were presented including the music of Karlheinz Stockhausen, Vladimir Ussachevsky, Toru Takemitsu and Luciano Berio. The audio system was able to make use of thirty-six to forty speakers and sound could be moved in circles and possibly trajectories with the help of a special rotary console (Zvonar 2006). The program was said to have become so popular, it was invited to the Brussels World Fair in 1958.

2.3.1 A Space of Moving Objects

Completely new viewpoints on music can only be given by completely radical minds, and one truly avant-garde head was that of Edgar Varèse. Paving the way for *musique concrète* after him, the "objectivation" of music (based on the ideas of his predecessor Igor Stravinsky (Harley 1994)) marks the foundation of a new musical organization. Defining music as "organized sound" is Varèse's great legacy to music today. Musical objects, in a way, replace the notion of a melodic motif: "[They] flow, change, expand and contract, yet they have a certain tangibility, a concreteness established by clearly defined boundaries." (Harley 1994) These objects are subject to modifications through motion, interpenetration, collision, etc, all spatial vocabulary, demonstrating the clear connection of musical objects to their spatial attribute. The boundaries are defined by the differences in timbre and dynamics, creating "zones" in which their tone color becomes an "integral part of form"²¹.

Spatial dispersion and sound object organization were already experimented on in his

talk to each other, but don't listen!", because there was nothing to listen to... (de la Motte-Haber 1986)

²¹Varèse (1936), as quoted in Harley (1994)

pieces *Integrals* for percussion instruments (Yildizeli 2010) and, later on, in his first electro-acoustic work *Déserts* (1954) for chamber orchestra and tape Austin & Smalley (2000). His true vision of musical space, though, was a continuous one, envisioning sound coming from any part of the performance hall and as many points as may be required by the score. The Brussels World Fair not only saw Belson and Jacobs' multi-speaker arrangement, but was the birthplace for a collaborative effort between Le Corbusier and Varèse, the highly influential multimedia extravaganza named *Poème Électronique*. It wasn't until 1958 in Brussels that Varèse could experiment with completely free movement of sound in space. For the first time, Varèse was to listen to the music "literally projected into space"²². Having had this experience, he would later call his earlier, instrumental efforts an aural illusion (Harley 1994). Unfortunately it was one of his last pieces and his only purely electronic work.

The Dutch electronics company Philips hired Le Corbusier for an all encompassing artistic work to showcase their innovations in the their world fair pavilion. The pavilion itself was designed by physicist, architect and composer Iannis Xenakis, a student of Le Corbusier's at the time. Being hailed a *Gesamtkunstwerk*, the *Poème Électronique* encompassed the architectural structure built for the media, which, in turn, was created for the structure. In other words, the work was site-specific with the site serving as a "delivery instrument for the multimedia content" (Lombardo et al. 2009). The whole complex was destroyed after the end of the fair. Nevertheless, Lombardo et al. (2009) tried to digitally recreate the experience in a virtual reality with limited accuracy but fair success. The architecture was a center-lacking²³, self-supporting "tent" (as it was Le Corbusier's original idea) with hyperbolic paraboloid walls standing on a contour depicting a stomach outline. Le Corbusier himself was mainly responsible for the visuals, which included a film supported by slides projected onto the walls, color themes and color spots, and two hanging figures, a female mannequin and a mathematical object.

The trajectories were realized using speaker arrays that were spotted all over the walls and arranged in predefined paths. The pathways were fitted to the structure, as sound was to travel up into one of the three peaks and along the parabolic lines of the structure. Just like the structure had no center, presumably the musical movement had no center of gravity as it used the building's walls for orientation. In total, an estimated 350 speakers were installed and controlled via a 52-step telephone selector driven by 20 12-channel amplifiers (Lombardo et al. 2009). The trajectories were grouped into arrays of five speakers, alternating one speaker at a time. The original sound was recorded to a 3 channel tape accompanied by a control score. The score was continuous with 5 second reference segments, describing the audio routing for each control channel. The whole exhibition was among the most popular of the exhibition and it was described as: "[O]ne no longer hears the sound[;] one finds oneself in the heart of the sound source. One does not listen to sound, one lives it"²⁴.

The *Poème Électronique* did not only contain the music of Edgar Varèse. During the

²²Varèse, as quoted in Lombardo et al. (2009)

²³...meaning no place inside will give an audience member the impression of being in the center of the building. This was intended, as to disperse the audience and prevent accumulations in one single spot (Lombardo et al. 2009).

²⁴As quoted in Lombardo et al. (2009).

transition phase, where audience members entered and left the building, a short mono tape piece by Iannis Xenakis, entitled *Interlude Sonore*, was dispersed among the speaker system²⁵. The piece would later be remixed for stereo playback and received more attention under the name *Concret PH*.

Iannis Xenakis first came to Paris as an architect working at the atelier of Le Corbusier. During work on *Poème Électronique* Xenakis had already experimented with his first symphonic work *Metastaseis* (1954). He would later be introduced to Olivier Messiaen from which point on his musical career would overshadow his architecture. But the influence through the collaboration with Varèse and his ideas on sound-masses and space in a musical context inspired Xenakis to constantly keep a notion of movement and direction in both his electronic and acoustic compositions (Harley 1994). Being an architect, Xenakis would think of music in similar ways as the buildings he designed. His fascination with mathematics, especially non-Euclidian geometry and stochastic models already were expressed in his architecture, but also later defined his musical style as a composer. Both the motivation to fuse mathematics with art and the close fusion of architecture and music came from his old mentor. As Xenakis stated: "According to Le Corbusier, architecture is linked to music by the concept of movement and the successive perception of volumes and spaces."²⁶ Furthermore, Xenakis would later state his growing concern for the aural experience of his audience (de Nooijer 2009), depicting another motivation for spatial composition.

The influence of Le Corbusier and Varèse is clearly visible in his series of multimedia installations named *Polytopes*. The name translates from Greek into 'multiple places' and each piece, as suggested, deals with space as a fundamental concept. By dealing with a clever interplay between architecture, light and sound, Xenakis tries to fuse space and time into one homogenous concept (Sterken 2001). The architectural space becomes transformative and temporal, while the "musicalization of space" is further developed in each *Polytope*. The immersion in each work would shift from total engulfment, as exemplified in the massive light show of the French pavilion at the Montreal Expo in 1967, to the simultaneity of the interior, virtual and exterior, real world of the *Diatope* pavilion at the Centre Pompidou in Paris in 1977. Through visual and aural bombardments, he shifts the perception of the site beyond the performance (Sterken 2001). Further on, in parallel to his work on the *Polytopes*, Xenakis was invited to create a work for the Japanese pavilion at the Osaka World Fair 1970. He was enticed with the chance the work with 800 loudspeakers grouped in 250 locations for the work *Hibiki-Hana-Ma* (1970) (Harley 2004). The source sounds for the tape piece consisted mainly of orchestral timbres mixed with traditional Japanese instruments. The wide range of sounds, governed by stochastic processes, were re-ordered using the spatialization and gave the individual, the moving listener, a personal "organized" subsection of the sonority.

Xenakis would later take the inspiration provided by the possibilities of concrete music

²⁵Contrary to other statements possibly made, current research by Lombardo et al. (2009) states that no one is fully sure how Xenakis' composition was really spatialized across the pavilion's system. His best guess is that *Interlude Sonore* was played close to the entrance and exit based on the few historical accounts in existence (Lombardo et al. 2005).

²⁶Xenakis, as quoted in de Nooijer (2009)

into the acoustic, instrumental domain, effectively shifting spatial composition from theatrical to pure musical means. While in *Eonta* (1963-64) he would simply order the 5 brass musicians to move around stage (Xenakis 1967), *Terretêktorh* (1965-66) would finally abandon the theatrical approach and create sonorous sound trajectories through dynamically balancing of instrumental timbres. Most astonishingly, he places the audience into the middle of the sound by dispersing the players among the public, encircling the conductor altogether. Taking his seating plan as a reference, Santana (2001) identifies several trajectories, especially spirals of different types, such as the logarithmic or hyperbolic spiral. The musicality and expressive intent of the spatialization becomes especially apparent when he reduces his approach to the same instruments and pitch per trajectory and only modifies the speed and dynamic at each rotation.

Similar to *Terretêktorh*, Xenakis mixes the audience with the performers in *Nomos Gamma* (1967-68), but, instead, returns towards a conventional orchestral organization and plays with spatialized clusters of instrumental groups (Hofmann 2011). The eight percussionists placed in a circle around the audience create the impression of motion through rotation. In turn, this would lead to his following composition *Persephassa*, with six percussionists, also situated in a circle around the audience, experimenting with static sound objects and how their accumulation through rolls and tremoli create quasi stationary sound masses (Hofmann 2006). Xenakis would later move away from radical spatial experiments such as *Terretêktorh* Harley (2004), but would never fully abandon his spatial interest. Further spatial works include *Windungen* (1976) for 12 cellos situated in a circle on stage, *La Légende d'Er* (1977) for seven channel tape, premiered at the *Diatope*, or *Alax* (1985) for three instrumental ensembles.

2.3.2 Serialization of the Spatial Parameter

While the movement of massless, sounding objects in space was a first step to bring spatialization to the consciousness of the audience, the methodology lacked a grammar. Sound moving with speeds higher than any human could run was fascinating for the public, but it was still up to analysts to categorize them and make musical sense of it. As so often in music, its enjoyment cannot simply be taken for granted, because, as Schaeffer originally distinguished, we have several modes of listening.²⁷ Without anything to hold on to, it was impossible for the analytic mind to grasp this as a radical new concept, and so a group of composers made it their goal to place this newly discovered musical possibility into a framework to base the compositional decisions on. This group was influenced by the thoughts of Arnold Schönberg fundamentally, but more immediately by the works of Anton Webern, called *Total Serialism*.

Serialism was thought of by Schönberg as a means to avoid the musical dead end he saw harmonic music of the late Romantic period leading to. Based on the demographic thought that no pitch should be favored, Schönberg defined 12 tone music, preventing any pitch to be played before all other pitches have sounded. A melody would be a short sequence of 12

²⁷Schaeffer identified four listening modes originally: *Écouter* (information-gathering mode), *comprendre* (identifying and interpreting information), *ouïr* (perceptive, sensational listening) and *emphentendre* (listening to sound (sub-)components) (Kane 2007). Today, one would rather refer to the three modes coined by Michel Chion, which are refined versions of those of Schaeffer: *causal*, *semantic*, and *reduced* listening (Chion 1994)

notes which could be represented in a series of numbers and mathematically transformed to create an ever growing series, hence the name. Total Serialism was the effort to bring this methodology and equality into every single musical parameter, effectively isolating each from one another.²⁸ With the creation of stereophony and the possibility to create a continuous space in which sound could move theoretically faster than any mass filled particle, the desire to serialize this newly discovered parameter grew quickly with the first electronic serial pieces. One of the first composers to actively preach the musical use of space as an emancipated parameter was the German serialist composer Karlheinz Stockhausen. Starting out in Paris during the beginnings of *musique concrète*, Stockhausen experienced the first sound dispersion experiments over multiple loudspeakers firsthand. After returning to Germany, forming the *Elektronische Musik* movement, his notion of musical space was not just a musical exploration but remained a fundamental parameter considered throughout his catalogue from his early major works onward.

Although seating plans were already predetermined in the scores of *Kreutzspiel* (1951) and *Zeitmasze* (1955) (Miller 2009), it is his acclaimed composition *Gesang der Jünglinge* (1956) that would mark Stockhausen's first use of space as a musical language. Aside from its serially varied pitch, volume and duration, the serial spatialization of the sound are said to be the most fascinating features (Smalley 2000), underlining the importance of the spatiality in this composition. The piece was originally written for six channels, but, due to technical difficulties, was mixed down to five for the premiere and later saw the fifth channel mixed into the first, further reducing to four channels for the commercial release. The first four channels had sounds moving all around the audience and the fifth channel was initially intended to be placed overhead. Unfortunately, only four channel tape recorders existed at the time. For the premiere, the fifth speaker was therefore placed on stage and both tape machines were synchronized by hand (Miller 2009). The surrounding four channel sound nevertheless impressed the public enough to mark *Gesang der Jünglinge* as a milestone in spatial music composition. On his compositional intent for *Gesang*, Stockhausen wrote: "Von welcher Seite, mit wievielen Lautsprechern zugleich, ob mit Links- oder Rechtsdrehung, teilweise starr und teilweise beweglich die Klänge und Klanggruppen in den Raum gestrahlt werden: das alles ist für das Verständnis dieses Werkes maßgeblich."²⁹ (Stockhausen 1959)

In *Kontakte* (1958-60) for piano, percussion and four speakers further he refines the spatial ideas used in *Gesang der Jünglinge*. Judging by Stockhausen's sketches, Miller (2009) concludes that *Kontakte* shows an intricate and emancipated relationship between space and the more traditional musical parameters. Considering the serialization of parameters, Stockhausen here is specifically concerned with both speed and space, space being the angular position within the speaker square. More interestingly, Miller (2009) notes that "[...] Stockhausen appears to have

²⁸Even if the series was multidimensional to interlace several parameters in one transformation, the parameters were still operated on separate grids. At this point, it is debatable if composers prior to this time considered every parameter available to them in a similar isolated manner. The minor point to be made is the clinical separation as a result of the use of mathematics and abstraction.

²⁹No matter from which side, with how many concurrent speakers, whether rotating to the left or right, sometimes static and sometimes movable, the sounds and sound groups are projected into the space: all of this is of utmost importance in understanding this work.

composed the spatial movements at the same time as other parameters, which is unusual for him.” This would indicate that the physical space in *Kontakte* has a much higher musical responsibility than in any other work. His sketches show detailed trajectories, some of which indicate specific locations for specific sounds. Furthermore, he defined six categories of spatial movement: rotations, looping movements, alterations, disparate and connected fixed sources, isolated spatial points. Primarily intended for the realization of the rotation movements, Stockhausen built a *rotation table*, consisting of a turnable directional speaker being recorded by four microphones placed in a square around it. Not much analytic work has been done on the spatial aspect of *Kontakte*, which could definitely be addressed as future work.

His work *Hymnen* (1966-67), written in three versions for either tape, tape with soloists or orchestra only, exhibits another step forward in Stockhausen’s sophisticated use of space. Miller (2009) notes: ”The techniques used for spatialization in *Hymnen* demonstrate that the increasingly sophisticated technology at Stockhausen’s disposal altered his music; in effect, Stockhausen’s compositional techniques evolved along with technology. If at first Stockhausen’s goal was to use space to clarify different layers of serial polyphony, it was now having the effect of giving the listener the impression of flying around in space in a way that was not physically possible.” Compared to *Kontakte*, Miller’s quote indicates how spatial composition shifted further from being an organizational expedient towards a differentiated parameter. Furthermore, it demonstrates how technological advances helped shape the notion of space.³⁰

Of similar importance to *Gesang* and *Kontakte* in the electroacoustic discipline is the isochronally composed piece *Gruppen* (1955-57) for three orchestras in the instrumental realm. The orchestras were arranged around the audience, with one in front and two on each side respectively, and each orchestra is lead by a dedicated conductor. Although it was not Stockhausen’s intent to mimic electronic timbral soundscapes (Stockhausen 1959), the notion of sound movement by acoustic amplitude panning is an important factor for the success of the piece. Harley (1994) correctly criticizes the approach, due to the instrumental groups being too few and too distant. Furthermore, complex timbres, such as the trombone, are hard to match up for a sufficient panning illusion to create a sense of moving sounds. Small differences between the two instruments between which you want to pan cause the impression to split into two auditory streams. Mainly, the piece demonstrates dense textures of enveloping sound, as *Gruppen*’s sparse spatial movement is restricted to the interludes, while the majority of the piece consists of three orchestras playing simultaneously. The same issue is exhibited in his later work *Carré* (1959-60) for four orchestras and choirs, placed around the audience to form a diamond shape. Similar to the interludes in *Gruppen*, *Carré* makes exclusive use of moving sounds between the orchestras in ’insertions’. British composer Cardew, who collaborated in a ’circular’ composition process for *Carré*, recalls his concerns about the circular movements. Instead of flowing in a circle, the music would jerk through the room, jumping from one ensemble to another: ”[I]n the insertions, the slightest lapse of attentiveness can have disastrous results; embarrassing gaps appear in the ’spin’ – and these are very noticeable” (Cardew 1961) Nevertheless, the sensation

³⁰To the knowledgeable reader this might not be a big surprise, as any parameter in electronic music – or even in acoustic music – benefits from advances, both, respectively, electronic and acoustic innovations.

2. Nonlinear transformations (performer-centered)

	(expand in all dimensions) equivalent	+
	(contract in all dimensions) equivalent	-
	(repeat an element) equivalent	=
	(spiral-sign: repeat and transcend a previous event) equivalent	±

Figure 2.5: Some of his symbols for a possible spatial notation (Maconie 2005).

of these effects still overshadowed the rest of the piece.

Stockhausen's ideas on the serialization of space were first published in his text *Musik im Raum* (Stockhausen 1959), after completing *Gesang der Jünglinge* and *Gruppen für drei Orchester*. In it, he already expresses his desire to build a spherical concert hall with speakers all around for full three dimensional space composition and even mentions a sound translucent platform that would float in the middle for the audience to sit on. Furthermore, he already mentions the need to establish a norm for these types of concert halls and envisions them in every bigger city to spread the notion of electronic spatial music. His visions are partially realized when he was invited to co-design the German pavilion for the Osaka World Fair in 1970. Also being the artistic director, Stockhausen, together with architect Fritz Bornemann, designed a spherical concert hall with 50 loudspeakers in seven rings, also including the acoustically transparent platform on which the audience could sit on. A dedicated sound engineer would control the spatialization live, using either a spherical sensor built by the TU Berlin or a ten channel rotary mill, while musicians playing acoustic or electronic instruments, such as shortwave receivers, would provide the sonic material (Fowler 2010).

Many of Stockhausen's previous pieces were dispersed over the course of the show. Among others (also by other German composers) were some of the works previously mentioned above, as well as his pieces for a variable number of shortwave receivers and one projectionist: *Spiral* (1968) for solo shortwave receiver, *Kurzwellen* (1968) for six, and two pieces specifically written for three dimensional dispersion in the pavilion: *Pole* (1969-70) for two and *Expo* (1969-70) for three shortwave receivers (Fowler 2010). During the work on *Spiral* Stockhausen further develops his graphical notation system already used for *Kurzwellen* by adding spatial instructions for the diffusion engineer. Due to the sophisticated technology provided at the pavilion, spherical projection and instantaneous changes between different spatial states, Stockhausen could further refine his notation to an "aural architecture" for *Spiral* (Fowler 2010): He divided the concert space vertically into layers, using + and - signs as the extremes. The transitions between loudspeakers were then noted using lines and polygons (see figures 2.5 and 2.6). Other than the few pieces written for Osaka, the spatialization of the other works performed, including *Spiral*, was improvised by the *Tonmeister*³¹ himself.

³¹Stockhausen was present during the whole of 183 daily live concerts as a performer and *Tonmeister* looking over the main soundboard in the concert space. Taking turns with other musicians, Stockhausen would play a five-and-a-half hour live program of his own music every day to about a million listeners. *Spiral* exceeded every other piece being performed over 1300 times.

At this point, it is evident that Stockhausen hit a climax in his spatialized thinking. His following works lack innovative spatial material. Somewhat notable is *Fresco* (1969) in which he simulates simultaneous auditory streams by having several orchestras in different auditoriums and hallways (basically offstage) play unrelated pieces (Maconie 2005). In *Sirius* (1975-77) four soloists and eight loudspeakers are placed around the audience, with a ninth rotatable loudspeaker in the middle, rotating up to 25 times per second (Miller 2009). The perceptual impression should have been no more than an enveloping point source modulated by the doppler effect.³² His seven cycle opera *Licht* (1977-2002) sparks out so far, as its spatial parameters were not considered in the serial process. Instead, Stockhausen balanced the spatial distribution according to his intuition (Miller 2009). The spatialization varies considerably in each of the seven operas and can be summarized in many circular movements of varying speeds with the main stage action happening in front of the audience. Finally, it is the thirteenth hour of the 24 cycle composition *Klang* (2004-2007)³³ that picks up on contemporary spatial ideas: "24 melodic loops [...] rotate in 24 tempi and 24 registers [...] each section of each of the 24 layers has its own spatial motion between 8 loudspeakers, which means that I had to compose 241 different trajectories in space."³⁴ As the piece increases in density of rotating layers, the perception of individual trajectories blurs until only a statistical mass of sound is created. Although supposedly a serial composition, the perceptual effect can be compared to those created by stochastic music, effectively reaching into the domain of textural composition.

Another spatially concered contemporary of Stockhausen and fellow lecturer of the Darmstadt School was the French composer Pierre Boulez. Similar to Stockhausen, Boulez considers spatial efforts of the past as anecdotal or decorative, but also dislikes the preoccupation with circular motion of his contemporaries, because, in his view, movement creates unwanted theatrical elements, negatively affecting the music (Harley 1994). His main concern is the creation

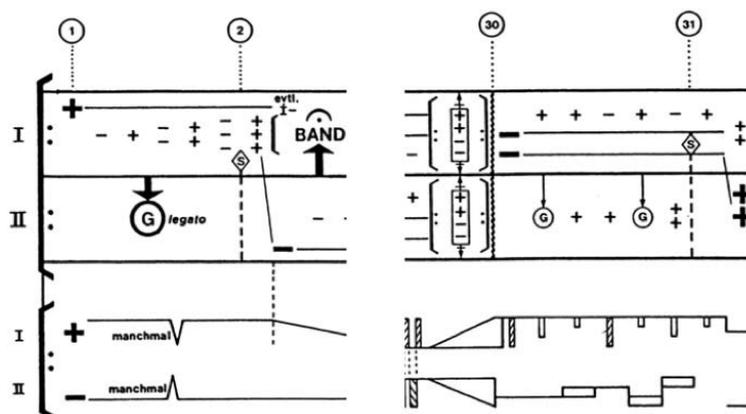


Figure 2.6: Score excerpt from *Pole für 2* (Fowler 2010).

³²Basically a leslie speaker, already invented in the 1940's, but with a bit more room projection.

³³The concept of *Klang* was to have a cycle of 24 chamber music pieces, each representing the hour of the day. Unfortunately, the work is uncompleted and there exist only 21 pieces before Stockhausen passed away on the 5th of December 2007

³⁴Stockhausen, as quoted in Miller (2009)

and alteration of spatial texture. Thinking about the dense sonority an orchestra can make, already traditionally situated in front of the audience, he alters the textures by repositioning the performers on stage.³⁵ He defines spatial intervals that are either *conjunct* or *disjunct*, meaning two spatialized chords sonically blurring into one another, or separated by an obvious pause. Furthermore, two facing groups can either be *homogenous*, of the same timbral quality, or *inhomogenous*. Together with his repositioning technique, the compositional methodology would have developed from *homogeneity* to *inhomogeneity* as the orchestra would mix and reorder itself. Later in his career, he speaks of space as the usual tool to disperse polyphonies, just like Brant and others before him (see section 2.1.2). Details about some of his pieces can be read about in Harley (1994), but since he could not set foot in the electronic music world (Normandeau 2010), this discussion will limit itself on his general methodology, given above.

The third of the Darmstadt Musketeers is Luigi Nono, and just like his colleges Stockhausen and Boulez, Nono, although not until later in his career, had formed his own thoughts about space: "No doubt, Nono will have felt Stockhausen's and Boulez' breathing down his neck. But it catches the eye how late Nono introduces this new parameter in his work, and how he treats it totally differently than his competitors." (Beyst 2003a) Like his forefathers of the Venetian Renaissance, the San Marco cathedral but also the city of Venice itself is a source of spatial inspiration, as it is the case with *Prometeo* (1984), intended for the San Marco cathedral. The unique acoustic landscape of Venice at night is described in Beyst (2003a), with its many canals and phantom footsteps echoing through the alleys. In contrast to the contemporaries of Gabrieli, though, his experimentations consider the full orbiting space around the audience trying to abolish the absolute center in the concert hall, forward-directed by our sense of vision. To Nono "perception and listening have the power to free the ears from monodirectional, sight-dominated, selective habits."³⁶ His overall work can be seen as a transition of the conception of space in the 1950's into contemporary views of composing the space itself, as opposed to placing sounds in space. While originally a serialist composer, his thoughts on spatial composition center around spatial perception instead of mathematical series. Sounds do not merely exist in space, but they interpret the space as the space opens itself to the listener (Guerrasio & Palme 2008).

2.3.3 Musical Space in the Computer Age

There is another factor besides the technological advances that shaped the use of space, which was virtually ignored by previous approaches, including that of Stockhausen. Not that serialism was considerably concerned with the perceptual effect the listener experiences in any aspect in music other than the intellectual decoding³⁷, but when it comes to spatial composition, psy-

³⁵It seems that this repositioning happened silently, and in succession, so that one musician at a time could rather unnoticeably move around, since Boulez tried to prevent any kind of theater in his music, including any sort of trajectory to appear.

³⁶Luigi Nono, as quoted from Guerrasio & Palme (2008)

³⁷Many postmodern movements share this view and consequentially break away from serialism. One of which is *Spectralism*, with Gérard Grisey and Tristan Murail at the forefront. For the interested reader, I would refer to my previous work (Schmele 2010) and the references given therein.

choacoustics play a considerable role. For example, Stockhausen refused to include distance as a musical space parameter, because distance was too intertwined with other musical parameters and thus was not serializable. What he did not care about, though, is the fact that, due to our pinnae acting as a filter, sounds change in tone color when moving around our head. More importantly to his serialist approach, most listeners are not at the center of the speaker circle, implying a minor sense of distance since, in this case, not all points on the periphery are equally spaced. Furthermore, as [Harley \(1994\)](#) points out, his segmentation of the perimetric space into a *spatial scale* is "perceptual nonsense".

A break from the trajectory-oriented view of sounds as objects in space is the approach of the American composer Alvin Lucier reverting the idea of spatialization and projecting the physical room back into the timbre. He did so most famously with his piece *I am sitting in a room* (1969). This piece, strictly minimal in its compositional process, exhibits the rooms reverberant frequencies by refortification through repetition of the original re-recorded sound material. At first, the listener is instructed by Lucier himself on the process, which, in turn, is the exclusive sound material for the piece. By continuously recording the projection of his voice (respectively previously recorded) he manages to amplify the room's resonant frequencies, until eventually rendering his speech intelligible. What remains is a tone timbrally morphing based on the general shape that the original speech dictated. For one, the piece is dependent on the tone of the voice and text that somewhat determines the speech melody, as well as possible imperfections in the hardware used. More interestingly, though, is the dependency of the room that this piece is played in, turning the performance space into a characteristic instrument.³⁸ His experiments also investigate the room in the sound on a physical level, meaning "[...] standing waves, phantom sonic images, diffractive properties of the sound, etc." ([Harley 1994](#)) His installation *Outlines of Persons and Things* (1976) demonstrates the effect of acoustic shadows, inviting the visitor to experience the timbral changes as they move around the space. Similar to the thoughts of Boulez, Lucier avoided all that invokes programmatic impressions. Lucier's intent was a very scientific one, being interested in phenomena of acoustic and psychoacoustic nature. Words that invoke metaphors would only distract from what is actually to be explored. "Yet, often, there is no concert to attend, no performers to applaud. The context for "musical" explorations has been drastically revised" ([Harley 1994](#))

One spatial dimension that has been problematic to reproduce electronically³⁹ is a parameter that is an important component in the creation of vast soundscapes and has been used in acoustic music frequently since Ives' thoughts on space at the latest. John Chowning is one of the first electronic composers to get a grip on spatial distance, while also being one of the first to use the computer for both a spatialization and sound synthesis tool ([Zelli 2001](#)). His goal was the gestural control of space, signifying a clear musical intent after having heard "plenty of electronic music [from] Europe [...] which attempted to utilize space in a fairly primitive way." ([Roads 1985](#)) Being one of the lesser prolific composers, his spatialization algorithm was

³⁸The author was able to listen to a recording of this piece at the ZKM in Karlsruhe, Germany

³⁹Still today, as can be seen in chapter 3.1 distance perception is a phenomenon which is not very well understood

mainly used in *Turenas* (1972), written for quadrophonic speakers, which was also the speaker configuration the algorithm was built for. Commenting on the innovation of his composition, Chowning points out the first use of the Doppler shift for a more accurate distance perception (Zelli 2001). The sound material consisted exclusively of his innovative FM synthesis method, and while a pitch to space relation was a natural bodily connection, a mere timbre to space relation was new ground for Chowning to explore. Chowning's spatial language mostly consisted of figures following trajectories. He is reported to have made heavy use of Lissajous curves in *Turenas* (Zelli 2001), and while the exact identification of such a trajectory is perceptually questionable, his report on the pleasant sounding characteristics of these curves seems to be of spatio-timbral nature.

A contemporary of Chowning, being born just a little more than a month earlier, and another American composer with a particular interest in spatialization techniques is Roger Reynolds. In his writing on spatial sound movement (1978), Reynolds searches for a meaningful grammar suitable for its interpretation. Compared to, for example, Stockhausen, Reynold's approach to azimuthal organization is much more established as he considers psychoacoustics and bodily experience. He points out several deficits in real world experiences that hinder our perception to draw appropriate analogies between pitch and space. But, when alluding to symbols and concrete examples of movement, the listener gets a chance to interact with his impressions. In short, Reynolds concludes that spatial building blocks, abstract, clinically isolated trajectories, cannot transport any meaning in themselves. The sonic space, nevertheless, can be composed as the mere notion of "I was moved" carries valuable information (Reynolds 1978). With that in mind, Reynolds points to dance and the efforts of Rudolf Laban to encode movement. Among his most famous works are the early *Voicespaces* pieces which explore the possibilities of sound location with the premise of the voice as a sound source being the most familiar sound to our ears. *Voicespace III: Eclipse* (1979), for example, was composed for the Guggenheim Museum and made use of elevated techniques by projecting sound from the top of the large central atrium (Zvonar 2006). While at IRCAM, Paris, he composed *Transfigured Wind II* (1984) using computer analysis to fragment, transform and spatialize the solo flute over a quadrophonic speaker setup. *Watershed IV* for solo percussion and interactive spatialization seems not only to move the sound around and above the listener, but "[...] sometimes the listeners themselves seem to be in the performer's position at the center of the percussion kit" (Zvonar 2006).

Going towards the formalization of sound movements, we have to look at Trevor Wishart, who, next to his favored use of the human voice, also made sonic motion a core concept of his compositional methodology (Zelli 2001). His categorization of different motions is done in six high-level groups (Zelli 2001): *direct motions*, encompassing different general path types, *cyclical and oscillatory motions*, including circles, spirals and swinging motions, *double motions*, which are complex motions of two types, *irregular motions*, *frame motions*, by which a grouped "super-motion" of all motions in parallel is meant and *counterpoint spatial motions*, the juxtaposition of two or more pathways. Out of his six cycle work *Vox* (1982-88), a short analysis of *Vox1* can be found in Zelli (2001): It consists of seven sections, each with a specific spatial category in

mind. There is a general development from chaos to organization throughout all sections. While the first two sections introduce the voices (two female and two male), having them emergence from silence and finally settling them in different positions, the voices are scattered within the whole space in section three, portraying a lot of kinetic energy. The fourth section, named *Relaxation* brings this chaos to rest until they are reorganized in the frontal regions, whereon all voices remain motionless throughout the fifth to the seventh and last section.

With the difficulties connected with defining a spatio-musical language in written form, tendencies in spatialization went from bring composed to improvised. Two major factors contribute to this movement: lacking any "true" approach to spatial interpretation one must start with an egalitarian attitude, and, secondly, spatial improvisation implies live diffusion, which requires interfaces for the performer to interact with. Considering the first point, John Cage, already briefly mentioned in section 2.1.2, is considered to be the major influential figure. Specifically, one can look at his *Gesamtkunstwerk HPSCHD* (1967-69) to understand the aesthetics of programmed chaos in his "five hour multi-media extravaganza" (Harley 2004). Seven harpsichords play pieces by Mozart and modifications of the like by Cage simultaneously. Accompanied by 51 tapes with music in different tuning systems, being projected from different locations respectively, the sonority, comprised of a set of indiscernible streams, forms a dense, enveloping texture. Part concert piece, part installation, the audience is the real performer as they are encouraged to walk, talk and experience the work from different angles. Pierre Henry and the *potentiomètre d'espace* were already mentioned in the start of this section 2.3. For Henry, the performance of *musique concrète* with only prerecorded sounds, organized on tape, remained in the act of live diffusion. Unlike other composers interested in spatialization, Henry's approach was not concerned with exact placements of multi-speaker systems to achieve accurate phantom images. Instead, he placed a battery of inhomogeneous speakers on stage, in front of the audience, controlling (performing) the source audio behind a mixing desk. The motivation of this approach was the intent to articulate his music by diffusing it through the appropriately sounding loudspeaker (Zvonar 2006).

This technique influenced composers and artists particularly in France, Great Britain and Canada, and poses an alternative trend to the precise, but fragile, homogenous surround sound composition. Here, exact positioning is not the issue, as Andrew Lewis wrote in 1998: "[...] 'real sense of space' [...] is a Western, classical, mainly nineteenth-century, mainly instrumental, academic approach which can only digest musical ideas which are formed from quantifiable, notatable, discrete parameters."⁴⁰ Lewis is one of the original members of the *Birmingham ElectroAcoustic Sound Theatre* (often referred to as BEAST), which, together with its predecessor *Acousmonium* by François Bayle at the GRM in Paris, represent the two major non-homogenous – and mobile – loudspeaker presentation systems. In the case of BEAST, there are over 100 loudspeakers available, differently arranged based on the space or the piece the system is constructed in and for. For example, there are different loudspeakers with different characteristics, but also care is hereby taken to account for psychoacoustics, using tweeters to but hung from

⁴⁰Lewis, as quoted in Zelli (2001)

above, for example. When writing about the compositional approaches to compose for such a large system (Wilson & Harrison 2010), its curator Jonty Harrison talks of pragmatic, non-specific spatialization. The flaws of any loudspeaker⁴¹ are turned into the strengths of complete system. The pragmatism is, therefore, not about any exact location in space, but the composer is rather concerned with the sound itself and how to articulate its inner structures through is (vivid) spatialization.

An active practitioner of live sound diffusion is Canadian composer Robert Normandeau. For Normandeau, the spatialization process is carried out with respect to the minute detail, and this has significantly affected his compositional workflow (Zelli 2010). Instead of working with non-homogenous systems, he employs similar aesthetics with more conventional loudspeaker arrays, such as 8 channel surround. Normandeau sees compositional space as a duality *Normandeau09*: first there is the *internal space*, which is the space, both spherical through multichannel composition and distanced with the appropriate techniques, imposed onto the tape while composing in the studio. This space is invariable and part of the composition. The space must then be opposed against the *external space*, meaning the performance space into which the previously composed space is projected into. Normandeau's need for live diffusion is, therefore, clear, as the *external space* can be variable if the piece is non-stationary – if it can be performed in different locations, and, while it is not necessarily part of the composition per se, the spatialization process is mostly improvised to react to the changing localities. When it comes to the actual act of spatialization, sounds are not simply diffused between a set of loudspeakers, but their spectral components are separated and the sound is literally smeared across the listening space to certain degrees.

One way of achieving this is with the use of bandpass filters. He exemplifies three gestural techniques using this approach (Normandeau 2009): in *StringGDberg* (2001-03) the source sounds are gradually introduced as the filters open up over the course of the work. *Éden* (2003), on the other hand, contains many superimposed timbres without a progression, so that attention instead is paid to the timbral behavior of the sounds. But in his discussion about *Kuppel* (2006-09) for the *Klangdom* at the ZKM in Karlsruhe, Germany, where the differences between the previous approaches and systems, like the Acousmonium, become apparent. While the *Acousmonium* approach considers the speaker as part of timbre, the effort of the *Klangdom* is to make the space completely virtual and transparent. Acousmatic music, in its effort to use the advances of audio technology, has the desire to completely separate the sound from its source, seducing the listener to concentrate on only the sonority itself. With the speaker being part of the sound the original intention has failed even before the music reached the ear. Coming back to spatialization in his music, there is a significant comment he made in a lecture given in 2010⁴², stating his resentment of his 5.1 channel DVD release *Puzzels* (2005),

⁴¹I may refer to statements made in Wilson & Harrison (2010) here and comment that I do have to agree, although the perception of these flaws (using a professional set of speakers) are questionable and probably depend on the situation. As Goßmann & Dombos (2003) state, even though every representation within a medium is always medium specific, the goal would be to have the medium become intuitive. Only then would it be transparent.

⁴²The author attended a performance of his multichannel pieces *StringGDberg* and *Éden* at DMARC, Limerick

not only losing the aspect of live diffusion, but also having to crunch the multichannel work into a commercial standard. For the spatial composer channel reduction can be compared to bandwidth reduction in the frequency domain: "[Robert Normandeau] feels frustration from the fact that his multichannel works are often produced in stereo format." (Zelli 2010)

Finally, we reach upon the ideas on spatial composition in the writings of Denis Smalley, composer and professor of music in London. Smalley's impact on spatial musical thinking is immense, as he coined the term *spectromorphology* to describe, for one, the temporal development of a sounds spectrum (Smalley 1997). In the same article, though, he also touches on the potential of the spectromorphology dictating a spatial movement through its inherent energy – or even its ability to direct itself exploiting psychoacoustic phenomena. Smalley recognizes the ability of space to change the sound's spectromorphology. Space is not just a parameter the composer can change at will, one needs to be aware of the impact it has on the sound and the changes that happen to the actual music (Austin & Smalley 2000). Space as an appreciative experience in itself is, in turn, coined *spatiomorphology*. He distinguishes spatiomorphology "[...] from regarding space only as spectromorphological enhancement." (Smalley 1997) Simply put, this is where Smalley delineates space from being a mere effect as opposed to a parameter suitable for musical expression. Similar to Normandeau, Smalley also actively engages in diffusion for the purpose to react to a specific performance space and its audience. Furthermore, live diffusion is a method for him to articulate the spaces composed into the piece on tape in the studio: "[...] one should be able to expand these dimensions [...] thereby adapting the space composed into the music to the dimensions of the listening space." (Austin & Smalley 2000)

For Smalley, a spatial image develops over time, but it remains in his memory as a single, collapsed moment of presence (Smalley 2007). A spatial scene needs to be experienced first through time. The listener needs to be given the chance to change into a state of mind of active spatial hearing. He must traverse the sonic image through space with his ears and collect different spatial impressions individually. But, finally, this temporal exploration is reduced into a single image. For this reason, Smalley (2007) writes: "Time becomes space." Furthermore, Smalley understands the perception of space as a full body experience and it would be a mistake to think of sound as an exclusively aural perception. Trans-modal perception, a light form of synesthesia that is present in everyone, lets you connect the things you hear with the abstract concepts or real objects that would best fit an association. When an audience member listens to sounds in an acousmatic context, he usually imagines what these sounds could relate to in the real world, perhaps having a visual or even a haptic perception. Smalley goes on to quote Michel Chion: "[...] the sounds which we hear are connected by us to our intellectual knowledge of context, to vision and the voluminous representation of sonic causes, and to our general sensory experience."⁴³

This might be one reason, why his categorial approach towards the division of space into a framework is predominantly earth and bodily bound. His three general classifications of space (Smalley 2007) are spaces that are created through human interaction, *enacted spaces*, which

on the 24th of March, 2010

⁴³Chion, as quoted in Smalley (2007)

include different scales of performance spaces. *Spectral space* covers the spectromorphological aspect of sounds and the relation to spatial images. Finally, *perspective space* covers the different types of spaces that may exist around the listener. In a nutshell, Smalley renames many spatial attributes of sound and, in this effort, tries to refine the definitions of each respective concept. For example, what one might refer to as reverberation, or, more simply put, the acoustics of a performance hall inevitably contained in a recording of a live performance, is dealt with by using his definition of *arena space*. The listener, in this regard, is not referred to as 'the listener', but instead as *egocentric space*, while the performance on stage, being in front of the listener, is included in the category of *perspectival space*. More importantly though, by abstracting these concepts he is able to group several phenomena within one set of terms and, in turn, by giving each concept a similar name, he brings all aspects of space into one new language. The hope is, that once the composer or analyst has incorporated this language into his normal musical thinking, the spatial parameter will gain importance as people are able to discuss the spatial aspects of a piece in a unified way. This is very well comparable to the situation, where one wants to describe a melody either with or without the knowledge of how to name the individual notes and intervals. The discussion over a concept is essential for it to have any chance of flourishing and nesting itself into the cultural consciousness.

Chapter 3

Spatial Composition Methodology

When we sit in front of a piano, we can press a key and hear a pitch. If we press two keys, we can appreciate a new quality labeled an interval, and two intervals together form a chord, which has the power to move us emotionally – be it as simple as the distinction between a major and a minor chord, or if we finally, by randomly hitting different keys together, going through several chords of dissonance, encounter a harmonic combination. The auditory world around us is mostly composed of noises and the refreshing sound of a harmonic overtone sequence makes us distinguish this phenomena as something unique. These sounds are mostly made by energy introduced in vibrating objects, which, both the energy and the object as such are usually of human origin. Producing a pitched sound, in the example of a piano¹, does not require skill. It is the temporal organization, the "correct" succession of these sounds, the ideas and intentions we identify with this particular human being for making the *musical* decisions on this object he calls his instrument that delights and excites us, possibly inspires us and makes us want to *listen* (and perhaps see) to what is happening and how his musical endeavor is going to continue.

Edgar Varèse once famously defined music as *organized sound*, a term that, with the liberation of all sounds in the course of the Modernist Period, we have come to accept. It is therefore not just the pitches or harmonies, and also not just the inharmonic sounds (the noise) that we organize with deterministic or stochastic methods² – but also the instrumentation, the timbres, the articulations,... – even the abstract associations that composers and songwriters have managed to organize in time. For example, in 1959 Giacinto Scelsi wrote his famous *Quattro Pezzi su una nota sola*, concentrating on all other qualities in music, most notably on the timbral frictions, neglecting any pitched composition by focussing each piece on a sole pitch respectively. Since the 20th century, music has experienced a separation of different musical parameters, gaining insights into the inner workings of its microstructure and effectively rendering every single one of them a valid entity of compositional value. Music of the 20th century, more so than ever before, wants to teach about the details of human hearing and tries to heighten the general level of auditory perception of the public. "Ich meine, [...] das formt neue Menschen.

¹In other instruments, such as the violin, the actual pitch production can be the lone reason for admiring the skill of a performer. But, what do we admire in the case of using the dials of a radio, for example?

²The mere conscious notion of placing sounds stochastically over time makes it a method of organization.

Mit denen kann man dann zum Teil auch über Sinnvolleres reden als vorher, das muß man schon sagen. Über bestimmte Bewegungen von Klängen zum Beispiel, und was durch sie geschieht.”³ (Stockhausen 1964)

As outlined in chapter 2, these trends also influenced the notion of musical space and spaciousness in music. Unlike pitch perception, though, spatial hearing happens with every sound around us that we are subjected to every day. Although pitch perception is no selective process either, spatial decoding of the auditory scene around us always contains many, mostly minute, sounds from every position within our auditory sphere, so that the process is as unconscious as our breathing (Guski 1992). Furthermore, the brain usually filters unwanted sounds and noise to clarify the actual sound source of interest coming from a specific position.⁴ The human is generally considered a visually oriented being (Neuhoff 2004, Zelli 2010) and – even though this filtering process can be consciously controlled, it, therefore, usually concentrates on the frontal area, the region that is covered by our sense of sight. Consequentially, when we hear a sound of interest, we usually, almost instinctively, turn our head to its apparent direction for better focus (Middlebrooks & Green 1991, Howard & Angus 2009). Despite our ability to hear spatially, we immediately reduce our consciousness to only that what we can see, naturally diminishing the spatial auditory experience, pushing the original sensation of a remote sound into the subconscious.

Generally, the listeners sensitivity to ”unfold spatial relationships in musical performance is [...] ill refined” (Solomon 2007). For the traditional instrumental composer, this implies a non-linearity if one were to directly derive a compositional methodology ranging from pitched or even any style of acoustic composition in general. Let us start from the beginning of this chapter, but this time, consider spatial composition: If we place a sound in space, we hear that sound coming from this specific location. If we place two sounds in different locations in space, we either hear these sounds coming from these two specific locations in space, or, if the sounds share a strong timbral similarity, we perceive a blurred impression of the sound being smeared across the uninterrupted straight line between the defined locations. When using loudspeakers, one can play the exact same sound through both speakers with either a time delay or a level difference, creating an effect known as a virtual source: the sound will appear coming from a position in-between the two speakers. If we vary this difference, we employ a technique called panning, effectively moving the sound on a trajectory from one speaker to the other. Initially, this ability inspires composers, like Stockhausen, to quickly think of complex trajectories as a gestural quality – today composers abstain from putting too much emphasis on this approach. A recent study by Marentakis et al. (2008) has concluded that the identification of spatial trajectories is generally poor and even decreases further, depending on the listening position offside the sweet spot. A later study also investigated the musical potential of sound trajectories and resulted in a very low recognition rate of spatial patterns, ”[...] even if we have only 4 possible patterns that are known beforehand, and that they are heard alone, without

³”I mean, [...] this forms new people, with whom you can talk about much more meaningful things, you have to admit. About certain movements of sounds, for example, and their causes.”

⁴This is called the *cocktail effect* and will be dealt with in section 3.1.3

coinciding sounds as we would have in a music or audiovisual work.” (Payri 2010)

Nevertheless, hope arises when “[...] there was a significant improvement in recognition with headphones as opposed to loudspeakers.” (Payri 2010) Furthermore, as was discussed in chapter 2 and as we will see later in this chapter, composers have looked for methods to go beyond the approach of sounds in space through trajectories. But, most importantly, other solutions have been found by merely being aware of the spatial limitations and general misconceptions of our auditory perception. Site-specific installations remove the notion of temporal development and invite the listener to move freely through the listening space, eliminating the sweet spot dilemma and having the listener explore the soundscape on his own, relieving the composer of the responsibility of delivering a stable soundscape. Furthermore, composers like Normandeau (2009) reject accurate spatial sound placement due to the generally poor psychoacoustic abilities and instead diffuse several spectral layers through different speakers for a changed, but reportedly improved spatial perception. It is, therefore, essential that I discuss the psychoacoustic implications of spatial composition, which will follow after this introduction. I will then scratch the surface of a philosophical discussion about the notion of space, aiming to arrive at a definition of musical potential. Finally, I will discuss three major approaches to spatial composition, followed by a comparison of analytic methods and arrive at a new interpretive possibility based around auditory streams.

3.1 Psychoacoustic Implications

Spatial composition can be a fragile construct. Virtual sources collapsing into one speaker, distances being overestimated or the audience shadowing each other in the acoustic field are just some of the more common traps, faults and misconceptions that the spatial composer encounters. The finished composition might not transport properly from the studio into a concert hall setting. Trajectories that were clear at first might be blurred during the performance. Textures could change their color or disintegrate into their individual parts. Stockhausen, for example, subdivided the circle around the audience into equal segments, obviously without any perceptual basis (Harley 1994). Our angular resolution is not equal in every direction, and so, while in the case of Stockhausen frontal segments might be differentiated, those at the side of or behind the listener might be confused. It all depends on the composers approach and goals, and knowing about psychoacoustics helps him model his aesthetic notions around the listeners limitations. The solutions to account for are as plentiful as creative and artistic, and it is therefore not the goal of this chapter to provide any single one. Instead, by knowing about how our hearing system works, the spatial composer should devise his own method of dealing with its characteristics – perhaps even for his aesthetic advantage!

3.1.1 Localization on the Spherical Periphery

In the horizontal half-plane, the horizontal area in front of us from ear to ear, the angular auditory stimulus are very well understood. We usually hear with two ears, located opposite

of each other on either side of the head. It is important to note that the ears have a relative mean distance of about 18cm between them, and also that this truthfully is the diameter of our head, a massive sound blocking object. Both these facts cause two effects to appear. The first is an *interaural time difference* (ITD) that occurs in a linear relationship to the diameter of the head. When a sound arrives at a certain angle θ to the listeners head, it will first arrive at the ear closest to the sound. The farther ear will not receive the sound until after a time delay that is proportional to the relative extra distance (see figure 3.1). Taking the diameter of the head d , that distance could be triangulated with $d \cdot \sin(\theta)$. But, as Howard & Angus (2009) point out, the calculation is inaccurate as it underestimates the delay. This model does not account for the sound having to travel around the (assumed round) head, another distance of Δd (cf. figures 3.1a and 3.1b). With $\Delta d = r\theta$, and c being the speed of sound (approx. 343.2m/s at 20°C) we end up with the ITD cue being (Howard & Angus 2009):

$$ITD = \frac{r(\theta + \sin(\theta))}{c}$$

Although the speed of sound is relatively constant and the same time delays should appear for any given frequency, this method does have an upper limit. It exists, because of the mechanism that resolves the time difference: due to the time delay, a phase shift appears when comparing the sound at each ear (this is why the ITD is sometimes also referred to as *interaural phase difference* (IPD)⁵). As of a certain threshold in frequency, this phase difference becomes ambiguous, which is at its lowest with 743Hz at 90° (Howard & Angus 2009). For higher frequencies, the head itself, being a physical sound shielding body, comes into play. Depending on the sound direction and frequency, the intensity ratio varies up to 35dB as a

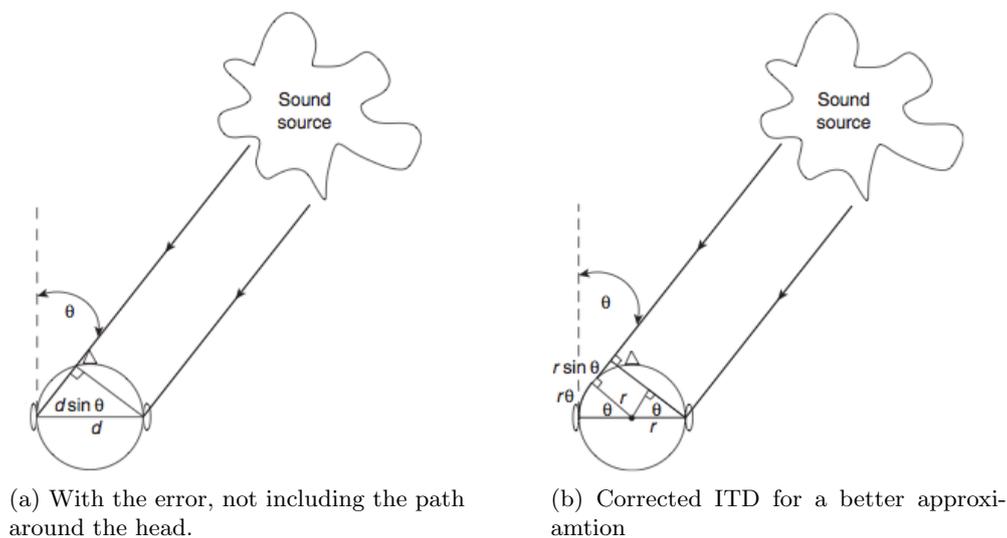


Figure 3.1: Two diagrams showing the ITD approximations (Howard & Angus 2009).

⁵Naming this phenomena can also be of historical motivation, as one of the first to investigate this phenomena is Reyleigh (1875), concluding if a low pure tone is perceived as being either left or right, the only alternative had to be a phasing phenomena.

result of shielding the sound at the farther ear (Middlebrooks & Green 1991). To become a shielder, though, the head must be at least two thirds the size of the wavelength. Therefore, a low threshold in frequency of 637Hz can be calculated (Howard & Angus 2009). Generally, a transition phase from about 700Hz to 2.8Hz is mentioned, in which our directional hearing is the worst, due to both ILD and ITD confusions. Using pure tones, localization worsens with increased angle towards the side of the head. Using the same angular change, ILDs with complex tones, though, locally vary considerably (Middlebrooks & Green 1991)

The above mentioned cues are dependent on the angle of the sound source to the head direction, perpendicular to the interconnecting line between our ears. It is clear that a source from behind the head at the same angle would give the same stimulus and requires an another cue. Furthermore, ILDs and ITDs do not consider elevation at all, creating a whole circle around the ear of identical cues called the *cones of confusion*. What helps resolve this confusions is the so called *head related transfer function* (HRTF). It describes a filter dependent on both the horizontal azimuth θ and elevation ϕ , as well as possibly the distance r (Howard & Angus 2009). These comb filtering effects occur when the sound first hits our pinnae⁶ and is reflected through its complex curvatures, causing a set of tiny delays dependent on the sounds direction, typically above 5kHz. Because our pinnae is like a fingerprint, the effect is different for every person, and the cues are learned from the early childhood on. Hearing the sound recorded through a different persons ears may change our ability to hear spatially both positively and negatively (Howard & Angus 2009). Asking the audience members to tie back any longer hair, for example, might benefit spatial hearing (Lakatos & Shepard 1997), as the hair may dampen the pinnae's effects⁷. Known sources are more easily detectable, because the position dependent filtering is easily detected. Since the HRTF is constant in time, moving the head will maintain the troughs and peaks for each position relative to the head in the filter spectrum. When listening to acousmatic music, which is concerned with unfamiliar sounds, the spatial location of these sounds can be improved when moving the head for a better distinction of the characteristically filtered frequencies (Middlebrooks & Green 1991).

The broad spectral dependency of the HRTF has several implications. Accurate vertical localization is most accurate with sounds containing a wide spectrum, especially with considerable energy in the high frequency region, above 4kHz (Middlebrooks & Green 1991). Furthermore, varying the source spectrum can have an effect on either the vertical localization or create front back confusions. Energy in the 4kHz region improves sound localization in the horizontal plane (Belendiuk & Butler 1975). An energy boost in the 8kHz region, for example, will elevate the source, no matter where it is projected from. These peaks in the HRTF are defined by Blauert (1997) as *boosted bands* and also roughly exist for the front and rear locations. Moreover, due to the spectral filtering of the HRTF, a received signal is changed in timbre based on the direction it is coming from. This effect is best noticed when circling

⁶The whole head and torso are also said to play a role in the HRTF, although the severity is still debated. The torso, for example, plays a crucial role in determining sounds from below (Cheveigné 2003)

⁷A study explicitly proving this assumption was not found and it might be worth investigating which preparations audience members should undergo to be best prepared for an acousmatic performance.

a source of white noise around the listener or moving your head in front of a fountain, for example. In turn, "[...] narrow bands of noise [may] appear to move back to front in azimuth as the center frequency of the band is varied" from 4 to 10kHz and again from 10kHz to 14kHz (Middlebrooks & Green 1991). Furthermore, Andreeva (2004) discovered that sound motion through frequency suggestion has a low temporal threshold, meaning sounds that are sustained longer will be better perceived as moving. His estimate of a lower threshold for radially moving sounds is around 400ms, for example.

But at the same time, what do you do with elderly people in such a concert setting? If the situation requires high frequencies, how is their ability to listen spatially? Malinina & Andreeva (2008) researched into this and found considerable deficits in elevation movement detection, unless it was attributed an increase in level. He lists several reasons as to why this is the case, one of which is of course age-related hearing loss, but another is also the loss of spectro-temporal analysis and simply forgetfulness of acoustic information. One critique, though, is the use of synthetic, non-individualized HRTFs, which are known not to work with every set of ears or give false impressions. Furthermore, there is no sign of training in his experimentation methodology for the subjects to accustom to the synthetic HRTF for possibly better results. Nevertheless, it is an important notion, that the age group of the audience might have an effect on the reception of spatial gestures and auditory streams.

Moreover, Smalley cites the research of Neuhoff (2004), suggesting that perceptual modalities interact with each other when experiencing space. Foremost, it is Smalley's intent to show that auditory and visual senses interact, both in experiencing space through everyday life, but more importantly, when listening to acousmatic music the visual senses are, in turn, triggered as well. While this observation implies a form of minor synesthesia in every human being, it also demonstrates the importance of multi-modal experience of space: "Auditory, visual, and somatosensory cortices have been show[n] to converge on multisensory neurons in the superior colliculus. The same multisensory neurons have also shown to mediate motor behaviours." (Neuhoff 2004) Therefore, mapping out auditory space is a neurological task: while everyone perceives auditory cues are based on physiological functions of the ear, we have to learn to associate them with the actual positions. As we explore the world from a very early age onwards, we start to align our maps of vision, audition and haptics, among others, with vision being the most dominant and of highest resolution within its limits. As we improve this alignment between each one, we improve our overall spatial orientation.

The HRTF is a learning process. This can be shown by inducing the recordings via headphones bypassing the pinnae of a listener from another person's ear directly into his ear canal (Howard & Angus 2009): the ear will readjust its spatial cues based on the HRTF information from this other person. The hypothesis here is, that the visual cues are important at the start of a child's auditory development, as the filtering cues created by our pinnae are not absolute. The brain has to learn how to associate these filtering ranges and changes in spectrum with their appropriate location. Given that this is a neurological process, the ability to narrow down angular resolution should be possible through training, but already can be achieved by the mere increased consciousness of perceiving our world aurally. Spatial music can increase this aware-

ness, and with an increased popularity of spatial music, this consciousness can be awoken. It is therefore important to further exploit the spatial parameter as an equal parameter and quality in music. The result is a win-win situation for both sides.

This claim can be supported when investigating the auditory localization skills of blind people. Studies have shown a better performance in blind people localizing sound sources (Ashmead et al. 1998, Lessard et al. 1998), especially for early blind persons, but more recent studies (Zwiers et al. 2001, Collignon et al. 2007) have reported an actual deficit in source location for blind people. As Zwiers et al. correctly point out, studies suggesting the supra-natural ability to hear spatially only investigates the horizontal plane and did not specifically test spectral localization cues. His study shows that the perception of elevation in early-blind is worse than that of sighted. Their reasoning is in accordance with my observation, that, since the HRTF is a neurological process, some form of training has to precede the correct perception, calibrating the neurological representation of the different spherical regions. They assume, that for early-blind persons the HRTF representation is blurred due to a reduced quality of cross-modal feedback.⁸ The study also finds that in the horizontal plane, localization of blind people was equal to, if not better than, that of sighted, as early-blind did perform specifically better in the lateral and rear regions, where vision is poor or absent for both parties and all rely more on haptic and motor feedback to calibrate their auditory inputs. Another study undergone by Voss et al. (2004) came to the same conclusion in the rear field for both early- and late-blind persons compared to sighted ones. This furthermore discriminates the HRTF cues from ILD and ITD cues, which instead can be viewed more as "physiological reactions"⁹ and seem to develop well despite impaired multi-modal sensations.

Finally, the comparison of late- and early-blind people undergone by Voss et al. (2004) supports the hypothesis that overall everybody can sharpen their spatial hearing. The results show that while both groups on average outperform sighted people, especially in auditory regions outside the visual field, but even more significantly in the estimation of distance. There was virtually no difference between early- and late-blind persons, indicating that multi-modal restructuring can happen even in the mature brain. Late-blind people even slightly outdid their early counterparts during the task of distance estimation. This is probably due to the initial benefit of visual feedback better calibrating the input senses at large distances, where haptics and motor sensory fail. Simply put, in a nutshell, if one were to stop and *listen* to his environmental world around him, closing his eyes and construct the visual scene in his head, effectively nurturing his cross-modal perception every day for one minute, a possible overall improvement in spatial consciousness could be achieved. At this point, I can state that I think of my own spatial compositions visually, in shapes and figures. Anticipating the compositional practice of section 3.3, these thoughts and allusions are personal, as I cannot guarantee that any

⁸If absolutely no visual stimulus is present, cross-modal associations between the haptic and auditory stimuli should still be possible for most early-blind people

⁹**Do not quote** me on this one. My use of the word "physiological" is incorrect and of creative nature! Although ILDs and ITDs are pre-filtered in the cochlear (Cheveigné 2003), they are a final matter of the brain, inter-aurally comparing the incoming neural signals. Still they seem to have the implicitness of a physiological reaction, being more easily learnable with specific deficiencies in multi-modal perception (Handel 2006)

audience member has the same sensations. Therefore, I have to compose my spatial gestures in layers of abstraction, making sure they conform with psychoacoustic principles and, while still maintaining my personal intentions, are still perceivable as meaningful even if the same shapes do not come to mind.

One further neurological process aiding in the localization process is the so called *Haas Effect*, otherwise also known as *precedence effect*, *summing localization* or *law of the first wavefront*. Its main function is to clarify source locations in spaces that exhibit reflections reaching the listener. The brain will locate the sound coming from the direction of the first wavefront that reaches the ear. All reflections arriving within a temporal window of about 30ms after the first wavefront are fused into a single sound, attributed to the original perception (Howard & Angus 2009). These reflections reaching the ear thereafter are perceived as echos. Among the severe acoustic implications, this phenomena is also important for the perception of apparent source width (ASW) and that of distance. ASW is usually contributed to the spaciousness, as strong first reflections seem to pull the sound apart within the auditory perception, similar to the effect of stereo panning (Cabrera et al. 2004). More importantly, and leading into the next section 3.1.2, is the impact early reflections have on distance perception. As the source retreats, the difference between the direct path of the first wavefront and the deviated path of the reflection decreases. This effectively decreases the delay between incoming reflections and the direct sound, which is said to contribute to an impression of larger distance (Shinn-Cunningham 2000). But because the mechanism that lies behind the correct auditorial estimation of distance is so complex, we will have a look at it in the following, dedicated section.

3.1.2 Distance Perception

First used by many acoustic composers, foremost by Brandt, in form of off-stage performers, then rejected by serialist composers, most famously by Stockhausen, because of its interrelation with many factors such as dynamics, and revived in the writings of Chowning (1971), by introducing it into the computational world, the perception of distance has turned out to be somewhat of a mystery and even to date scholars are unsure how to achieve an accurate simulation of the effect (Middlebrooks & Green 1991, Kearney et al. 2010). The different cues are not only numerous (decreased overall level, air absorption, direct to reverberant ratio, familiarity of the sound source, etc... just to name a few), but their behavior over increased distance can be quite complex, not to mention the intricate interrelations between them. Coming from a musical motivation, the problem with many studies is that their reported success is achieved in controlled environments, so that in an anechoic chamber a sheer amplitude change is perceived as a change in distance. But in a concert setting these results are irrelevant because they barely to never work. For this reason, the following section tries to agglomerate all possibilities in the hope to achieve a perceptual, stable improvement in the future. To start this discussion, it should be said from the start, that source familiarity is a strong factor for the appropriate perception of distance (Zahorik 2002). A composer of natural soundscapes should, therefore, have lesser problems with the correct estimation and localization of sources than a practitioner

of acousmatic music.

The cue easiest to handle is the loudness of the first wavefront. It is governed by the inverse-square law, which describes the remaining energy of a source in relation to its distance to the receiver. Assuming the sound source is floating in free space, its energy spreads in all directions, forming a sphere. The initial energy W is then spread over the area $A = 4\pi r$ of the ever growing sphere, with r being the traversed distance. Therefore, the intensity of the signal arriving at the listener can be described as:

$$I = \frac{W}{A} = \frac{W}{4\pi r}$$

Concert halls, performance spaces or any real space that humans usually experience are hardly ever floating in free space. Usually there is at least a floor present, if not proximate walls, against which the source is standing. Depending on the amount of boundaries present, the above equation has to be corrected as the sound intensity loss decreases. This is done with a simple multiplication factor (Howard & Angus 2009):

$$I = \frac{QW}{4\pi r}$$

Q is called directivity of the source, and describes the amount of extra reflections, and effectively amplification, the source experiences by not being able to fully expand into all directions. If, for example, the sound is split into a hemisphere due to the presence of a floor, a directivity of 2 is to be applied. As a rule of thumb, a source increases by 3dB with every added boundary (Howard & Angus 2009).

Sound is a phenomenon that traverses through matter, usually air. Every material dampens specific frequencies, which is the basic principle why acoustic instruments produce harmonic tones! For air, the absorption is dependent on four factors: the frequency, the temperature, relative humidity and the atmospheric pressure. This effect is usually not noticeable in enclosed rooms and is said to take effect as of about 15m (Zahorik 2002). The full absorption formula can be found at Sengpiel (n.d.)¹⁰, but the usual practice is to approximate the dissipation with a simple high shelf or low-pass filter (Hollerweger 2006).

Furthermore, "[a] cue that serves to specify absolute distance, even on the first presentation of a stimulus, is termed *reverberation*" (Loomis et al. 1999) Truthfully, reverberation is most likely the most stable and widely used cue to denote a sense of distance. It is a common tool in contemporary audio production techniques and essential for any notion of realistic spaces. Physically simulated reverberation can give us a sense of room size and possibly even dimensions. The most common practice is controlling the amount of direct to reverberant ratio. It has been shown that the "[...] perceived distance increases rapidly with the number and amplitude of the reflections" (Kearney et al. 2010) An improvement would also be the correct reproduction of the direction of early reflections from the perspective of the listening position. As already

¹⁰The direct link to the formula is <http://www.sengpielaudio.com/AirdampingFormula.htm>. On another page on the same site, one can find an air absorption calculator and further information at <http://www.sengpielaudio.com/calculator-air.htm>

mentioned when discussing the Haas effect in section 3.1.1, early reflections have been shown to play a significant role in the perception of distance: "As the sound source approaches the listener, the reflection will travel much further than the primary and will be relatively weaker due to the inverse first power law. As the source moves further away the reflection distance approaches the primary distance and the strength of the reflection approaches the strength of the primary." (Brungart 1993)

But when getting into details, the direction of source in relation to the listener becomes a major logistical problem, as the direct to reverberant ratio has to change considerably in each ear: "For instance, when a source nears the head along the interaural axis, the floor/ceiling echo no longer dominates the reverberation and the reverberation level varies less with source location [...] For sources in the median plane, the left and right ears receive approximately the same reverberant energy, but the energy level varies with source distance. Similar results are obtained for the energy reaching the right ear when the source is positioned at 90° azimuth. However, when the source is positioned laterally, the reverberation at the far ear is essentially independent of source distance." (Shinn-Cunningham 2000) Nevertheless, one low-cost solution could be to apply a virtual ghost source below every actual source by default that varies in delay, simulating an early reflection in relation to the distance. The problem with this suggestion is that its implication has not yet been studied. One effect might be that the perception floor always present will persist (although this is usually the case in a natural environment), which might not be desired by the composer at all times. In total, reverberation is a very characteristic timbre that could very well play a large part in a spatial composition. It is therefore extremely necessary to leave this parameter in the hands of the composer for him to maintain the highest level of control. A search for alternatives to approach the problem of distance perception without the use of reverberation was undertaken for this reason. Details can be found in the first experiment below.

A rather unusual parameter is acoustic tau, "[...] a computed variable associated with the increasing sound level of any source that is being approached." (Loomis et al. 1999) The acoustic tau is a temporal measure, specifying the time to collision based on the source's distance, no matter if this source is approaching, receding or stationary. It was used in experiments for virtual reality simulations, as a hands-on control mechanism for test subjects to adjust as they see fit. Originally sought out by ecological psychologists to explain models of reaction times, its practicality and real-world relation was heavily questioned by (Guski 1992). Back in the field of virtual reality, acoustic tau also did not seem to provide considerable improvements in observers' estimations of source distance (Loomis et al. 1999). Sure enough, the variable seems to disappear in literature not soon thereafter.

Distance is usually best perceived when given a chance to relate to two positions of the same source with one another. This is especially true with unknown, synthesized sounds, where the timbre is unknown to the listener. In this regard, the best impression of distance is created when a continuous trajectory is created, employing changes in dynamics and spectrum. Sounds that approach and/or retreat from listener demonstrate a phenomenon that is always best described with the pitch changes of a passing ambulance: the *Doppler effect* (Chowning 1971).

The Doppler effect is essential, since it is naturally inherent in every moving source that shows any partial vector along a distance change from the listener. The frequency difference can be calculated by:

$$\Delta f = f_0 \frac{c \pm v_{observer}}{c \pm v_{source}} - f_0$$

The Doppler effect does alter the pitch, on the one hand, but on the other, if audience member is not familiar with the sound itself, the Doppler effect can be be unnoticed as it might be attributed to being part of the (synthetic) sound. This, for one, should be the case when moving the sound source in distance at constant speeds. It's typical effect, as known with the passing ambulance, is usually only perceivable when it actually *passes* the listener, giving a direct comparison of approaching and receding timbre. Of course, the comparison to the original timbre is only possible if we are familiar with it. Our conception of an unknown ambulance siren is an estimate between the two versions, the approaching and receding sound. On another note, by using very high speeds and oscillations in distance, the Doppler effect could also be used in other creative ways.

Finally, [Speigle & Loomis \(1993\)](#) and [Waller \(1999\)](#) both speak of one perception type that is simply impossible to recreate in psychoacoustics-based spatialization systems¹¹: "When an observer translates through space, additional information about source distance becomes available under the assumption that the source is stationary. One of these is the absolute motion parallax, the changing direction of any source that is initially off to one side" ([Loomis et al. 1999](#)) The effect is analogous with visual parallax, which describes the displacement of an object. This can best be visualized when imagining two objects of different distances to the observer. While moving the head past the direct line connecting them, the foreground object seems to move past the one in the background, making the one in the back appear on the other side of the one in the front. This parallax is also present in sound sources. My hypothesis is, that its importance is similar to the problem of externalizing sound perceived from headphones. When moving the head in relation to the stationary loudspeaker, the perceived parallax will denote the sound coming from this loudspeaker, no matter which other cues it might display to simulate an illusion of distance. Especially when the sound of the source is unknown, the timbral alterations done to the sound in order to have it appear distantly will be assumed to be part of the actual timbre. Investigations will have to be made as to how far minimal head movement can trigger this cue. First estimations are that its sensitivity is not as high as it is the case with the externalization problem with headphones, although it does improve distance perception ([Speigle & Loomis 1993](#)).

In section 3.1.1, possibilities of localization refinement on the periphery of the surrounding aural sphere were discussed. A learning process can furthermore be observed in subjects during distance perception:

"Another effect of reverberation in distance perception is that the accuracy of

¹¹I will refer to psychoacoustics-based spatialization systems to technologies that make use of amplitude panning, ranging from stereo listening to VBAP, but perhaps even low-order ambisonics fall under this category. Their main characteristic is the aural illusion they create, as opposed to a physical recreation, which is the underlying principle of WFS and high-order ambisonics (if not also low-order ones).

distance judgments improves with experience, often over a course of hours and days of practice. The exact details of how this effect occurs are not clear. However, it is known that the learning is general in the sense that performance improves even if the listeners' position in the room changes during the experiment. Still, how exactly this learning occurs, what are the critical characteristics necessary for it to occur, and what is the time course of this learning, is not well understood. [...] [I]t can be hypothesized that a process similar to memory consolidation is observed that leads to improved performance and that requires a long break between the measurement sessions." (Kopčo et al. 2004)

While Kopčo et al. relates his quote to the improvement of experimental data, this learning process is highly interesting in a greater overall window of a human development. It shows that if consciously perceived, the localization in distance can be improved. It also leaves room for the hypothesis that there is still much potential of aural improvement in the general public, and, reiterating the fact that we are so heavily visually oriented beings, our hearing might actually degenerate in the long run.¹² Brungart (1993) also cites another study, which investigated the distance estimation of either 15cm or 1m in both adults and six month old infants, based on intensity cues. Although, based on other findings that a change in amplitude alone does not contribute to the impression of distance alone, the study did show a considerable superiority of the adults in judging the distances compared to the infants. Another point of critique is that the study could be broken down into loudness estimation alone, ignoring any sort of distance sensation. Nevertheless, a learning curve for distance estimation is given and needs to be acquired and developed through multi-modal experiences (see section 3.1.1), as is later written about all perception in general in Handel (2006).

An argument for the development and refinement of distance cues is the strong variance in subject's impression of distance: "These results suggest that the cue-weighting process is flexible, and able to adapt to individual distance cues that vary as a result of source properties and environmental conditions [...] The manner in which listeners process, or perceptually weigh two principal distance cues intensity and direct-to-reverberant ratio varies substantially across the same stimulus conditions." (Zahorik 2002) It shows that every subject has a conception of how to interpret the cues, suggesting individual developments. One clear trend observed in many studies is the general underestimation of distance cues (Zahorik 2002, Kearney et al. 2010). This was not necessarily the case for real world examples, most likely due to the availability of a complete spatial picture: "These results showing that far physical distances are underestimated and near distances are overestimated are consistent with an influence of a perceptual organization factor known as specific distance tendency. This factor describes a tendency towards a specific 'default' value of perceived distance under conditions where all distance cues are removed." (Zahorik 2002)

In the following, two experiments relating to a possible improvement in distance percep-

¹²This phenomena was not observed nor even studied. It could very well be, that auditory evolution is resting in an equilibrium, perceiving enough to survive, but leaving the details to our eyes to decide.

tion in mimetic spatialization systems are reported on. The goals and main motivation for the following experiments were of practical nature. It was not so much an investigation into the complex psychoacoustic perceptions of the human nature, but more an investigation into possible tools for the spatial composer, for an increased level of control over distance parameters. Therefore, the results are not so much quantitatively but qualitatively documented. Furthermore, they exemplify how the knowledge of psychoacoustics can aid in the spatial compositional process.

Experiment I: Assessing the impact of linear signal expansion

An usually highly stable factor for an illusion of distance is a simple change in the direct to reverberant ratio (Shinn-Cunningham 2000). The following experiment's intent, though, is to create an impression of distance without relying on this method. This motivation stems from the desire to detach the reverberation algorithms from the distance algorithms within the spatialization mechanism. The reason for this is the fact that reverberation and the space it represents can be very characteristic and a compositional tool in itself. It should, therefore, stay in the hands of the composer, the user, and ideally not be hard-coded into any spatialization code. Other than in anechoic chambers, a form of reverberation always exists, be it even the reflections off the ground in free field.

At the time of this thesis, the VBAP algorithm in the studio at Barcelona Media only incorporated an overall level change with distance according to the inverse-square law, with the doppler shift being under development, but not available at the time. This cue alone was not enough to create any slightly realistic impression of distance. This fact was also assessed in the experiments by Vartanyan & Andreeva (2007), although it was found that the approach was more distinctly perceived in relation to the withdrawal of a sound source. Even theoretically speaking, there would be no separation between a reduction of volume and an increase in distance, which is undesirable within my own compositional approach. From my own experience using WFS systems, I knew it was possible to create realistic distance impressions without the use of artificial reverb. Hence grew the interest to seek for new cues one could implement for a distance simulation in mimetic spatialization systems.

The inspiration came from the introduction in Chowning (1971), in which he speaks of a cue for distance perception being "[...] the loss of low intensity frequency components of a sound with increasing distance from the listener". Chowning never goes into detail on this point, and the sentence above is the only mention of the effect in his paper. While it acoustically has no foundation, the possibility that it could be a valid psychoacoustic cue was given. When one thinks of a voice far away, so that it is barely comprehensible, one will perceive the necessary components to discern the message. On the contrary, the same voice up close to the ear bears much more detail in every little noise, such as the smacking of the lips and tongue, or the air hissing through the teeth. A vague basis of plausibility was given and the statement was interpreted as an expander, which was further implied in discussions with sound engineers. For

this first experiment, a first single-band, linear expander was implemented in the following way:

$$\min(0, (s + 1)(d \cdot RMS - d - 1)^m)g$$

The normalized RMS¹³, being the base analytic reference as to how far the current window should be reduced, is scaled by the distance d . For $m = 1$ the equation is linear and the expansion ratio increases with increasing distance d into infinity. In the experiment, the main parameter of interest was the exponent m , to determine the weight of the effect in relation to the inverse-square law, as this idea was very adventurous. Another scaling value s was used to tamper with the equation¹⁴. Finally the gain g is the overall level reduction imposed by the inverse-square law of the regular distance algorithm.

In total there were 10 subjects, all considered expert listeners as being students of the Sound and Music Computing master program. For each subject, an automation would make a sound of a drum loop disappear to a distance of 10m or 20m and return. In each turn the weighting m would be increased. The subjects were to mark the turn at which they felt the source had moved either 10m or 20m. In the end, no data was collected as the experiment failed. None of the subjects had a true impression of the source disappearing into the distance. The inverse-square law alone resulted in a minor to no distance impression (in accordance with the findings of [Vartanyan & Andreeva \(2007\)](#)), and, while the added expansion ratio did not impair this impression, it did not seem to improve it either – and it wasn't until the expander turned into a gate that the subjects noted a difference (which was not related to distance, of course).

It can be concluded that single-band, linear compression very likely has no impact on the perception of distance. Furthermore, the experiment itself was based on vague assumptions and the formula is based on a simple linear expansion model with added tampering. It can also be said, since distance cues show a complex interrelation with each other that assessing the inverse-square law and this expansion experiment alone is not sufficient for a realistic distance impression. Future investigations could implement a multi-band expander, as this really attributes to the correct interpretation of [Chowning's](#) statement and is also closer to the psychoacoustic thought experiment. Considering WFS, reasons for the realistic impression of distance is, for one, the inherent doppler effect when moving a source away and towards the listener¹⁵, but, more importantly, WFS displays motion parallax, an effect which is impossible to simulate in mimetic speaker systems and a irreplaceable effect in estimating distances in everyday life (see discussion of motion parallax above). Furthermore, through experiments with WFS and distance perception, researchers came to think about the importance of the

¹³RMS stands for root-mean-square, and computes the average over a given set. In signal processing, it is a standard measure for power within a given window.

¹⁴This was a quick&dirty method to get a grip of the [average~]-object in MaxMSP and keep the values within reasonable levels.

¹⁵When testing a WFS system implemented by Darragh Pigott at the University of Limerick in 2009, the impression of distance was convincingly realistic. Surprisingly, he wrote in a recent conversation via email: "I had no reverb or doppler programmed in the time you heard the demo. The max delay object does not interpolate between delay times, so no doppler.", which means, that the strong distance impression arose even without the help of a doppler effect!

wavefront curvature. Although not proven psychoacoustically, experiments suggest that, due to the physical reconstruction of the wavefront in WFS, that the curvature in relation to the distance of the source might be perceivable (Bates 2009). Both effects of motion parallax and wavefront curvature are two cues, that a stereophonic system will not be able to simulate. It is, therefore, very likely that, in terms of distance simulation, WFS will remain far superior.

Experiment II: Panning audio into the center of the listening space

A further limitation in using distance in the compositional process in mimetic spatialization systems is the inability to create dramatic gestures by panning the audio into the center of the audience. Proximate distance is usually considered different from regular, far distance, as at least two effects come into play that would otherwise not be considered. The first is called the *proximity effect* which is said to be attributed to the curvature change of the wavefront (Hollerweger 2006, Bates 2009), but also to the equal-loudness contours (Hollerweger 2006). Due to decreasing distance from the listener, a source inevitably becomes louder due to the inverse-square law, which flattens the equal-loudness contour, effectively applying a bass boost to the incoming signal. This effect was approximated using a bass shelf with a cut-off frequency and drop-off similar to that of the equal-loudness contour.

The second effect is the sudden possibility to use the ILD as a cue for distance estimation as of a distance of about $d < 1m$ (Shinn-Cunningham 2000, Kerber et al. 2004, Hollerweger 2006). This effect stems from the cause that when the source approaches the head, the loudness of the direct wave front increases, again, according to the inverse-square law, while the reflections remain about the same. It is reported that for distances between 87.5cm and 17.5cm the loudness may differ by as much as 20 dB (Zahorik 2002). To approximate this effect, the levels in the far ear had to be reduced compared to the increase in the near ear. This, of course, had to be done without the use of headphones and, if possible, keeping room acoustics in mind. One possible, but not viable, solution is mentioned in Stefani & Lauke (2010), who use small speakers underneath the seats of the audience to create a sound within the personal range of the listener. Furthermore, the paper only mentions the possibility as a kind of effect, and no psychoacoustic study and compositional effectiveness is demonstrated.

The possibility studied in this experiment was the effect of a level drop-off through the use of phase inversion. The two sources were panned hard left and hard right. One of the sources played the source sound in regular phase while the other source was given the same sound in opposite phase (see fig. 3.2). This resulted in a total cancelation across all frequencies at the exact center between both speakers, at which point the center of the head was. By figuring the standard width of the head to be roughly 18cm, a time delay of $\frac{0.09m}{340m/s} \approx 0.26ms$ was applied to the phase-inverted speaker to shift the cancelation maximum of the two signals towards the far ear (the ear, closer to the phase-inverted speaker, see fig. 3.2). At a sampling rate of 44.100Hz, this was equivalent to 23 samples in delay. Finally, by applying a gain normalized from zero to one to the phase-inverted signal, one could control the distance between the real speaker and the listener. By gradually increasing the gain in the phase-inverted speaker, the level in

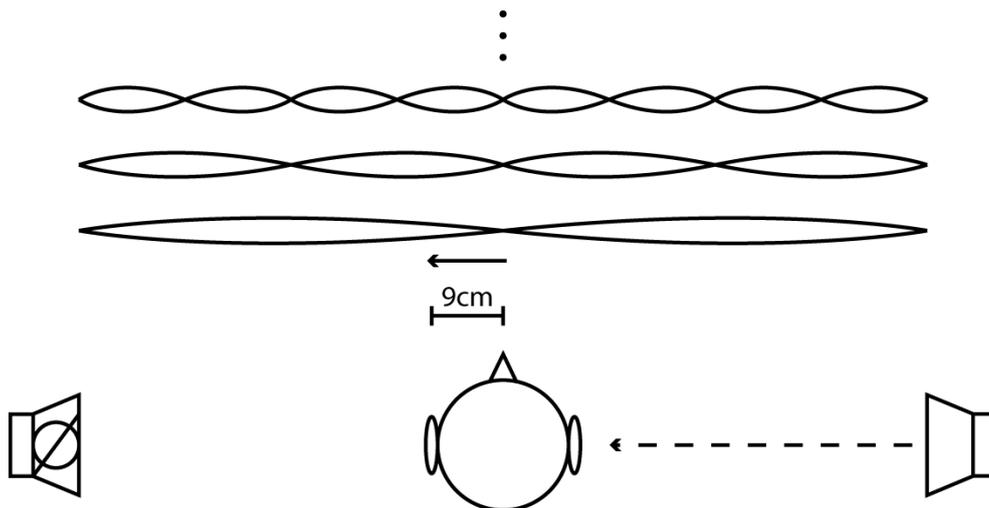


Figure 3.2: Illustration of the center panning experiment. Two speakers exactly opposite of each other, with one playing the exact same sound as the other, only in opposite phase. The top pressure curves visualize the effect in an harmonic series of a possible sound. The arrow denotes the shift of the pressure minimum if the signal is delayed by half the diameter of the head.

the far ear would effectively be decreased, resulting in an impression of increased ILD and the impression of an approaching source.

Test subjects consisted of both 10 students from the Sound and Music Computing master’s program and 4 employees at Barcelona Media. All are considered expert listeners. The demo consisted of a female voice singing a short melody on repeat, with the phase inverted signal slowly being added and removed by simply by using its gain. Each subject was asked to sit in the center between both sources. To most subjects, nothing was mentioned as to what was going to happen and the only thing asked of them was to talk about their impressions after the short demo was over. Unfortunately, test subjects from Barcelona Media knew the matter of the test and, therefore, can’t be considered completely unknowing. Nevertheless, *all* subjects reported an impression of the sound moving towards them and receding again. 11 subjects reported the sound as almost being inside their head, while 3 people reported a sensation of the sound ending up overhead. The same experiment was repeated, only this time both signals were in phase. Although the difference was difficult for most to pin down, every test subject reported the effect as being similar. Nevertheless, only 2 subjects reported an improvement in perceiving the sound approaching. Most subjects noted, while the sound was centered, it felt somewhat more enveloped and around the head; harder to localize than it was the case with a phase inverted signal, which gave more of an impression of being right next to the ear.

A second demo tried to pan the source sound from the front into the center of the listening space. Due to the nature of the effect, panning sound with opposite speakers both in front and behind the listener does not work. The phase needs to be canceled not at the nose but at one of the ears. The failure of this technique was verified during post-discussion with some of the subjects, but can also be found in [Shinn-Cunningham \(2000\)](#), where it is shown that ILDs

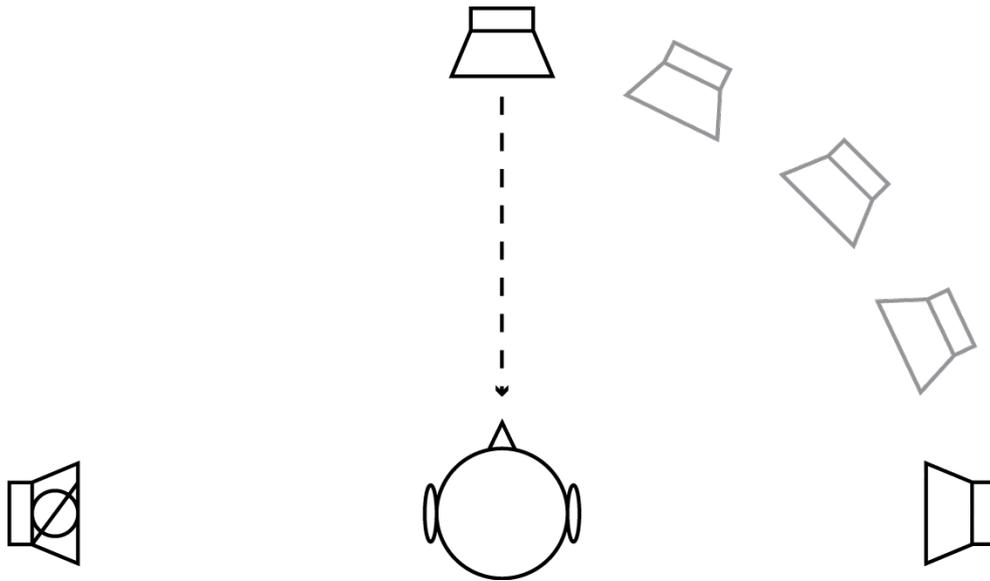


Figure 3.3: Illustration of the center panning experiment, this time panning sound from the front into the center. The phase inverted source is kept stationary, with the regular gain change as it was applied in the first experiment in fig. 3.2. The phase corrected source, though, was panned from the front around the listener as the gain was increased.

act best as a cue for distance if coming from 90° . As the approaching source varies its angle of action towards the front, the effect goes towards zero. Therefore, the moving sound source needed to approach the listener by moving by his side. A method was tried as illustrated in figure 3.3: While the phase inverted source was kept stationary, the phase corrected source was panned in a quarter circle around the audience. The panning was done in accordance with the gain change of the phase inverted signal, so that both maximum gain was at the phase inverted source while reaching the extreme lateral position of the phase corrected source.

Results show that most subjects heard the source approaching, but generally it would move along a curvy path. The curved impression probably resulted from inconsistencies, making the precedence effect assign the source coming either from the panned source or the stationary one. Also, the source would not start approaching the listener, but would generally curve about 20° before starting to move into the center. Several methods were tried to calibrate the source movement and the gain increase for a straighter path through the middle. Applying a square root to boost the gain increase in the beginning of the demonstration helped stabilize the path for some. But the effect seemed to vary much between each subject, so that a consistent stable image could not be created. Other methods that worsened the panning quality were, for example, also panning with a phase inverted signal from the center. This caused the impression of jumping from left to right even more.

In conclusion, one could say that both methods produced very pleasing results, as they achieved the desired effects. While the panning from the front towards the center demonstrated some difficulties, the panning from the far side of the listener has proven very stable. This stability is only given, though, within a very small sweet spot and tiny audience (probably no

more than one listener). Listeners that were too close to the phase inverted speaker heard the signal coming from the unintended direction and listeners too far from to the original signal, hardly heard the effect. One method to increase the sweet spot is a limit in gain increase on the phase inverted signal. For an impression of the source to reach the listener just outside the ear no full amplitude is needed in the phase inverted signal. This dampens the precedence effect and widens the sweet spot variably, according to the limit in gain. It should be noted that shadowing of different audience heads can be a big problem, also limiting the audience size. Nevertheless, no other method to pan a sound into the center is really known for mimetic systems, and the only other technologies known to do this are Ambisonics, within limits, and WFS with seemingly impressive results. Therefore, the results can be seen with value, in that respect. It is not widely applicable, but it works very convincingly within its limits.

3.1.3 Ecological Psychoacoustics

”Human ears can neither be closed nor moved around independently from the head, and they receive auditory arrays from all directions: Focal hearing (listening) can be accomplished only by selecting partial information from the acoustic array.”
(Guski 1992)

Traditionally, the act of music listening neglected spatial parameters as a viable carrier of musical meaning. Most concerts were, and still are, front view oriented and concert halls are designed to mush the complete ensemble sounds together to one timbre mass. The *listening* had only to concentrate on one fixed point – the front of the *acoustic array* – and the engulfment was a separate, almost unconscious, garnishing sensation. When spatial music comes into play, the language of the acoustic array explodes in expression, and the listener of spatial acousmatic music suddenly has to scan the space around him. This is not just a new impression, but also a new effort, which composers such as Schafer and Smalley deliberately ask for from the listener. The auditory involvement of the, usually passive, audience becomes a prime factor in spatial music, and their ability to follow multiple stimuli in physical space in a new act in music to listening. While the localization of auditory stimuli is well represented and a large field of research in scientific writing, ecological research has received less attention and, although it forms the basis for localization cue extraction, is still only partially understood (Darwin 2008). What this exactly implies will be introduced here, as this final section on psychoacoustic implications will scratch the surface of the auditory mechanisms that ”make sense of it all”.

So far, the auditory world was considered to consist of static sources in any location around the listener, with the exception of perhaps peaking at sources moving into the distance. Time and level differences between both ears are neurologically compared and the frequency resolution of the cochlea gives the brain the chance to map spectral composites to patterns that relate to different positions and distances in space. But this is only half the truth. The localization models consider an isolated part of the frequency spectrum and would relate to real world situations only if the brain would receive a single isolated source. Instead, as the quote above states, our ears are constantly bombarded with all sorts of sounds. Multiple sources,

distorted reflections, background noise, etc; it all arrives at the brain in one agglomeration of frequencies – “[...] the single pressure wave that is the sum of the pressure waves coming from all the individual sound sources [...]” (Bregman 2004) For the brain to even determine the spatial cues of one single sound, it needs to decide which frequential parts relate to their respective source – despite its constant timbral change – using a set of neurons dedicated to receiving the neural signal as a function of frequency, space and time (Handel 2006). It is important to note, that these neurons do not act on isolated parameters (for example, on frequency and time only) and spatial hearing is an inherent feature of the incoming neural stimuli from the physiological part of the auditory complex. This thesis will *not* go into the neural details of the auditory pathway, which can be read in Cheveigné (2003), Handel (2006), among possibly others. Instead, we will head on to higher level of psychology and look at the most prominent theory in this field, labeled *Auditory Scene Analysis* (Bregman 1990)¹⁶.

One prominent example of how we filter our surrounding world is the so called the *cocktail effect* (Arons 2001). The effect is generally attributed to the field of speech recognition and takes its name from our ability to focus in on a particular conversation during a party, while many other voices are simultaneously sounding in the same locality. It is an intuitive ability and we use it, mainly unconsciously, but continuously as it enables us to focus on a specific sound source of interest, filtering out the rest of the “noise” around. Typically the sound source of interest lies in our frontal area of vision as we tend to look at what we want to listen to. But we are also very well able to follow a conversation around the corner, in another room or just behind us. The attention towards multiple sources, though, is debatable. While Zelli (2001) mentions the listeners ability to concentrate on multiple sources at the same time, Arons (2001) cites studies that show that multiple sources compete for the attention of the listener and simultaneous attention is only possible if the information rate is in accordance to the listeners ability to rapidly switch from one to another. Therefore, two messages with only little information content can be dealt with simultaneously.

In the early 1950s, this effect was explained with mainly vocal properties: different speaking voices, means speeds and pitches, different accents, even lip-reading, gestures – but it also included the notion that the cocktail effect selected areas based on location (Arons 2001). In 1990, Bregman then formulated a general theory around the perception and segregation of *auditory streams*. His theory was based on the principles of Gestalt theory (Koffka 1955). Gestalt theory lies within the psychology of perception, mainly the visual domain, with the word “Gestalt” being German for shape. It makes the effort to holistically describe the brain as it distinguishes objects from one another, wherein gestaltists abstract the human perception to several basic principles. Gestalt theory is often criticized for being mere observant and not predictive, but nevertheless it did not fail to inspire and survive for over 100 years. Auditory Scene Analysis’ essence builds on the five founding principles of Gestalt theory around which the theory of grouping and segregation, separating the figure from the ground, are formed (Arons 2001):

¹⁶Alternatives to Auditory Scene Analysis are introduced in Macpherson (1995), Arons (2001), Cheveigné (2003)

Similarity: elements that are similar in physical attributes tend to be grouped

Proximity: elements that are close together in space or time tend to be grouped

Continuity: elements that appear to follow in the same direction tend to be grouped

Common Fate: elements that appear to move together tend to be grouped

Symmetry & Closure: elements that form symmetrical and enclosed objects tend to be grouped

In classical Gestalt theory of the Berlin school, the brain is thought of as self-organizing and the principles of the theory were described as innate mental laws, forcing the brain to see shapes. The perception of these groups was competitive. Bregman defines a perceptual distance d , that describes a weighted distance between several comparative auditory dimensions, such as frequency or time. Depending on the distance between two events in relation to other surrounding events, these events might group with each other, or with other the other elements, leaving one separate from the rest. In Gestalt theory, a situation where several interpretations of a scene are on the verge of fighting for the attention of the observer (listener) is termed *Multistability*, in which the perception seems to flip seemingly uncontrollably between both possibilities. In Auditory Scene Analysis, this effect causes the grouping criterion to narrow down and the perception falls back into a stable state (Bregman 2004).

Segregation is caused through contrast. Two objects separate one from another not from their relation to each other, but in their relation to their background. If two sheets of paper have differing shades of gray are on the same table and the darker one, if lit up to measurably match each other, they would perceptibly still differ. This is because the table is also lit up, and it is the contrast with its background to which each paper is measured. "[T]he important auditory properties are those that signify changes in the relative vibration patterning that characterizes objects. [...] The change must be predictable and that predictability must be able to be derived by the observer." (Handel 2006) This means, auditory stimuli are segregated when patterns are broken from the background noise. While it marks a fundamental function to discern sounds from the environment, this also becomes especially interesting when dealing with textural composition in section 3.3.3.

Moreover, Bregman defines two types of auditory grouping phenomena: primitive segregation and *schemas*. Primitive segregation describes our natural abilities to segregate sounds in the environment from one another, similar to how Gestalt theory describes the urge to see patterns: "I find it impossible to put the noise back into the tone to hear an integrated sound or to put the snowflakes back onto the object. I believe that the auditory and visual segregation is obligatory and represents the first step in achieving the objects of perceiving. Bregman has termed this process primitive segregation." (Handel 2006) Segregation causes individual events to be split apart and grouped into auditory streams. "Streams are a way of putting sensory information together." (Arons 2001)

"To recognize the unique timbre of the voice we have to isolate the frequency components that are responsible for it from others that are present at the same time. A wrong choice of frequency components would change the perceived timbre of the

voice. The fact that we can usually recognize the timbre implies that we regularly choose the right components in different contexts.” (Bregman 1990)

Spatial cues are a major component in the process of primitive segregation and include both spatial location and spatial continuity among other cues (Arons 2001). More interestingly, there seems to be a segregation time constant, which defines a number of seconds before a stream is established, and also before it is dismissed after the last event of the stream. Among the other cues are harmonic, pitch and loudness cues, for example. Grouping in Auditory Scene Analysis creates auditory streams that compete with one another sequentially (temporal) or when sounding simultaneously (frequency). The above cues used for segregation seem to aid as indicators for grouping in the process.

Finally, *schemas* come into play where primitive segregation fails. (Bregman 1990) observed, that, despite the principles setup behind primitive segregation, subjects were still able to filter different streams of speech produced with a synthetic voice, that should theoretically be fused. He went on to define an additional model of learning, a way of discerning learned patterns from previous events that involved attention. He places this on a higher level of central processing. For example, studies have shown that infants are not able to connect speech to a cohesive stream, because of when the infant repeated, what the mother said, specific sounds were missing, which it must have segregated (Arons 2001). It is not until it developed a schema, a blueprint, that it would finally connect the dots.

The schema is said to surpass the primitive segregation and directly pick exactly what it needs out of the total spectrum. ”There is also evidence that a scene that has been segregated by primitive processes can be regrouped by schemas.” (Arons 2001) The schema is therefore superior to the primitive segregation process. This insight leads to significant and popular misconception in spatial audio when dispersing frequency components of a source. During investigations in spatial sound synthesis undergone in chapter 4, a source sound was broken up in up to 64 frequency bands, each dispersed in a dome of speakers as its own virtual source. The effect, at first, could not have been more disappointing: instead of having the sound completely disjointed into separate auditory streams, composed of its components, one would hear the original source sound, with a slightly added quality envelopment. Even though separate frequency components were all sounding distinctly from different points on the hemisphere, the brain would be able to fuse all the components into one cohesive image. This effect is also documented in a just recently published paper by Kendall & Cabrera (2011)

The implications of Auditory Scene Analysis return to the beginning of this section giving us a chance to finally discuss the motion of sound sources. While some alternative theories attribute auditory motion detection directly to dynamic aspects of localization cues, the Auditory Scene Analysis approach lets us discuss a relatively accepted approach called the *snapshot theory* (Middlebrooks & Green 1991). It describes motion in terms of aural ”snapshots” that the brain then compares against one another. Considering the coherence with the principles of Gestalt theory in audio, it is very well possible to have events of audio of the same type but in different places in space to create a mental trajectory. The closer these events are in time and

space, the smoother this trajectory becomes. At the same time, there exist minimum audible movement angles, and they vary considerably in terms of source location and source type. Our frontal area, for example, has a higher angular resolution than any other position around us. The movement of broadband stimuli, noise, is generally more easily detectable than a tonal sound. Moreover, the speed of the source will decrease the sense of motion at both ends of the scale. The upper limits of rotational motion were studied by Féron et al. (2009) to be around 2.5 rotations per second, while lower boundaries exist at around 0.2 rotations. The study itself states technical problems, though, so the values should be seen as "soft" points reference.

One interesting effect that occurs from being subjected to a period of perceiving motion are *motion aftereffects* (Handel 2006). They occur due to a loss of sensitivity of neurons within auditory (also visual) neurons. They are best described with a visual analogy: when looking at a spiral turning inward for 30s and then looking at your hand, for example, the hand would seem to grow. The effect is reversed if the spiral turns the other direction. A similar effect is the rising landscape, if one first stares at a waterfall for a long period of time. Since these neurons are responsible for inter-modal processing, the effect can apply to auditory stimuli as well. Audience members during the premiere of Stockhausen's *Kontakte* reported a loss of orientation, as if they were in an antigravity room (Féron et al. 2009). The exact effects are still to be studied, but perhaps this resulted from the conflicts between the aftereffects-causing neurons and the composition pushing the perceived rotation in the other direction, effectively heightening the experience beyond the auditory stimuli. Depending on the intentions of a skilled spatial composer, these feelings could be quite impressive.

3.2 On the Notion of Space

"Finally, it is sounds and their own materials which generate, through, projections or inductions, new musical forms." (Grisey 1987)¹⁷.

In the previous section 3.1, we investigated the state-of-the-art knowledge on psychoacoustics with respect to spatial hearing, starting just after the physiological part of the auditory pathway all the way to higher level psychological processes that decode the aural scene around us. What remains is the discussion of how we actually perceive *space* itself, how we feel and acknowledge the space that we are in. Beyond the musical context, on which this thesis will try to limit itself, space as a concept is a widely discussed topic from philosophy to mathematics with blurry boundaries. What is interesting is that many theories about the perception and structure of the "fabric" that makes space are often interrelated, or even inseparably placed

¹⁷This quote is torn out of context. Originally, Grisey talks about *Spectralism*, a school of composers mainly connected with the French IRCAM institute from around 1970 onwards. Nevertheless, as we will see in the end of this section, my notion of space hails from similar motivations and motives, which this quote should foreshadow. Furthermore, I have composed *spectrally* myself (Schmele 2010) and feel influenced and inspired by this aesthetic. *Spectralist* composers sought out to define musical structures and models that would be based around the spectral analysis of sound. An entire piece could, therefore, be built around the overtone series of a single note played on a specific instrument, such as the exemplary *Les espaces acoustiques* (1974-1985) by Grisey, with the famous third part *Partiels* (1975) being written all around the analysis of a low pedal E2 on the trombone. I may reference the interested reader to my previous thesis and the references given therein (Schmele 2010).

on the same level with the notion of time. Not seldom do composers of spatial music, like Xenakis or Smalley, talk about the fusion or interchange of musical time and space in their works (Sterken 2001, Smalley 2007), while in other fields, the German philosopher Immanuel Kant, for example, opened up the door to a phenomenological view of space, giving importance to the temporal aspect in his view that spatial experience was intuitive (Harley 1994) and physicist Albert Einstein famously brought the two together into one *spacetime* (Zelli 2001).

Foremost, space is a geometric concept, a measure between two points. This is what our bodily experience teaches us. For the longest time, geometry was defined by the *euclidian space*. But in the efforts to prove euclidian space in the 19th century, new concepts of spaces that violated the euclidian concept of parallelism arose (Harley 1994), and soon many geometric spaces that would aid in reducing the complexity for specific tasks emerged. Other fields of research each formed their own concepts of space and spatial organization, such as the abolishment of absolute space in general relativity, the *Hilbert-space* as a euclidian generalization, or the abstraction into different views of density in urban planning, just to give a few examples. Looking at philosophy, phenomenologists discussed the notion of space in terms of time as a mediator of this experience. Summarizing the findings in Harley (1994), the phenomenologists saw the notion of space as a personal, egocentric perception with movement being the essential bodily experience. Humans live in spatial spheres based on distances centered around the human corporeality. Despite the mathematical or physical abstractions, for the man on the street, for the audience in the projection space, for the perception of the immediate space around oneself, the intuitive notion of euclidian space is still the most related, most useful and most valid. But this euclidian spatial notion is also said to be a false impression, a cultural distortion of space: related to the visual experience of space, Patrick Heelan states it being primarily hyperbolic (Harley 1994). In that sense, how would one construct a mathematical representation of our audible perception?

Composers and music philosophers have thought of space in several dimensions, ranging from none to multidimensional (Harley 1994). Many of these thoughts, though, circle around the projection of musical parameters into a virtual, perceptual framework. The perception of music in particular, in that case, goes beyond the three dimensions of our surroundings. If viewed within this model, the mental complexity of organizing a spatial image including the significance of physical space in the music would go beyond our three dimensional thinking and would have to be projected back into a space we can imagine. Perhaps this is one reason for the reluctance to fully exploiting the spatial parameter to its fullest extent. The inclusion of the physical space in music was, first and foremost, a method to avoid spatial masking of different compositional streams (see chapter 2). Space was seen as a delimiter to separate different groups and later spread textures apart. Later on, the spatial notion evolving around movement emerged. This concept of space most likely refers to our bodily experience, similar to the phenomenologist view of the physical experience of space. The cross-modal satisfaction comes from the movement of sounds, reversing the physical experience. And it is movement, a change through space, that best triggers our spatial consciousness. But Xenakis, for example, is said to try and work against this multi-modal bond and play with the diversity of the senses (Sterken 2001):

Speaking about his multi-layered *Polytopes*, the audience is expected to contribute actively to decode the multimedia extravaganza, in his effort to try to open the audience's mind to diversity and simultaneity. Similar things could be said about *HPSCHD* by John Cage, in which the audience is not just required to participate, but shape the piece personally by physically moving through the space. This type of spatial experience is one core concept of any installation, for example. Further down the timeline, composers think of spatial composition in actual spaces. Soundscape compositions bind external spaces onto the recording and contemporary composers like (Smalley 2007) categorize the space in different spatial hierarchies. In this, he coheres to the ideas of spatial spheres of philosopher Merleau-Ponty (Harley 1994), keeping the spatial notion within the boundaries of bodily experiences.

The necessary active involvement of the audience is not just a way to shape the audience in their listening abilities, but it is a necessary behavior to make them experience the space. Space is made conscious through the act of exploring, be it with your eyes, moving through a space – but in acousmatic music this exploration happens solely with the ears. Without this active involvement, the space itself would fly by unnoticed. Music that is primarily concerned with space, but fails to address the spatial engagement will be completely misinterpreted. Therefore, the music must necessarily invite the audience to engage with the space. Usually, one might think this is done by using multiple speakers in a concert hall setting, but, of course, there are shades of success in which the composer can find himself portraying his intentions. If a piece on its outer surface, the final part of the music-making process, the phase where sound travels through the listening space and reaches the listener, does not communicate its intentions and take care of its primary aesthetics, presumably it will eventually fail in the long run (Morgan 1980). Especially in the case of spatial music, where logistics and performance possibilities are a major hinderance, a movement of arrogant intellectuals would eventually stab and finally bury this idea. For Blesser & Salter (2007), there are two major factors that constitute inattentive listening and disinterest. First, there is the lack of information given to the audience: "[...] the sound and acoustic space may be without meaningful content for a particular listener; there is nothing being communicated." (Blesser & Salter 2007) Clearly, this depends on the audience and their ability to decode the intentions, but exposure and acceptance is a prime factor for an idea to carry fruit (Bigand & Poulin-Charronnat 2006). The fact that even most artists turned away from the audible surface that Serialism created (Morgan 1980), is a good, negative example.

The second factor, as obvious as it first may seem, might just as well be a very important notion considering the complexity of music: "[...] the listener may not be paying attention to the sound and space. Even if these are emotionally charged, you may not be engaged in focussed listening [...]". While this might be lesser the case in the actual sound in a concert setting (the quotes from Blesser & Salter (2007) are taken from an architectural context), the spatial composer needs to focus the attention on the space. Stockhausen and Xenakis did this through visual dramaturgy, by either constructing a large sphere of speakers or dispersing the orchestra throughout the audience, for example. Being a complex, uncommon notion in music, the spatial parameter needs to be obviously exploited. Cultural development will then drive the

idea forward into further spheres of complexity. For this, the state of spatial awareness that the composer has to achieve through his composition can be compared to that of *reduced listening* (Kane 2007, Chion 1994): the aesthetic of Pierre Schaeffer, likely a cause of the awareness of the radicalness of his creation, *musique concrète*, required the composers of this new approach to place the listener into a new, unexplored mode of listening that would enable him to focus on and perceive the traits of the sounds themselves. In *musique concrète*, the sound should be detached from its source, meaning all connotations and allusions to what caused the sound, any meanings, or effects should be avoided, effectively steering the attention to the inherent perceptual characteristics of the sound itself. Similarly, the spatial composer should steer the attention to the space, concentrating on the spatial perceptions and how it shapes a sound. Effectively, both spatial and "[...] reduced listening has the enormous advantage of opening up our ears and sharpening our power of listening." (Chion 1994)

Spatial hearing is a natural mechanism, more primitive than that of pitch, for example, and it differs in this with *reduced listening* in so far as Chion (1994) writes "Reduced listening is [...] hardly natural. It disrupts established habits and opens up a world of previously unimagined questions [...]" . But the problems of spatial listening lie in the mode of perception: Blesser & Salter (2007) identify three kinds of perception: *detectability*, *perceptability* and *desirability*. The first means a mere detection of auditory stimuli, sine tones, high energy noise bursts – sounds that bare no cultural, or even real-life meaning. After the Second World War, artists wanted to create an aesthetic that would break away from the past, avoiding any connotation to the horrific images that shook the whole world. Pierre Schaeffer went out to define a new language of concrete sounds and spoke exactly about a listening state that is not of emotional, nor intellectual motivation, but about the pure stimulus and the sound itself. Stockhausen, detesting the recent abhorrent past of his own nation, went against periodic rhythms that held the remains of commanding military marches and embraced the new, clinical language of electronic sine tones to build a cultural identity, free of the historical baggage. Today, *post-digital* artists, such as Ryoji Ikeda and Carsten Nicolai, make heavy use of this vocabulary (Cascone 2000). It helps define a sense of transparency, a language without cultural identity – although you might argue that this cultural identity is built as the language is further exploited, and the motivation for post-digitalists is of another nature than that of their post-war colleges. For example, as our culture identifies itself more with our technology, *Glitch* becomes its true language, coming from deep down among the dark holes of virtual reality. Coming back to space, for the most part, aural spatial perception lies within this (usually) subconscious realm of *detectability*. Our auditory system can not *not* detect a sound.¹⁸ It is constantly receptive, but it is our brain that filters out what we think we do not need (Bregman 1990).

When we bring the unconscious to light, we *perceive* sounds by partaking in *attentive listening* (Blesser & Salter 2007). Judging by the theory of *Auditory Scene Analysis* (Bregman 1990), this perception is foremost a neurological necessity, an automatic process, and the spatial position is inseparably connected to each *auditory stream*. But, on top of that, it is honed by

¹⁸If our physiology is capable of perceiving it, the brain does to. Here, I talk about the ignoring of sounds around us, which does not happen, but instead the sounds are pushed into the unconsciousness.

learned *schemas*. The primitive process may first discerns sounds and sound fragments on a coarse, basic level, but it is the *schemas* that group these fragments to a more cohesive whole. In fact, when *schemas* are learned, they operate *ahead* of the primitive cognition by filtering the sound, leaving only the remains to be analyzed. This explanation goes in accordance with [Blessner & Salter \(2007\)](#), as they claim cultural influences and personal experiences to be the cause of us forming the auditory world into a more meaningful context. In other words, we shape our own higher level aural perception of the world. We group and re-group the auditory scenes around us, localizing, separating and categorizing different streams. Our world becomes a scene of events, ranging in different degrees of information content and importance. Generally, we would then *perceive* the space within our eyesight, the auditory scene in front of us, leaving our surroundings to the merely detected.

Space, for the unperceptive listener, might be related to the transparency of sine tones, or abstinence to musical meaning that concrete sounds first portrayed, for example. Moreover, and unlike sine tones, it does not serve the purpose of a traditional view of music in terms of major-minor distinction, simplistically speaking. One needs to look at new developments in music, the noise of the Futurists, the liberation of sounds in *musique concrète*, the structural freedom in chance music, the separation of parameters of the Serialists, and, especially, the continuity of space by means of technological developments and computer technology. Physical, musical space is a logical conclusion of our Modern times. At first it seems that it is a pure sensory experience, a *l'art pour l'art*, as the parallels to the reduced listening mode postulated in *musique concrète* were drawn. But if we look more closely, likewise as mentioned earlier with the post-digital movement, space and spaces are attached with deeper cultural meaning, and its exposure through culture in history probably goes beyond the spiritual atmospheres of religious sites. Every visitor of a church stands in awe of the auditory space of a cathedral, the reverberation and the soundscape of small footsteps in the distance, the mumbling of soft prayers – the occasional camera clicking away. "In many situations, listeners may not be consciously aware of the affect induced by listening to engaging sound or spaces." ([Blessner & Salter 2007](#)) It is exactly the search for these spaces that soundscape composers seek, and the goal acousmatic composers strive for their synthesis. Through reduction of other musical parameters the audience has to come to a conclusion that it was not the sounds that moved them – it had to be the space. Auditory spatial awareness stretches from *detection* to *desirability* of sound in the auditory space. My conclusion is that the composer can steer the attention to shift the listener from the *detection* of space to the attentive mode of *perception*, but the language of "[...] high-impact, emotionally engaged listening" ([Blessner & Salter 2007](#)) can only come from a rich pool of culturally established norms – and a true musical spatial language is still to be established.

One might say, space can exist without music and is omnipresent throughout time. This is a usual state and for the space to become musical is a notion so alien, one heavily has to push oneself against all usual habits. Moreover, that person might insist that music does not have to exist in space but can come from one point. Its exploitation of pitch space and timbre space is sufficient to create musical gestures and deliver meaning. But this is only partially

true – if not completely false. Spaces (that humans can experience) always exist with sound, and sounds always exist in space. We always listen to space. This becomes apparent when entering an anechoic room. The reactions upon entering an anechoic chamber for the first time are sensations of pressure, disorientation and sudden discomfort. Some people even describe an instant feeling of nausea and even feelings of spacelessness (Blessner & Salter 2007). The impact on our awareness of space is immense. Suddenly the information that our eyes gives us is not in coherence with what our ears tell us.¹⁹ The impression can be a sense of cross-modal *Multistability*. In accordance with Bregman, though, the ears started to adjust to fall back into a state of stability, eventually giving way to the eyes.²⁰ Nevertheless, the sense of enclosing space from behind, the space can not be seen, prevails for minutes until the body finally relaxes to the given situation, if at all.

Furthermore, spaces have a characteristic tone. A bathroom sounds different from a living room or a closet. An empty house *sounds empty*, while its comfort is elevated when furniture is present. While shopping for groceries, you are constantly bombarded with background music and subliminal advertisements. A parking garage has a sense of openness, with reflections arriving much later from other levels, while churches suggest an enclosed space of spiritual experiences, where most reflections bounce from walls within our sight. Church bells delineated spaces by defining the *inside* of a community or the *outside* in the wilderness (Schafer 2007). Spaces can invoke all kinds of feelings: anxiety, tranquility, socialization, isolation, frustration, fear, boredom, aesthetic pleasure, etc. (Blessner & Salter 2007). We navigate through space and we do so using our ears. Spaces influence our behavior, or mood, or sociability. More interestingly, while Blessner & Salter (2007) speak of musical enhancement through space, they also mention the aesthetic notion of space through aural embellishments and richness. All these symbolic meanings and feelings invoked through spaces, though, are subject to change, and to the modern listener the sounds of bells do not necessarily bare any spiritual meaning anymore (Blessner & Salter 2007). Hence, a culture adapts to the spatial soundscapes and their collective meanings. Consequentially, music cannot live without its space. Today, concert halls strive for exact acoustics and people often talk about its sound and *tuning* a concert hall. An ensemble purposefully makes use of the acoustics, as the reverberation aids the acoustic performer in minute details, the combination is inseparable. These details become evident if one listens to recordings done in anechoic chambers.²¹

Musical spaces also shape musical styles and periods (Oliveros 2003, Byrne 2010). The sustained sounds and long pauses in chorals are in direct relation to large churches with long reverberation, while the more percussive styles, those intended for dances, were usually performed in intimate, rectangular ballrooms: "The composer [...] often enjoyed the advantage of

¹⁹Another interesting, but non relevant effect in anechoic chambers is the complete silence that can be experienced. With no reflections and completely detached from the outside world, the bodily functions are no longer masked. The sound of the bloodstream and nervous system can be heard if listened to carefully. This experience is said to have inspired John Cage to one of his most famous compositions *4'33"* (1952).

²⁰The visual priority, when both senses do not match is also called the *Ventriloquism Effect* (Zelli 2009).

²¹Denon released an anechoic CD in 1989 entitled "Anechoic Orchestra Music Recording", with Masahiko Enkoji directing the Osaka Philharmonic Orchestra.

creating a piece of music with a particular space, or kind of space, already in mind.” (Beranek 2004). This spatial influence was often unconscious. While the inspiration was, perhaps, of more conventional musical origin, the existing space and its notion unconsciously shaped its style (Byrne 2010).²² This is plausible, since the first concert halls built specifically for concerts did not show up until 1863 (Beranek 1992). To summarize the previous paragraphs, music has always existed in space and the spaces defined the music that it contained, to a degree. Furthermore, there is no space without sound, and since all sounds can be considered musical, hence there is no space without music. The conclusion has to force one to the notion that music essentially is, and always has been, a spatio-temporal art form.

But the notion of space in music goes even further: ”Musical space [...] is a space of relationships.” (Morgan 1980). Although Morgan (1980) does not necessarily speak of physical space in this particular sub-sectional context, he does include all possible spaces in a later context of his writing, including physical space, and brings up an important hint. In the background of the main discussion, his article specifically discusses the origin of applying the term *space* to musical parameters. Indeed, many terms used to discuss music are borrowed from spatial properties. Besides meaning the physical space, the term *space* can be used in many other musical contexts, such as tonality, also called *pitch space*, or the orchestration, *timbral space* (Einbond & Schwarz 2010), or even *notational space*, creatively using the space available on a sheet of paper. The connection pulled together by this nomenclature is of an interesting notion: ”These considerations suggest that music is apparently unthinkable without the presence of some spatial, extratemporal dimension [...]” (Morgan 1980) Space, as a concept, is projected into every musical domain that shows at least two points of reference. Starting with the theory of Hermann von Helmholtz, which states that the close relationship between space and pitch, scholars and artists alike discussed the notion of *pitch space*, being part of what constitutes *musical space*, which, in turn, is clearly distinguished from physical – or *auditory space* (Harley 1994). Harley further concludes, that this characteristic of ”intrinsic musical space [...] is perceived in the experience of music ’as heard’ in temporal motion [...]”. Soon this notion of temporal development turned into a spatialization of time itself, especially in the writings of Theodor Adorno and George Rochberg (Harley 1994), but also Christian Calon (Zelli 2009), comparing the temporal development of music to the spatiality of a painting – also demonstrating the influence of physics and *spacetime*, being ported into the musical thinking (Harley 1994).

Reverberation can be considered part of a sound, and, as such, the sound carries the space (Kendall 2010). But this notion stems from the lack of acoustic and psychoacoustic knowledge and the influence recording technologies have on today’s musical thinking. As the reverberation is, in fact, many copies of the same sound, the sound itself does not carry *the space*, but it is the space that is sounding. Instead, as illustrated in the cross-modal perceptions of *pitch space*, among other spaces, there are other notions of spatial dimensions in sound – and even alluded to

²²This thought could be taken further, that both music and architecture (and all art forms of the time, really) influenced and shaped each other, imprinted by the time that they were created in. The whole period (any period) progressed like a huge construct, eventually morphing into the next.

through the sound: "Energy modifies space or generates a new space [...] Sounds are generally source-bonded in particular, therefore, carry their space with them – they are space-bearers."²³ Sound events are intrinsically spatial, because of the physicality of the cause of this sound, which, due to our bodily experience, must have a spatial context (Kendall 2010). We apply this notion also when listening to unknown, electronically generated sounds, since, if we can not identify the source, our brain usually scans our experience for any kind of connotation (Smalley 2007). Therefore, next to their cross-modal allusions, sounds also carry a deeper, philosophical notion of imaginary spaces in them. In terms of electronic sounds, though, these thoughts have to be completely reformulated. By filtering, distorting, modifying and synthesizing sounds, the electronic vocabulary has expanded beyond anything thought of before the emergence of *musique concrète*. The most elaborate theory that brings this new universe under one sonic umbrella is the concept of *Spectromorphology* (Smalley 1997).

Spectromorphology provides a framework to describe sounds in terms of their spectral development to reach into their intrinsic features. Especially aimed at the perceptual understanding and apprehension of electronic music, Smalley separates the momentary spectrum from its temporal shaping. Furthermore, the connection of the sound origin, the technology, is not necessarily given, nor is *technological listening* necessarily part of the listening experience, as is the usually the case with acoustic music (Smalley 1997). When engaging in *reduced listening* in acousmatic music, therefore, the sounds needed to be described with a new language. Performing spatial music myself, I came to a notion of space in sounds which I later found supported in (Smalley 1997): "[...] spectromorphology is concerned with motion and growth processes, which are not exclusively or even primarily sonic phenomena: sonic motion can suggest real or imagined motions of shapes in free space." Another spatially inspired aspect of spectromorphology is the classification of *spectral spaces* which may act as a discussion basis in spatial composition, especially when looking at techniques, such as timbral spatialization (see chapter 4.1, (Normandeau 2009)).

Smalley (1997) finishes his discussion with a section on *spatiomorphology*, marking his spatial framework, which he would elaborate on further in (Smalley 2007). His intention is the description of spatial development in a piece, analogous to spectromorphology in the frequency domain. Although slight connections are hinted at between the two, they remain mostly two separate domains of composition. Perhaps he evades the discussion since the projection of a spectromorphology into spatiomorphology most likely is a private experience, a purely artistic act, hindering objective evaluation. And still, this notion of space within a sound – embedded and revealed within the temporal unfolding of its spectromorphology – are hinted at in so many writings about other quantitative and qualitative parameters of music. Considering acousmatic music specifically, Zelli (2009) interprets a quote by Christian Calan the following: "It is hence a music which serves subjective simulation through sound metamorphosis in a narrative context."

Unlike harmonic sounds, which are naturally merged to a pitch and timbre sensation, electronically generated or modified sound can exhibit much more complex spectral composi-

²³Lefebvre, as quoted in (Smalley 2007)

tions, that, in fact, the auditory stream can easily diverge into several streams, similarly and more pronounced to a formant sweep. When projecting the intrinsic characteristics to the spatio-morphology, one engages in the effort to reinterpret the microscopic intrinsic movements of a sound, and emphasize them through the act of spatialization, turning the sound inside out. One moves away from the extrinsic features (if they were present before or not) and aids the listener during reduced listening to focus exactly on those particular features of the sound.²⁴ At the same time, one motivates the listener to engage in a form of reduced listening for spatial cues. Unlike one of the dangers Smalley (1997) states (that the "[...] 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."), the background is brought to the foreground, flipping the aural Gestalt upside-down (or inside-out) by distorting it.

Not all sounds necessarily have a spatial connotation. It is usually long sounds, with a vast and wild spectromorphology that nourish the imagination. Given the ideas of spectromorphology and connecting it with the auditory scenes and Gestalt principles, a "long sound" can be seen as the development through several subsounds, if we break the stream into separate chunks.²⁵ The inter-spectromorphological relationships between each chunk then each create points in space based on this inter-modal allusion. The experience is clearly cross-modal, as the sound punctiliously performs a dance through the head. Spatializing this sound according to its apparent dance is no less an emphasis as an effect. Because of the degree of intimacy of this personal sensation, the act of spatialization is how the composer intends the listener to listen to a sound. The spatialization becomes the compositional language.

We finally reach a point, in which the sound is not just in the space, and the space is also not just in the sound: there are sounds, especially in multichannel electronic music that *are* the space. These space-sounds are not just created through illusion, with the means of placing sounds and ambiances into the listening space. This sound, perceived as a whole coherent auditory stream, but completely surrounding, also differentiates from traditionally enveloping or engulfing sounds, in so far as that the space-sound is on the barrier of being enveloping itself while simultaneously being perceived as many sub-sounds from all areas within the mass. Careful balancing hinders the ear from moving too far from perceiving the sound as a whole and prevents full recognition of its subparticles. In effect, the sounds are "much more" spatial than enveloping ambiances, as they draw the mind of the listener simultaneously into each direction, while the enveloping quality of the ambience is quickly swallowed by the subconsciousness.

Hagan (2008), having a similar notion of this type of sound, refers to the term *texture*. Originally stemming from acoustic composition to describe the combination of melodic, rhythmic and harmonic development (Morgan 1980), texture is described as a type of complex supra-spectromorphology, neglecting time by reducing the complete musical development into

²⁴I do not claim that this approach will make music objective. In the end, it is an approach like any other, and if it invokes different mental representations in the listeners mind, then that is fantastic as well, as long as the listener enjoys listening to the piece, of course.

²⁵Although Gestalt theory does state that the perception of a Gestalt is finite and irrepeatable, i.e. you can not *not* perceive the Gestalt, once brought to mind, the clear notion that coherent dots on a page that form a Gestalt are still only dots, exists.

one momental perception. In the electroacoustic context, [Smalley \(1997\)](#) and [Hagan \(2008\)](#) both refer to texture as a temporal stasis. Completing the loop, we return to the agglomerate spatial notions found in [Harley \(1994\)](#), where Susanne Langer is quoted: "[...] music makes time audible and its form and continuity sensible." While the temporal stasis stands in antagonism to this (generally accepted) musical definition, we should not forget about Adorno or Rochberg, who spatialized time, comparing the temporal act of musical composition to the spatial act of painting. In this, the piece as a whole is seen in stasis. While the spatialization of time works on music that is pointillistic²⁶ in its compositional methodology, the act of spatialization brings a new dimension to this "picture framing approach". While time is spatialized, the spatialization, in turn, is seen as a development through actual time. The analogy drifts from painting a picture to creating a movie. The temporal aspect of the music is brought back to the overall view, but it is perverted coming from other dimensions.

3.3 Composing the Space

"Let's say there are rules that come with the genre, and if you try and go against these rules you're likely to fail. That's one of the difficulties here." – Smalley²⁷

When it comes to spatial composition, the following point, made in a lecture given by Erwin Roebroeks, sums the problems of spatial composition very well: composers starting to engage with spatial composition quickly want to have sounds flying all around and over their heads. But this is not what spatial composition is all about – spatial composition is about composing the space.²⁸ Smalley (quote above) also remarks on this aspect in spatial composition. Composing the space comes with certain limitations and techniques that were established in almost 100 years of electronic spatial composition. The limitations should be clear after the previous sections 3.1 & 3.2. Considering techniques, one can say that, when playing with pitches, for example, the succession of different pitches might not conform with the norms and expectations of the current culture. They will either be perceived as harmonious, disharmonious, conclusive, coherent or disagreeing, etc... – but they will always be *perceived*. As discussed in the previous sections, the spatial perception bares many traps and misconceptions. "The rules", as Smalley puts it, are not comparable to the conventions of a cultural period, the spatial parameter lacks a sophisticated culture and exposure of its own for this comparison to be done just yet.²⁹ But, if one does not consider certain characteristics of space and its perception, the compositional intentions will be subconsciously ignored, masked, or are simply misperceived.

That said, it can be stated that the spatial compositional technique is not straight-forward. Still, many think of spatialization in terms of sounding objects flying through space – and in essence, that is what we have to deal with. This elementary object comprises the core of our

²⁶...as acoustic music is, if one concentrates on the instruments only.

²⁷As quoted in [Austin & Smalley \(2000\)](#)

²⁸The quote can not be reconstructed correctly, therefore, it is rephrased from my own memory. The lecture was held during the Sónar Festival 2011, 16th–18th of June, Barcelona, Spain.

²⁹One hope of this thesis is that it will aid in taking a step further for this culture of spatial composition to emerge and fortify

thinking, still, because the electronic composer will, at some elementary level, have to deal with audio channels, one for each speaker. One level of abstraction beyond are discrete streams of audio that are somehow distributed over the lower situated speaker channels. Only when the composer moves up in abstraction, creating a complex network for audio streams being distributed over lower levels, will he move away from the cliché flying sounds in spatial approach. Furthermore, it should be stressed that we want to look at space as a compositional language. In the past, space has often been seen as a way to clarify dense musical textures or organized in layers that zoom in and out of one another. As Boulez puts it: "Space, topography is like instrumentation: it clarifies something else. For instance, one can play a polyphonic piece on the piano without timbral contrasts and polyphony will be there, but this polyphony will be much more distinct if different groups will play different parts. Space works exactly the same way." (Harley 1994) Although his approach is of lesser interest, one important aspect of this statement can be extracted: spatial composition is comparable to timbral composition. It defies the traditional linear organization that the one dimensional conception of pitch and dynamics might depict.

Nevertheless, these techniques should also not be ignored just because they might be cliché. What makes them cliché is how they are used and a skilled composer might project them into new hemispheres of complexity. Take this quote from Stockhausen, for example:

"Wenn ich den Raum topologisch orte und entsprechend strukturell komponiere, um Klänge im Raum sich in ganz bestimmten Ordnungen bewegen zu lassen, so kann das zu einer unerhört lebendigen Musik führen."³⁰ (Stockhausen 1964)

Although this structural approach might be belittled from today's viewpoints³¹, it is still a "romantic" idea and people maintain a hope that precise gestures will somehow develop into an objective interpretation. A recent study undergone by Marentakis et al. (2008), for example, investigated the success of the correct identification of gestures in a concert hall setting. The study will be looked at in more detail in section 3.3.1, but it shall be anticipated that the results were generally very poor. Knowing about auditory distance perception, perhaps this is not all that surprising, and considering the amount of knowledge out there, maybe even a bit naive. But distance is a desired factor, for one, to produce sophisticated gestures and, more importantly, the composer can perfectly play with different zones of proximity from the audience. (Smalley 2007) cites Edward Hall who distinguishes public from personal space, that are only an arm-length away; an intimate space, relating to contact. Notwithstanding, despite good control over distance when working with reverberation, the distance parameter should be dealt with care, as it needs more investigation (see chapter 3.1.2). That said, there are valid reasons for using trajectories (especially when using them as elementary building blocks), and if done with the right intentions and approaches, one can overcome both critique points of being misperceived and too theatrical. One popular technique, for example, alludes to more reminiscent techniques

³⁰"If I were topologically fix positions in space and structurally compose these accordingly, in order to have sounds move through space in specific orders, then this can lead to an unprecedented lively music."

³¹Stockhausen's approach to segment the circle into equal segments of a spatial scale was ridiculed "perceptual nonsense" in Harley (1994).

in melody composition: being called *spatial counterpoint*, it makes use of different points in space to create a counterpoint development articulated in space (Bates 2009, Stefani & Lauke 2010).

As the composer encapsulates different gestures into groups to form complex trajectories, each with their respective perceptive impacts on the spectromorphology itself, gestural spatialization can become quite overwhelming from a compositional side. Controlling these spatial movements can already be difficult in the studio alone. In terms of performance of a spatial work there are mainly two approaches almost opposing one another (Wyatt 1999, Normandeau 2009, Barreiro 2010): the first being the *fixed medium* performance, where a piece is composed in the electronic studio and simply played back in the performance space. Opponents to this approach mostly argue that the acoustics and the complete concert situation, such as the relation of the audience to the speaker array can not be attributed for and, effectively, the composition will be heard differently, if not with a disappointed ear. *Live diffusion* on the other hand gives the performer, which is often the composer himself, a chance to react to this concert situation. The means of live diffusion are usually a normal mixing board, if no experimental interfaces are specially developed for this purpose. I personally feel musically impaired, if the spatialization is limited to diffuse several precomposed channels of audio over manually controlled amplitude regulations. The effect I observed when attending such concerts is a washed down spatialization – a spectromorphological cloud that moves around the audience³² – a means to create a slightly shifting feeling of being enveloped, and so far I have not encountered radical new approaches to this technique, making pieces, from a spatial aspect, that are diffused in this way sound cliché and sometimes dull. Furthermore, it often limits the performer to, usually, eight speakers (one for each finger on the mixing desk) in the plane around the audience. Three dimensional source management poses a large logistical effort and poses new problems in the act of the actual spatialization. Unless specialized systems that allow for complex spatializations were previously developed, I regard the timbral diffusion by means of mixing desk techniques a mere enhancement – which, nevertheless very rapidly, is on its way to become a performance language of its own.

What many spatialized approaches try to achieve is to place the listener inside a real or imaginary sound world. In this, the importance of the core aspect of immersing the audience is at the (sometimes subliminal) heart of the composers intentions. The research undergone in (Lynch & Sazdov 2011b) is very interesting in this regard that a distinction between two dimensional immersion and three dimensional *engulfment* is drawn. The audience feels a separate sensation if the sound is actually coming from overhead. When placing the audience inside the sound world, a whole new set of sensations are triggered that a frontal oriented performance can

³²Due to the complex nature of so many aspects in electroacoustic music, I often have the feeling that live performances are still in the fledgling stages. Composers spend months and years in electronic studios producing complex electronic pieces for tape which simply can not be reproduced in a comparable manner in any live setting. Of course, both approaches have created an aesthetic and appreciative notion of their own, but in the end it is the fixed medium composition with its, as I see it, larger pool of possibilities in the shaping of the sound that attract me, due to the time offline the composer has, instead of having to compose in real-time with limited possibilities. This holds particularly for the spatialization aspect of the composition

not deliver. I want to give a good example that somewhat bridges the gap between traditional acoustic concerts and surrounding performances. In his instrumental work, such as *Terretêktorh*, Xenakis created a piece in which the audience can experience what a member in a traditional orchestra would experience – everytime. Taking the analysis method developed by Solomon (2007) and his view towards spatial gestures in any acoustic music into account, the perception of spatial gestures are most dramatic when inside the ensemble, where panning effects are the most extreme. The audience becomes part of the performing body, but is restricted to their ears alone; or, restricted to the perception and may not contribute to the soundscape³³, blurring the difference between audience and performer. In turn, questions arise such as, does the performance member feel as part of the audience, since, audibly, there is only a minor difference between him and the true audience sitting right next to him? In this case, he does not listen differently to the piece as any audience member in the mean. What will the performance member perceive once he returns to traditional performances, but with this set of mind, considering spatial gestures within the ensemble manifested in him? Will the perception of music coming from the front (on stage) in a more traditional concert affect the audience perception thereafter, and will the spatial cue of constantly being in front suddenly receive a meaning? Similar questions like these are what a soundscape artist and sound installation artist are asking with their art, as they are concerned with the altered perception of spaces, real or projected, not only in the present state, but also in the continuative perception of both the current space and other spaces (Zelli 2009).

For my own practices, I discovered a method to compose close to the auditory mechanisms that make our world appear in our mind – whatever that might be for the individual. But, in this, I strive for techniques that reach more into abstract domains, play with our perception and spark creative connections, hence the elaborate discussion in the sections 3.1 & 3.2. When composing spatially, the abstract thinking usually goes towards Gestalt principles and effects, especially playing with emergence and reification. Kendall (2010), for example, speaks of a *Path* schema, which is constructed out of a combination of fragments, which can be described as an emergent path appearing due to the common fate the fragments exhibit together. Furthermore, there is a strong tendency to make use of psychoacoustic effects in my approaches to bend the space around the head of the listener. These underlying principles are not limited to expand the trajectorial methods, but I also work them into the more spatial sounds and spatial texture synthesis, which gives me the ability to construct warped spaces. Multistability, for example, could be achieved with techniques inspired by Brant, to play with background and foreground textures.³⁴

In general, it should be one of the long term goals of spatial music of any approach to shape the perception of space and challenge the auditory system to push the average listeners ability. That this cultural shaping towards improved musical hearing of more complex compositional

³³Of course, one might argue, considering the aesthetics of John Cage, that the audience is truly part of the performance. But, it is to be investigated if the works of Xenakis allow these aspects to become part of the composition.

³⁴Although finding suitable sounds and maintaining the effect is quite difficult.

practices in music is a natural phenomena was shown by [Bigand & Poulin-Charronnat \(2006\)](#). The space needs to be emphasized at some point in the composition, because without this emphasis it will fly by unnoticed. Spatial composition calls for the reduction of other parameters to steer the focus of the audience towards the space. I call this a form of *spatial reduced listening*, that, instead of listening to intrinsic characteristics of a sound, I listen to a sound's behavior in space, how I group sounds within the auditory scene and how this perception changes as the composer flips certain aspects in the composition. Eventually I sharpen my ears to the environment around me in the long run.

”Wenn ein Mensch etwas Akustisches erlebt, wird er verändert, weil er durch die Schwingungen moduliert wird, seine ganzen Atome werden moduliert; er kann nur zum Teil den Zustand wiederfinden, in dem sie vorher geschwungen haben. [...] Da gewinnt die Musik nun eine ungeheure Bedeutung. Sie kann zum Beispiel durch die Erweiterung des räumlichen Parameters dazu beitragen, daß eine Wahrnehmungsphänomenologie, die ebenso auf dem Akustischen wie auf dem Visuellen basiert, allmählich wieder entwickelt wird.”³⁵ ([Stockhausen 1964](#))

3.3.1 Point Source and Trajectorial Approaches

”[O]ur turning towards a source is calculated to bring our ears into maximum use [...] motion is essential to our knowledge of the spatial layout of our environment.”
– Edward Lippman ³⁶

Point sources and trajectory based approaches build the core of any spatialization aesthetic. This can be shown in the way we think of audio: every sound essentially ranges from one (or many discrete) channels. In the most basic way, every channel is directly connected to a physical speaker that holds a unique position in space. Therefore, when spatializing our sound, we employ several point sources to do so. Introducing amplitude panning between each speaker creates a continuity between all our speakers within the array, and requires a first level of abstraction where a sort of virtual speaker is created that we can reposition in any position between the two real ones. But, since the source audio – in its essence – is comprised of discrete channels, no spatialization synthesis method can avoid placing these discrete channels in specific positions in space. Based on the way we think of audio makes this approach intuitively straight forward. Point source spatialization and trajectories are without a doubt the most frequently used approaches in spatialization – and the most cliché. Sounds flying through space have a clear sense of causing the listener to *look* at the sound ([Barrett 2010](#)), hence the theatrical effect and the general reluctance of using trajectories. From the extensive discussion about Stockhausen's approach to a spatial language in chapter 2, one can conclude how old the

³⁵”If a human experiences something acoustic, he is changed, because he is modulated through the oscillations – all his atoms are modulated. He can only partly return to the previous state in which they were swinging before. [...] Suddenly, the music receives an enormous importance. It can, for example by amplifying the spatial parameter, be conducive to slowly developing a perceptive phenomenology that is based on acoustics as well as on visuals.”

³⁶As quoted in ([Zelli 2001](#))

notion of using solely trajectories is and that many have sought to establish a language around this methodology. Movement is one of the prime sensations that makes us aware of the space, because of our life-long bodily experience exploring and learning the space around us (Kendall 2010). At the same time, being so basic, knowing about trajectories seems fundamental. It is as if every composer of space music somewhat needs to arrive from the trajectorial approach. It is comparable to a student in a classical musical education having to study the old masters. You need to know your counterpoint harmonization before you can start going wild with Cage'ian chance. Therefore, although seen as cliché by some, its discussion is mandatory.

First of all, the desire that the sense of direction could give a notion of expression is prevalent; that the sheer movement of a source constitutes a message to be delivered. Every novice probably has thoughts similar to these, but also respected composers such as Truax base a large part of their work on this topic (Zelli 2001). Nevertheless, since these efforts were already seriously, although not necessarily successfully pursued by Stockhausen and contemporaries, it shows that a simple, linear mapping to human reactions is probably a romantic dream.³⁷ Most ideas similar to those back then do not find support in auditory perception (Harley 1994). This equality towards direction of our spatial perception can be demonstrated with the problems children have learning to distinguish the left from the right. This distinction is usually established with mnemonics, such as remembering which hand one writes with, but the initial confusion shows the equal balanced symmetry of both directions and that the words are merely exchangeable terms. Another egalitarian attitude towards direction is the apparent flip of the left with the right when we look into a mirror.³⁸ It does not disturb us much when checking that our tie is sitting properly that the whole world we're looking into is seemingly flipped around. But, although we perceive the illusion that the sides are switched, what is more important is that we clearly notice that the top and the bottom are still intact. Like with pitches and melodies, one tries to define an analog language by which space can act as a replacement to the like. Pitches go up and down and jump in intervals to form melodies as the intervals are split in time. Space has an *up* and a *down* on its own account, but like East is the West for the even farther East on the globe, *left* and *right* can only be defined in perspective to the audience. Furthermore, the up and down in melody can evoke feelings and emotions, which the relativity of the left and the right cannot simply replace. If ones intentions were still to produce a spatial melody, then a system with an elevation parameter is needed to give any sort of possibility to offer a pitch replacement or equivalent. The movement azimuth, on the other side, can only denote a change in position in relation to the body.

The third dimension, distance, avoids this discussion. The notion of *far* and *near* evoke reactions more so than *up* and *down*. In section 3.2 spatial notions of different zones of proximity were discussed, and the closer the listener is to having his privacy threshold broken through, the more distress the gesture might cause, while things that move into the distance give the

³⁷Of course, respected thinkers like Truax show much more sophistication due to the collective experiences made with this approach.

³⁸Although, this is only a perceptual illusion. What is really happening is that the front is the back and vice versa. The image we actually see in the mirror is a version of us pushed through ourselves.

listener room to relax. Smalley (1997), for example identifies five paths that identify breadth and depth: approach, departure, crossing, rotation, and wandering. But, also from the discussion in section 3.1.2 it is known that the distance perception is not well understood and its simulation in amplitude panned systems is inadequately implemented. This, for one, hinders the implementation of complex trajectories which make full use of all three dimensions. The accurate perception and distinction of such complex movement is questioned (Marentakis et al. 2008, Payri 2010). Secondly, though, the absence of a distance parameter limits one to the sphere and in modes of expression when using trajectories alone. This quickly leads to a heavy bias towards circular movement, which is only characterized by its direction and speed. Since, for example, Stockhausen rejected the distance parameter for reasons motivated by the Serialist aesthetics (Harley 1994), this might be one point why his theories evolve these limited possibilities only. Looking at historical accounts, distance has been frequently used for organizational purposes, though. Several ensembles, one on stage and the other off-stage, for example, would pass musical material back and forth, bringing different qualities selectively into the foreground. One important insight is that the listener would generally underestimate distances (Brungart 1993, Zahorik 2002, Cabrera et al. 2004) and it is recommended to overemphasize gestures and trajectories to compensate, or counteract this misperception.

Concluding from the paragraphs above, one can see that a similarity between the note and a space is hard to see, but probably impossible to establish. In the beginning of this section 3.3 the comparison between spatial and timbral composition was made. Considering that simple spatial gestures hardly carry any meaning with them, the comparison to timbral composition could be much more fruitful. A difference I see in this comparison, is that instead of a top-down approach, in where we start with an array of different instruments and timbres and define their difference to one another, spatial composition creates its *spatiotimbres* from the bottom up, starting with somewhat basic building blocks out of which a specific spatial image can be created. Segmentations of spatial groups, for example, only make sense in coarse regions and, of course, if psychoacoustically motivated, i.e. in relation to the listeners head. These regions can then be hierarchically ordered, so that the front, being the primary area of focus, is preferred to the back and the sides. Similar ideas can be found in the Wisharts methodology (Zelli 2001). Also similar to his *double gestures*, the basic blocks are then combined to form more complex gestures. "Speed of movement and regularity and shape of motion, too, are important factors in the characterisation of space. Fast motion expresses energy; slow movement may imply passivity; regular motion suggests something mechanical; angular, irregular motion may convey hesitancy or uncertainty."³⁹ What is important is the spectral changes the composer imposes on the spectromorphology of the original when placing and moving it in specific directions, specifically into the distance with the possible volume, permanent spectral (dissipation) and temporal spectral (doppler-effect) changes opposed on the sound. I suspect that with growing knowledge about the effects specific types of trajectories have on the spectra of a sound (and with an increase in quality of spatialization technologies), the minute details between different

³⁹Alistair MacDonald, as quoted in Zelli (2001)

trajectories will be discerned, not by the auditory spatial stimuli, but the ear will become more sensitive to the tiny changes, that after the coarser spatial information is extracted, trajectories will be consciously held apart by their color. If this effect is a mere timbral sensation or a true refinement in spatial resolution and distance estimation will have to be investigated once we reach such a general level of sophistication. Further on, "[a] trajectory is not necessarily a concentrated point-source. As the head or bulk of a gesture moves through space it can leave residues behind. Trajectories can therefore leave trails, which can be smeared across space or be spread in a more egalitarian way through space. It may be that the establishing of a residue is part of a transformation of a gesture into a spread setting - the spread setting is introduced by a trajectory." (Smalley 1997) The emergence of this effect, like many sensations that composers like to describe based on personal experiences, would need to be investigated through studies to establish a qualitative measure of its usefulness.

Further sensations described by composers and listeners is the feeling of antigravity that Stockhausen supposedly achieved during the premiere of this piece *Kontakte*. One might wonder if feelings of nausea or something similar are possible, quite like the excitement of being on a roller-coaster. Another effect, limiting the speed of trajectories is a sensation I call *flickering*. This disturbance of the trajectories continuity happens as the trajectory most likely moves beyond the auditory threshold of motion detection (see section 3.1.3). Spatial flickering is a sensation in which the smoothness of the fast moving trajectory is broken into segments, similar to the flicker of a broken television screen. This limits the speed of the trajectory considerably, if the artifact is to be avoided. Another observation, though, is that this effect does not (or, not until higher speeds) occur when, instead of using the source position system of the VBAP algorithm, employing an amplitude panning based on a virtual speaker setup. While this might suggest an effect emerging from within the algorithm, the observation should be viewed with a critical eye for a particular reason: both techniques could only be compared in a circular motion, which reduced the artifact due to the continuous nature of the trajectory. Fast, discontinuous motion exhibited more problems with spatial flickering. The exact nature of this effect could not be further investigated and was written off as being an effect occurring due to perceptual motion limits. Unfortunately, a more precise investigation into this effect is, therefore, left to future work. Further on, the speed of a trajectory is also limited if the composer wishes to achieve exact source placement in space: "Although front/back and up/down confusions are fairly common for a listener with stationary head, these confusions largely disappear when the stimulus lasts long enough to permit the listener to derive additional information by means of head rotations" (Loomis et al. 1999)

Grouping of sources becomes important when one wants to perform generalized tasks or create agglomerate gestures in a counterpunctual manner. A technique to minimize cue confusions might be a slight vibrating movement of all sources irrespectively, as this simulates an exaggerated rotation of the head and possibly allows better spatial cue resolution, even for the stationary head. Trajectorial approaches will be used as a means to govern the details of the greater super-trajectory, as each new spatialization technique could, as a group of interrelated point sources, be moved around in space to add new degrees of spatial expression. This en-

capsulation approach can result in a complex Gestalt of several spatial techniques. Examples of this encapsulation are spatial granular synthesis methods, in which each grain is governed by a complex, often chaotic or even stochastic movement. The trajectories themselves, even the particles themselves, may not even be perceivable as such, but, just like in regular granular synthesis, the algorithm used to control the cloud of grains in space has an impact on the spatial texture perceived (Bates 2004). In another example (Kendall 2010), he describes how image schemas help us group hard to grasp events into a cohesive whole. Fragments of different sounds, all with a general sense of the same direction, can be brought together in the mind to create one large trajectory. The example demonstrates a common application of the Gestalt concept of *common faith*.

In trajectorial approaches, the technique could involve the method of identifying the spatial energy that the spectromorphology of the source sound suggests. As I argued in section 3.2, the projection of its spectromorphology onto a spatiomorphology could constitute a complete compositional approach. The reception of the audience could be the mere appreciation for the composers ability to identify a certain solution that is surprising and new. Smalley (2007) writes about a notion called the gravitation in spectral space, first introduced by François Bayle. He relates to pitch music, where there is sense of up and down, cadence (latin for 'fall') and the general notion that we are on safe ground. Spectral gravitation is explained as being one of the driving forces for spatial illusions and consists of *planes*, residual spectral points of reference, that act as attractors for *diagonal forces* along which spectromorphology moves along. Surely, this is one approach to a spectromorphological mapping, and every composer should have his own notion of how a sound behaves, making up the creative process, the personal, appreciative effort of sound spatialization.

One different notion of how a mapping of a spectromorphology into a spatiomorphology could look like, is by the illusion of rhythmic pattern within the spectromorphology. Every timbre depicts rhythmic beatings to some degree, close frequencies that interact in close frequency space with one another. Further on, complex spectromorphologies might include a large development on different formants appearing and disappearing, possibly even in a repetitive manner. On the other hand, a spatialization can always be seen as a rhythm in space. Every change in direction or reoccurrence of an oscillating motion denotes points of reference, which, when compared with one another build a rhythm alluded through space. When writing that for spatial reduced listening, a sound should, if possible, be removed from all other musical parameters, it is then clear that rhythmic patterns can not be avoided. They are inherent in the spatial interrelations between sounds. Furthermore, a sound which *sounds* can not be removed from its spectromorphology, which, in turn, denotes rhythmic fluctuations. If movement alludes to rhythm, the illusion can clearly go the other way, and we create a rhythmic interlocking between the spectro- and spatiomorphology. The spatial composer then has the option to play with or move against this rhythmic suggestion, effectively creating a single or several audible streams of different musical interpretations. Working with the rhythm creates a fortified gesture and allows the concentration to a single agglomerate gesture throughout all musical perceptions. This technique can be quite rewarding if used sparsely.

A recent study by Payri (2010) uses this rhythm-to-space analogy and they even transcribe their stereo gestures in term of rhythmic notation. The study differentiated between experienced and non-experienced listeners and investigated simple (for non-musicians) to somewhat sophisticated rhythmic patterns (for musicians) created by panning both a synthetic or real sound between the stereo panorama. The simple patterns were either periodic patterns and accelerandos of the like, while the more complex rhythms exhibited closely similar rhythmic phrases that would probably challenge a non-musician in a more conventional situation, played by acoustic instruments. It should be emphasized, that the sound stayed the same as it was only the panning that would contribute to the rhythmic illusion. This was done either continuously through amplitude panning or discontinuously as point sources. The results are bloodcurdling, though: a pattern recognition rate as low as 30% was determined. They studied many factors that would contribute in recognition of the simple gestures, and it turns out that musicians and non-musicians alike would not listen to the spatial interplay, but rather try to listen to intensity variations and other more traditional factors first. Moreover, seating position in this context was apparently irrelevant, which maybe is the case as they investigated rhythmic illusions instead of precise trajectories. It can be concluded, therefore, that, due to the confusions over the intensity, the space was not perceived as a musical gesture, hence the low recognition rate, as it was the main contributor to the differentiation of the patterns.

Questions arise, though, like those of the speaker separation, and speed of the rhythmic patterns. The study does state that sounds were panned hard left and hard right for each beat, how far were the speakers separated in relation to each listener position? Furthermore, although sine tones are hard to locate accurately, the results show that gesture detection was best with synthetic sounds than with real-life sounds. For one, rhythmic patterns can be nicely created with continuous sounds and basic building blocks. The reason for the bad recognition of real-world sounds is the dynamic variation these sounds naturally occupy. Continuous sounds do not distract the listener with other parameters, shifting the focus completely towards the only variability, which is in the spatial location. The detection of patterns created with these sounds were around 75% and higher. However, dynamically active samples, such as speech, instruments or other concrete sounds, have a rhythm of their own, exhibited through the timbre or temporal development, which counteracts the intended rhythm created through the spatialization. Not only that, but natural sounds with their conventional stimuli distract from lesser conventional gestures, especially when poorly spatialized. The study simply ignores this fact, and concludes the impracticality of spatial trajectories in music. Instead, what the study really depicts, is the inability to treat the spatial parameter in a completely independent fashion, which is a notion Stockhausen had to face at some point – and this is nothing new. As quoted in section 3.2 already: "Musical space [...] is a space of relationships." (Morgan 1980) The spatial language is tied to the space its speaks into, and to the vocabulary it uses. Spatialization happens within a projected space, and it is done with sounds that bare space within themselves. Furthermore, the listener has to be sensitized to the spatial parameter. The simultaneous attention to two sources is debated and requires trained mental skills (Arons 2001), the listener is most likely further distracted by listening to melodic and rhythmic modulations within the source itself,

pushing the spatial location into the subconsciousness.

Another study in the perception of different shapes of trajectories was undergone by [Marentakis et al. \(2008\)](#). It investigated the feasibility of using trajectories as expressive means to carry meaning to the listener. For the test, they used straight, arced and complex gestures. They identified three principal concepts that diminish the perception of specific trajectories. Foremost, as expected, there is the complexity of the trajectory itself. This goes hand in hand with minute details between gestures, as more complex gestures show minute details within the trajectory itself. Easy gestures are those that cover large spaces, particularly within and through the listening space. Based on the distance discussion in section 3.1.2, it is clear that gestures that reach outside the loudspeaker array need to be treated with care. Secondly, the number of simultaneous trajectories and interfering, masking sounds hinder the perception of the right gesture. For one, this could be limited by the listeners ability to concentrate on multiple sources, Finally the position of the listener can have an impact on the gesture recognition. This is to say, while an existing sweet spot limits the perception, if a spatialization system without a sweet spot would be included, the listeners position would still affect his perception if his relation in distance and angle are within or beyond the range of some minimal noticeable difference. The notion in this case is similar to that of an installation, with the restriction that a listener has to sit through the whole composition in one fixed spot. In that sense, one can understand Xenakis when he stated that his spatial concerts should be attended several times to experience the pieces from different points of view ([Santana 2001](#)).

Interestingly, but also (in the light of this research) somewhat disappointingly, [Payri](#) speaks explicitly about the learning curve during the experiment, with no good prospects: "[...] musical training does not improve spatial trajectories recognition [...]" ([Payri 2010](#)) But they also point themselves to the study undergone by [Bigand & Poulin-Charronnat \(2006\)](#) which investigates the effect mere exposure can have to the general "unexperienced" music listener. [Bigand & Poulin-Charronnat](#) conclude that our culture is very well able to absorb new musical ideas and idioms so that people with no musical training are just as well able to identify and perhaps even appreciate musical gestures in contemporary movements. Nevertheless, the exposure itself can not be stressed enough. Specifically thinking about spatial relationships in music in physical space, the heightened awareness of spatial perception is still to be achieved: "How the listener experiences artistic meaning in auditory spatial art depends greatly on the everyday experience of space, and clearly the listener's most important spatial framework is anchored in the body. [...] The deeply meaningful sense of space that is aroused when listening to electroacoustic music has its roots in a lifetime of embodied spatial experience." ([Kendall 2010](#)) For a better understanding of trajectories and movement in space in music, the exposure for spatially meaningful music can improve the perception of space, which, in turn, enables composers to push the limits of perception even further. Both the audience gains from a heightened spatial awareness and the composer in an increased freedom of spatial gestures.

[Marentakis et al. \(2008\)](#) further observed that gestures can be improved when the sensation is cross-modally stimulated. Since our life-long spatial experience is done over all stimuli, this conclusion is perhaps not very surprising, but in acousmatic music the tendency and general

attitude towards musical display is the removal of all audible causes. That said, a performer sitting in front of a mixing desk probably is not enough to trigger this spatial impression. The gesture needs to be clearly associated with the audible stimuli for this effect to interlock. "The interpretation of the gestures is to a large extent a social phenomenon and depends very much on the existence of a common ground between the performer and the audience." (Marentakis et al. 2008) One problem that Marentakis et al. (2008) observed is the placement of the performer with respect to an expressive spatial performance. While the performer needs to hear the spatial distribution in order to react in an appropriate and expressive way, placing him in the middle of the audience is not an optimal solution with respect to the sweet spot of the distribution system. Ideally, the audience should be experiencing the best quality sound and concert halls are usually forward oriented, making the effect of seeing the performer obsolete for the audience members sitting in front of the performer. One solution they propose is the use of a HRTF encoded monitoring version that is fed to the performer on stage over headphones. Here, the problem, though, is the distorted relationship the performer has to the performance space, into which the sound is projected into.

Finally, this idea, translating the physical movement of a performer into sonic movement through virtual space is an interesting, and, considering our bodily experience of space, probably the most true approach to trajectorial gestures: "[...] every approach ought to be informed by the realities of spatial perception and also acknowledge that listeners are making meaningful sense of spatial events." (Kendall 2010) At this point, I want to return to the idea of creating rhythms with spatial movement and the notion of significant points in space within a trajectory. First, rhythmic suggestion through mere spatialization can easily be considered a musical act. Of course, rhythmicity – just like spatialization, then – does not require periodicity. This is solely dependent on the composer and his intended audience. Moreover, parallels could be drawn to another temporal art that is concerned with movement and rhythm on its own account. If the on-stage action of a live performer translates into virtual sound movement in space, then analogies may be drawn with the language of contemporary dance. Reynolds (1978) compared a list of *effects* by dance artist Rudolf von Laban to the movement of sound. Effects are products of *impulses*, which describes a notion that all movement emerges as the conclusion of some preparation. Reynolds ponders the possible meaning of impulse in a spatio-musical context, its respective preparations and expresses the idea that "[...] a particular arc of movement in space, sonically described, [might] be more comprehensible if appropriate impulses were involved in articulating its elements." (Reynolds 1978), A direct translation of these effects is nonsensical, as dance is an art form that lives of both the dimensional restrictions of the body and the play with gravitation, two limits that virtual sound space in a speaker array does not exhibit. It is also not the intent that spatial music should imitate dance, but, instead be inspired by the linguistic characteristics. Space is a full body experience, and, in such, the comparison and inspiration from dance seems quite appropriate.

Dance is an art of and with the body. It lives from the rhythms of the bodily functions, that of heart and the breathing, and with the room defined by the physical limits the body imposes; the *kinesphere*, after Laban. Furthermore, being an art form that is delivered visually,

it can make use of filigree movements, from only moving the fingers to facial expression – spatial details the audible space simply can not compare with. Instead, the combination and interrelation of the spectromorphology and its spatialization would have to be viewed together, for any parallels to the language of dance to be drawn. What signifies mimic in postmodern Butoh dance, for example, then might have a similar meaning in the spectromorphology of a commonly spatial piece. Currently, the notational dance system by Laban is being further developed in the dance company of William Forsythe: "There is still no standardized system of notation for sequences of movement and choreography. Several years ago the choreographer William Forsythe created a new and very promising system of notation with which one can study or reconstruct an entire choreography without actually having seen it." ([German Federal Cultural Foundation n.d.](#)) Being granted a fund, the four year project, having started in 2010, is currently still under development as of the writing of this thesis ([The Forsythe Company n.d.](#)). Its development should, nevertheless, be followed by spatial composers alike. Generally speaking, there is a clear sense of development from simple to complex gestures, from which one may then create further modalities. Forsythe seems to work a lot around the principles of attraction and repulsion, as well as creation and destruction, as he demonstrates in his *room writing* techniques. Room writing is a technique to work with the further-reaching space the dancer finds himself in. Objects, as simple as a doorknob, are created through abstract gestures and are immediately worked with, destroyed and thrown away.⁴⁰ Movements described as torsions are examples where the ideas of spectromorphology could exhibit a similar concept in a spatio-musical language. Direct use of the vocabulary could be made in the case of *spatial reorganization*, in which one body part defines a plane in three dimensional space that, when moved and twisted, all other parts follow the spatial distortion. It is a salient example in which higher order gestures that group smaller trajectories together could be explained with.

"Unter Berücksichtigung konventionalisierter Hörerfahrung bezieht sich der Begriff der Klangbewegung auf die sukzessive Verbindung von musikalischen Elementen, die entweder durch Dauern, Klangfarbe oder den musikalischen Kontext so stark aufeinander bezogen sind, dass sich auch über große Raumdistanzen hinweg das Nachfolgende aus dem Vorhergehenden ergibt und demzufolge eine eindeutige Zeitrichtung der klangräumlichen Entfaltung deutlich wird."⁴¹ Nauck-Wicke⁴²

Decorrelation effects

In terms of trajectory-based composition, I see decorrelation effects, as basic as time delays as, an important extension of the spatial vocabulary. Many additional gestures arise with a creative, and especially dynamic approach, which would not be possible with mere amplitude panning.

⁴⁰These actions are motivated by their resulting movements.

⁴¹"Considering conventional listening experiences, the term *sound movement* refers to the successive connection of musical elements, which are so strongly related to each other, either through their length, timbre, or musical context, that, over greater spatial distances, the subsequent results out of the antecedent and a definite direction through time, therefore, becomes clear."

⁴²As quoted in ([Zelli 2001](#))

In music production, techniques such as stereoized reverberation and chorusing is quite usual (Kendall 1995). In virtual surround environments, usual application of decorrelation is the control of source width. When decorrelating the source between two speakers, the impression is a psychoacoustic spreading of the audio in the space between them. Potard & Burnett (2004) identifies several different techniques for achieving this effect: the first is a time delay, although this technique is limited to a delay of about 40ms and it introduces comb-filtering effects, which might not be desired. Other techniques are based on all-pass filters, with a noise-like phase response. Next to the simple static filtering technique, time-varying approaches in decorrelation are highly desirable, artistically speaking, due to the added modes of expression they allow. Potard & Burnett (2004), however, introduces a novelty by applying decorrelation to selected sub-bands, allowing for specific degrees of decorrelation in certain frequency ranges. The sub-band decorrelation is reported to have been well perceivable by the author, but seems to require some training beforehand. Perceptual effects of decorrelation are identified in Kendall (1995). For one, decorrelation eliminates coloring and combing effects. Introducing a delayed signal close enough for combing effects to happen, a study in Kendall (1995) reveals that a gradual increase of decorrelation by the means of a variable FIR will eventually restore the impression of the original timbre. Furthermore, Kendall uses decorrelation techniques which produce "cheap" diffuse sound fields, that avoid the use of system-heavy, multichannel reverberations. For performance of multichannel music, decorrelation is especially interesting due to the loss of a sweet spot in the concert hall. For example, a signal is shifted perceptually between two speakers as the time delay between the two is increased up to around 1ms. Any value further than that will be subject to the precedence effect and, beyond its threshold, turn into an echo. When decorrelated, the signal will remain equally perceptible in both speakers for all audience members until a clear temporal difference between them is perceived.

Decorrelation by means of delays can act as a trajectorial approach, since the use of ITDs will have a similar effect as normal amplitude panning by the means of ILD – but it has a different spatial flavor. For one, due to phase resolution limits of the auditory physiology (see chapter 3.1), frequencies below $\sim 1\text{kHz}$ are affected. Having the lower frequencies panned only with an equal volume in both speakers is generally an unknown sensation, because we are either used to both horizontal cues, or, coming from industry-style production techniques, mostly hear ILDs in panning techniques alone. This effect seems to define the space the sound will move through (Vaggione 2001), broadening the image across the respective speakers. The trajectory is somewhere between being blurred and defined. When applying phase decorrelation, the image is further broadened over all frequencies. Different approaches to source decorrelation lead to significantly different perceptions, and moving through these images of alternating "spatial color" can lead to refreshing and perceptive results, meaning a sound source is not just panned through space, but the space is shaped in the course of its movement. In effect, the space is opened up and presented to the listener, in which his attention is steered towards the width and height of the space. This is what Vaggione (2001) meant with "defining the space". In balance with amplitude panning techniques, the expression can lead to meaningful messages.

In my composition *The Common Perkins Scream* (2010)⁴³ for headphones or stereo speakers, heavy use was made of different spatialization techniques to create complex and intricate movements with the evolving auditory Gestalt of a prerecorded feedback loop. Being a mono signal to start with, the main challenge was to compose a spatialization for both channels. Broadening technique by means of decorrelation was done in many, but selective parts. In the beginning of the piece, at around 1:40, one of the piercing sounds suddenly spread across the whole listening field, after being subjected to many similar sounds that were only slightly, almost unnoticeably amplitude panned to either side. Minute 2:47 features a section of the audio with a slightly time-stretched copy in the opposite speaker, being aligned at a transient-like sound at the end of the gesture. As the time-stretched copy catches up to the original, circular warping effects around the head speed up to reach a climax, followed by a release due to the final transient. The sensation of a phase shifted sine tone can indeed be a warping sensation around and behind the head as opposed to only panned between the speaker pair. This effect is prominently used by Ryoji Ikeda in his work *Matrix (for rooms)* (2001). The sound even seems to be coming from inside the head, even though it is played with external stereo speakers. As experienced in the second experiment in section 3.1.2, phase inversion can be used to pan a sound into the audience. When creating a trajectory which should travel through the audience space, good care should be taken to have the trajectory at speed low enough for the effect to be perceived. As discussed, though, the effect is not stable enough to create a uniform sensation, especially for fast movements.

***3dSpamlePlayer*: a Max/MSP spatialization tool**

Working with samples in Max/MSP is not as straight forward as it would be in digital audio workstations (DAW). Being more a programming language than an audio tool, one needs to synchronize the spatialization manually with the audio stream. Furthermore, the positioning information is sent in spherical coordinates, which quickly leads to purely circular movements that evolve around the listener in several revolutions. After having worked for a few weeks in the studio, I quickly came to the notion that a more intricate and delicate control over the spatialization needed to be implemented. Also, wanting to move away from predominantly circular thinking, a small subpatch that would translate from cartesian to spherical coordinates had to be implemented. The first version that could be sequenced when synchronized was a three-dimensional vector interpolator (figure 3.4). It takes a 4 element list, a cartesian vector with an interpolation time, and returns float values of a three-dimensional ramp within the time specified. Inspired by the standard [line]-objects of Max/MSP, this subpatch would take and store any incoming 4 element list in an internal FIFO queue, so that complex sequential motions could be programmed.

But the level of control was insufficient. Furthermore, while cartesian coordinates did result in a more cubic way of thinking, the distance parameter was not implemented in the main VBAP system at the time and the whole approach seemed a bit pointless. When working

⁴³This composition was done prior to this thesis and premiered publicly at the Electroacoustic Music Night in Dublin, 12th of June 2011, as well as at the ISSTC 2011 in Limerick, 10th–11th of August 2011.

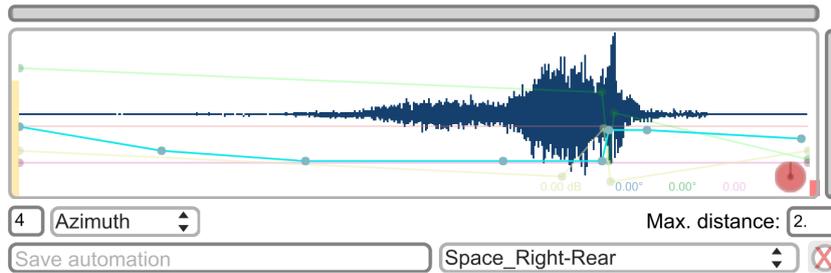


Figure 3.5: Screenshot of the `3dSamplePlayer` subpatch, presented within a `[bpatcher]`-object.

appears during playback at the exact location within the waveform and automatization curves, allowing for easy identification of prominent points in the sample, as well as verification and fine tuning of the automation curves. Using the `[function]`-object as automation curves has the nice advantage, that the right output from the `[groove~]`-object, giving the current position in the sample between 0 and 1, can be fed directly into the respective object, which returns an interpolated value between the user-specified point in the UI. Samples are loaded into the `3dSamplePlayer` by simply drag&dropping the file from the filesystem over the `[bpatcher]`.

Next to the drop-down list to switch between the automation curves in the bottom left of the lower UI elements, there is a number input on the far left, specifying the source number on the receiving VBAP server. It specifies the audio channel over which the audio and the spatialization information should be sent. The maximum distance input on the right gives the user a chance to better fine-tune the distance parameter in the automation curve. Because the distance can go into infinity, theoretically, this measure was implemented to have better control over lower distance values if the maximum distance to be used is known beforehand. The measure can be modified at anytime without losing or changing previously set values. One idea here would be to implement a logarithmic scale in the future and remove the UI element for a more simplistic interface. It should be mentioned that, for the convenience of the high level user, the whole UI adjusts itself proportionally to the resizing of the containing `[bpatcher]`. The user can, therefore, resize the work area to his liking in the superordinate patch.

Finally, one of the most important features of this module can be found in the bottom UI elements. The large disadvantage of the `[function]`-objects is namely that there is no built-in way to store the automation. Therefore, a self-built database structure using `[coll]`-objects was implemented. Unfortunately, the `[coll]`-objects contained in the subpatches also do not store values on their own. The user, or composer, has to include a "mother"-`[coll]`-object in the superordinate patch representing the whole composition. All values from all `3dSamplePlayer` modules will be stored in this one `[coll]`-object. Each `3dSamplePlayer` module therefore needs a unique name in the patch, as well as the name of the "mother"-`[coll]`-object given as the initial parameters. To avoid possible conflicts there are actually two databases maintained within the whole system. One is dedicated to the automation values and the other holds other important features that should be set when re-opening the previously saved patch (i.e. composition). These include loading the sound-file used as the sample, as well as the last used automation curve, the associated channel number and maximum distance value (see table 3.1).

%SP name%_automation name%	domain "%(float)value%" azimuth "%(list)values%" elevation "%(list)values%" volume "%(list)values%"
%SP name%_sampleFile	%(string)path to file%
%SP name%_LastUsedCurve	%(string)automation name%
%SP name%_channelNumber	%(int)value%
%SP name%_distanceVal	%(float)value%

Table 3.1: The formatting convention for the [coll]-style database. Left column are the keys, while right column are the associated values. The first entry, holding the automation curve values, would be stored in its dedicated database.

The usage of the management interface is done to be very simple and intuitive. Upon hitting enter after typing a name for an automation curve on the left, the curve is automatically stored and can be recalled at will. Entering the name of an existing automation curve will overwrite this curve. Loading an automation curve is as easy as selecting it from the from drop-down menu, and upon clicking on the red 'X' on the right deletes the currently selected automation. The user will have to save the patch, for the values to be stored before closing it. Unfortunately there is no easy way to have Max/MSP itself ask the user to store all changes upon closing the patch. The only time this happens is if the [coll]-objects are modified themselves, other changes are not stored. A highly desired feature to be implemented in the future is the ability to save all data into the currently selected automations, as individually saving each curve on every change can become quite tedious and is easily forgotten. The [coll]-objects will store the data directly within the *.maxpat*-file for easy distribution.

The subpatch also has inlets and outlets, although the features are currently still very limited. Sending a 1 or a *bang* into its only inlet will make the *3dSamplePlayer* play the sample and its automations from start to finish. 0 or *stop* will stop the playback. To crop the sample playback, one can send a *playselection* message with two floats between 0 and 1, relative to the length of the sample. Determining exact values for this message is close to impossible and mostly done through experimentation. This will have to be revised in the future. The left outlet does give a reference value between 0 and 1 during playback, from which one can derive estimations. The last feature, the inlet currently takes, is the *preset* message, which, when sent together with the name of a stored preset, allows for algorithmic changes of the automation curves. Finally, the right outlet sends a *bong* upon finished playback of either the whole sample or the subsection. Other planned features for the future is the input and output of audio streams. Currently, the domain, the length of the horizontal axis of the *3dSamplePlayer* is only set based on the loaded sample given. In the future, this should be possible independently and without loading a sample. The player would then take the incoming audio stream upon playback and spatialize it according to the set automation curve. Conversely, every *3dSamplePlayer* should output the audio stream currently playing. This would allow the stream to be fed into a second, but empty *3dSamplePlayer* to be sent to a different location in space, after having done some treatment to achieve possible decorrelation and other psychoacoustic effects.

One last remark should be done on the workflow that this new widget enables and motivates. Although the *3dSamplePlayer* itself is inspired by usual designs in DAWs, the user environment in Max/MSP does not have an overall timeline. The composer needs to synchronize several *3dSamplePlayer*-[bpatchers] in his way. Samples are, therefore, not just placed on a timeline, but need to be triggered in some order. That this order does not necessarily have to be linear is given in Max/MSP. A delineation of how this might look like is given in figure 3.6. The samples can be synchronized by a global *bang*, delayed respectively, or samples can trigger each other, either by a *bang* at the end of the current playback region or by monitoring the current position. Programming further possibilities around this second method can quickly lead to algorithmic composition, autonomous playing and particularly *installations*, a compositional technique which we will look at in the following section 3.3.2.

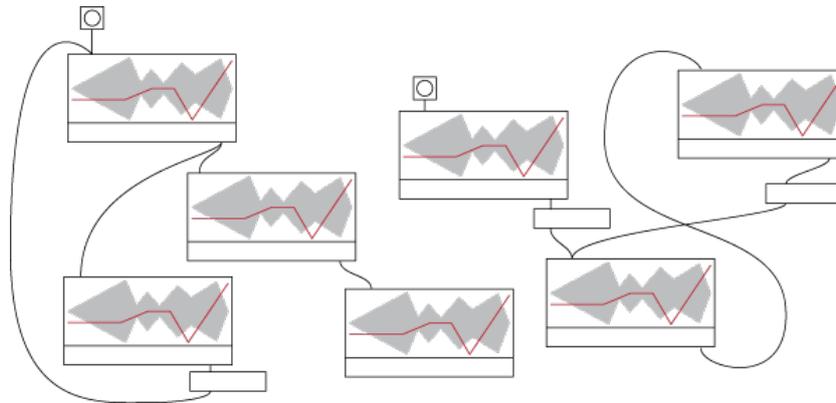


Figure 3.6: A schematic example of an asynchronous organization of 3D samples in Max/MSP.

3.3.2 Architectural Pieces and Soundscapes

Aside from a moving source that defines the space by its limitations in movement, there is also the space itself, with which one may compose. These approaches, nevertheless, are left aside to an extent, as they are not necessarily acousmatic. In soundscape composition, the usual practice might be capture and juxtapose recorded soundscapes to one another. The idea of distorting a sound from its source is a minor notion, if not non-existent and the exact opposite is true. The only source that should be veiled the best possible is the reproducing loudspeaker itself. Spaces can be seen as resonances, reverberation times or complex ambiances, to give a few examples. A complex space, of course, usually does not lack any of these parameters, with the exception of an anechoic chamber, perhaps. Nevertheless, there are two clear distinctions made, which are based on the performance space itself: a soundscape composition is usually mainly concerned with the content on tape and is performed in a usual concert setting. Site-specific, or architectural pieces, on the other hand, are usually installations that exhibit a much closer, to inseparable and fundamental relation to the space that the audio is projected into. While these two notions are usually distinguished, they share one fundamental concern: real spaces. In this, it is not digressive to look at both artistic movements in relation to each other and draw conclusions of the use of real spaces.

Soundscapes are made up of simply much more than just the correct placement of events in space. A soundscape is sublimely governed by the background noise, the ambience, the little, natural sounds that fuse together to a texture. In terms of reverberation, one especially has to consider the correct simulation of the characteristic first reflections, to allude to a specific space. For soundscape compositions in particular, having precise control over the reverberation of a space would be the ultimate control. This is meant not just internally, i.e. on the tape, but, more importantly, externally, in the performance space. Unless the performance space is an anechoic chamber, this level control seems virtually impossible, at the moment. Active cancelation of room acoustics, meaning the creation of an anechoic chamber by simply projecting sound in the the performance space, has, unfortunately, not yet been discovered.⁴⁴ What site-specific installations do, in turn, is play *with* the acoustics of a space instead of against them. The physical space, hence, becomes part of the composition.

One argument for composing for unusual spaces, usually not allowing standardized speaker configurations to be installed, is the necessity of the composer to engage with the space, therefore, in order to determine more conceptual and interactive positions for each speaker to placed in (Lauke 2010). A specific care is taken in architectural installations to enhance the experience of the actual physical space in which the music lives. The space is, therefore, not so much contained inside the sound, on tape, so to speak, but the sound is rather placed inside the specific place. A point made by Smalley (2007) in particular is taken into account, in that listeners usually do not instantly engage in spatial listening, but need time to adapt and become accustomed to the spatial dynamics of a piece (Lauke 2010). The particular attitude towards the use of time in installations, namely that due to the nature of having no official beginning and ending, each visitor experiences the installation differently and installation pieces usually become indifferent toward the temporal development on the macroscopic scale (Normandeau 2009), works in favor of this perceptual notion. But, for Stefani & Lauke (2010), the site-specific view, the consideration of the performance space, also goes beyond that of the installation. While it is an integral part of any installation, be it considered or not, the same applies to any performance space as well. Even though Normandeau (2009) draws distinctions between the two approaches, i.e. the temporal aspect of each, or the ability of the audience to freely explore the space on their own account, he does note the similarity that the space into which the music is projected plays a particular and characterizing role in the performance of an acousmatic piece. Stefani & Lauke (2010), therefore, argue that even a few days preparation for the composer in the respective space of an upcoming performance might enhance the performance quality.

But site-specific approaches also hand over the responsibility and control over the piece to the physical space completely. The site-specific composer can, of course, vary in degree of acoustic knowledge of the space from giving particular considerations based on its properties to composing a complete framework around the space. In soundscape compositions, conversely, the space lives rather *inside* the sound recording, enabling many more degrees of freedom in its manipulation. Soundscapes can include simple methods such as reproducing recordings

⁴⁴Although, the technique is blatantly mentioned in Stefani & Lauke (2010) as a possibility, but without further details on its realization.

of external spaces and bringing them into the concert hall, or artificially recreating acoustic spatial pieces, for example the *Symphony of the Factory Sirens* by Arsenij Avraamov, creating points and splashes of military sounds around the audience as if they were back in Baku, listening to the piece. Artificial soundscape synthesis can be found in [Verron et al. \(2009\)](#), for example, where enveloping, environmental sounds, such as air and waves, are created using an efficient stochastic model using the inverse FFT algorithm. Composer Pauline Oliveros ([2003](#)), for example, engages in compositional approaches that tamper with the acoustic properties themselves. Embracing the emergence of powerful personal processors, she makes use of the ability to simulate room reflections and reverberation coefficients in real time. The degrees of possibilities in which this can be achieved are a mere growing or contracting of a room. But audible barriers and obstacles, such as virtual walls and other larger objects can be alluded by the mere alteration of spatial properties of a room. Besides moving sounds in space, the audience can have the impression of being moved through space themselves. Furthermore, she writes about the use of the (artificial) room resonances as an instrument, creating a specific harmony or timbre. This effect was already displayed by Alvin Lucier in his famous composition *I am sitting in a room* (see chapter [2.3.3](#)), in which he amplified the room resonances through constant modification of an original single, being projected and recorded together with the room resonances over and over again. With the advent of computation and simulation of these resonances, the virtual room can be changed and the resonances altered to the composers liking.

For Hildegard Westerkamp, soundscape composition is a means to exaggerate spatiality ([Zelli 2009](#)). In her composition *Kits Beach Soundwalk* (1989) she demonstrates the microphone as an acoustic zooming tool, being able to focus into sounds around the Kits Beach in Vancouver, Canada. Sounds and spaces of proximity are brought into the audible and faded away again, as she captures the different, personal impressions of the space in relation to the different moods with which one can experience the space. In *Kits Beach Soundwalk*, the demonstration is accompanied by a narrative speech voice, in particular, making this piece an excellent demonstrative example. In general, it is the relationship to the space of the piece, which, for her, defines this art form: "Soundscape ideology recognizes that when humans enter an environment, they have an immediate effect on the sounds; the soundscape is humanmade and in that sense, composed. Soundscape is the acoustic manifestation of place, where the sounds give the inhabitants a sense of place and the place's acoustic quality is shaped by the inhabitants' activities and behaviour. The meanings of a place and its sounds are created precisely because of this interaction between soundscape and people."⁴⁵ Taking the delineation of Barry Truax into account ([Zelli 2009](#)), the approach taken by Westerkamp would then be called a *framed* soundscape composition, partially removing sounds in the soundscape from their greater context. The opposite classification, for Truax, would then be the unaltered playback of soundscapes and represents the truest form of soundscape reproduction, trying to establish an impression of being at the external space as good as possible with the conceptual and technological means given.

⁴⁵As quoted in [Zelli \(2009\)](#)

3.3.3 Virtual Spaces and Spatial Textures

”Sirius is based entirely on a new concept of spatial movement. The sound moves so fast in rotations and slopes and all sorts of spatial movements that it seems to stand still, but it vibrates. It is [an] entirely different kind of sound experience, because you are no longer aware of speakers, of sources of sound – the sound is everywhere, it is within you. When you move your head even the slightest bit, it changes color, because different distances occur between the sound sources” – Stockhausen⁴⁶

The notion that movement and space are linked to each other has been extensively discussed in the preceding sections. Movement can create a mental representation of space through exploring or by following the movement of others. But, while experimenting with spaces themselves and their characteristic sound(s), the idea that composing the space goes beyond the movement of sources finally arrives at the notion of the space in the sound and finally at the sound being the space (see section 3.2 for a detailed discussion). Sound sources are not just seen as point sources anymore, but effort is made to make a sound appear like a carpet, or a wall, to the point that the single sound completely envelops the listening space. The source movement becomes irrelevant, and instead the space itself, being the sound, is moved. Furthermore, though, other spatial developments become more important than the sheer change of a sound's position. For the impression of a sound being completely surrounding to arise, each area around the head has to exhibit a degree of decorrelation to not be fused together in the brain, losing the enveloping and extrinsic feeling. As the spectromorphology of these sounds vary to minute degrees in different areas in space, the temporal development of the sound might be seen as spatially static, but these spatially disjunct events make the sound morph, shape and pulse like a thick cloud of ever changing entropy. It is, therefore, not the movement of the whole audible cloud around the audience but the changes and development in the spatial sub-regions, and the relation of these changes to each other within the complete development of the spatial sound. Moreover, (Smalley 2007) takes the thought a step further and argues that the spatial image can be collapsed into one singular moment, beyond the human noticeable time. When the focal point lies on the spatial development, the temporal development becomes irrelevant. The exact timing of events blurs into lesser importance, up to the point of being neglected completely.

Live sound diffusion, for example, deviates away from precise source positioning, and towards the approach of enveloping spaces. Although the sounds are somewhat moved around the audience using the faders on a sound console, the desired effect is less defined. In contrast to creating exact positions and trajectories, which supporters of live diffusion believe to easily become ambiguous and collapse (Normandeau 2009), they create diffuse sound fields that change shape and pulsate through the audience instead. They somewhat smear the spectromorphology across the listening space by playing, or scattering it through all loudspeakers together. But, this attitude towards sound diffusion does not have to be limited to a live scenario, but other, more complex methods of sound diffusion may be developed in the studio. These experimental

⁴⁶As quoted in Bates (2009)

techniques, such as spatialized granulation and spectral diffusion, which will also be introduced in chapter 4 in more detail, would require more intricate and convoluted methods of control in a live setting. Nevertheless, all approaches have one common goal in mind, namely to reach a level of envelopment by which the sound itself loses its point source association. "One possible interpretation of this process is that the sound creates its own space – or that the sound becomes the space." (Barreiro 2010) For composer Trevor Wishart (1986), these sounds were differentiated from point sources as *textures*, alluding to haptic and visual analogies. Smalley (1997) further argues that the texture forms when one "zooms out" and groups the perceived details, possibly temporally dispersed, into one bigger perceived picture. It is this movement outside human time and the resulting loss of physicality that forms a texture sensation.

In the general case of composition, one could say that composers are usually interested in 'abusing' the technology, touching new grounds and finding new ways to organize sound – or even finding new sounds themselves. In acoustic music, for example, the increased interest in the use of experimental extended techniques to produce a whole array of previously unknown or unconscious sounds can be given as an example.⁴⁷ In the digital world, post-digital artists push the technology to its limits and use the whole range of sounds that emerge from provoking the error (Cascone 2000). Within the aesthetics presented in this section, composers are not so much concerned with creating 'realistic' sound events, as they would rather be in section 3.3.2, but rather interested in presenting the listener with spatial images which do not occur in nature (Baalman 2010). Unfortunately, quoting Damian Murphy, Otondo (2008) concludes that this has not happened in the spatial domain so far: "[...] the technology is there but there has not been a development in terms of artists pushing the boundaries." Nevertheless, the approaches introduced in this section lean towards such an attitude in the technological usage. In a way, the composer creates spatial sounds that emerge from an extreme usage of the space available, reaching into the marginal details of the technology and balance along thin lines of many possible parameters.

One example that, at least moves towards pushing the boundaries of the human perceivable and possibly the machine computable is the approach used in Hagan (2008), she describes as *textural composition*. In her composition *real-time tape music III* (Hagan 2008), she constructs eight granular textures by stochastically sampling from ten different source files using Gaussian processes for the texture generation and Markovian stochastic processes for the higher level organization of the different textures. Due to the high computational effort required, each texture is sent on a simplistic circular trajectory at random speeds, with an additional copy of the texture located constantly at 180°, fading in and out of phase to give the impression that the texture is constantly jumping through the listening space. Finally, she employs a variable reverberation algorithm on each grain to avoid any constant sense of distance for any texture. All the randomization put together creates a dense soundscape of grains around the listener, constantly changing in detail, but maintaining relatively stable on its surface. Trying to concentrate on one single stream is nearly impossible due to the stochastic nature and rate of change in

⁴⁷The interested reader might want to look into the work of Helmut Lachenmann in this case.

many dimensions beyond the spatial ones. The perception resigns from focusing on the minute and, instead, is steered towards the whole. For this reason, Hagan concludes that the whole soundscape can be seen and perceived as one single sound, she calls the *meta-object*. The term *meta-object* is borrowed from the view that sounds are seen as objects moving through space, with the difference that the *meta-object* is no single sound in itself. It is made up of many, individually spatialized sounds that are carefully chosen to balance between being perceptually separated into each individual component and becoming a homogenous mass.

Hagan (2008), here, clearly puts her approach in opposition to trajectorial spatialization approaches. Textural composition is not so much about the quantifiable distances between sound events and a temporally development, as it is more about the internal activity of the meta-object outside of human time. Moreover, she takes care not to be put into the same category as soundscape artists, by delineating her approach from the directionless lo-fi soundscape, as coined by Murray Schafer (Hagan 2008). Instead of becoming an agglomerate soundscape, the *meta-object* should still be perceivable as single sound object, that lies outside of the trajectorial domain. As the lo-fi soundscape is defined by its individual sounds, the dense *meta-object* forms a single sound of its own. Textural composition for Hagan (2008), therefore, is the careful balance between the unintelligible mass of sounds and an enveloping soundscape of discernible events. For example, while the *meta-object* maintains an implicit sense of motion, due to its individual moving components, its greater picture as a single sound remains static. Furthermore, for Hagan (2008), the sound *meta-object* tries to be present in all dimensions at once. In this, it avoids all boundaries and is as circumferential as it occupies all spaces in depth. At the same time, it must take care not to be too densely packed, creating textures and individual grains with care, as density creates a sonic wall that binds the space.

While Hagan (2008) gives a prime example for enveloping spatial composition, the given definition of textural composition is very restrictive and specific to the approach taken. For one, many methods are based on the possibility to create textures from gains and, more specifically, the inherent stochastics connected to this method. The reason why the *meta-object* can exist in the numerous intermediary positions, is due to the equal distribution of each parameter throughout its possible range, making sure that no location in angular position or specific sense of depth through reverberation, for example, is in any way preferred. But her definition is also very limiting, as she concentrates on a subset of spatial composition, taking Smalley's definition of texture (1997) as the exclusive, restrictive basis and framing it into a compositional methodology. The act of composing virtual spaces and spatial sounds does encompass more than the bounds defined through textural composition. This also gives room to question if every spatial sound, a sound that occupies an area of the space without being reducible to a single source, is necessarily a texture. While texture is connected to its illusion to sight and touch, it is exactly that our senses need to be cross triggered, for this analogy to make sense. But Handel (2006) also points out the correspondence problem in texture perception, that is, correspondence is a matter of point of view and distance to allow a certain amount resolution. Gestalts can appear and disappear not on the account of the object changing its form, but on account of the listener. Given the open nature to the problem, the different definitions of

texture as well as the potential to find completely new sounds with unforeseen spatial aspects, complicates the answer. If every spatial sound needs to have a haptic or visual illusion, as (Trevor 1986) defines texture, or if all textural sounds need act outside human time, as in the definition of (Smalley 1997) are both still individualistic, relatively new definitions that still seek a general acceptance. Hence, one should rather refer to the sounds that lack any point source association in more general terms as spatial sounds and regard textures a part in this category.

A spatial sound can have qualities that a normal point source can never exhibit. Being associated as a space in itself, its spectromorphological balance throughout space can have a spatial impact on its own. Using decorrelation techniques in frequency sub-bands, for example, one can create a sense of the space *pulling* the auditory attention to one particular direction. The effect is achieved by playing with backgrounds and foregrounds, first creating a canvas with the original spatial sound, on which the spatial imagery and modifications are then displayed upon. An example would be an opposing temporal decorrelation in each the lower and higher frequencies, pulling or distorting the sound through the listening space as if skewed along the axis defined by the method. But it is not just the spatial imbalance that one creates, but a change in perception of the spatial sound itself. We listen to sounds differently, depending on their apparent direction of origin in relation to our viewing direction. Even though all the frequency content of the original sound is present during the spatial imbalance, the impression of the current sound being different to the previous version stored in the memory of the listener arises due to the spatial dislocation of subcomponents in the spectromorphology. Put in the words of Smalley (2007), the spectro- and spatiomorphology is fused into one super-construct, as the spectral development through time is also spatialized, constituting both morphologies tightly interwoven in one sound. As I consider the space of a sound – its location, movement and arial occupation in space – a fundamental component of its perception, just as its spectromorphology is, I will use the term *perceptual morphology*, to include this property in the spectromorphology, while still delineating each from one another.

Relating to Xenakis' composition *Terretêktorh*, Santana (2001) writes of the *rhythm of timbres*, for example. Similar to the rhythmic connections in trajectorial approaches (see section 3.3.1), where rhythm is implied through the temporal development in a sounds motion, the perceptual morphology of a spatial sound can create a similar sense of musical, rhythmic development. Timbrally implied rhythm suggests spatialization through associations given to the composers personal experience of spatial listening throughout his life. A perceptual morphology, on the other hand, includes the spatial displacement already and changing this property of the sound would change the sound itself. The perception of many, sometimes minute fluctuations on the "canvas" of a fully enveloping sound can be quite rich in activity. What the listener, for example, needs to understand, and what the composer also needs to imply through his music is a sense of defocussing the aural impression of the space, taking and perceiving the whole sound without concentrating on a particular area in the space. Only then is the whole perceptual morphology apparent, as the focus on a particular point (usually a point source acting as an audible "pixel" or "voxel" in the auditory scene) which would only bring the reduced spectro-

morphology of this area into consciousness. The effort is a defiance of the cocktail effect: a listening mode that reduces itself to the general perception of all sounds around oneself. The listener does focus in the usual sense, but concentrates on the overall scene around him and how this scene shapes in audible color if only even the spatial displacement is altered. "If I put myself in the 'musical attitude' and listen to the sound as if it were music, I may suddenly find that its ordinary and strong sense of directionality while not disappearing, recedes to such a degree that I can concentrate upon its surrounding presence."⁴⁸ For example, if one can imagine a noisy bar, and would not focus on the conversation in front of him but would soak up the complete soundscape restricting himself not to tune into any one conversation. For one, he would engage in reduced listening as he perceives not the meaning of the words and noises but only their sonic qualities. Secondly, the noise level of the bar suddenly rises significantly, as the cocktail effect is annulled. But, furthermore, he also engages in spatial reduction as no spatial position is favored and the complete spatial scene collapses into one single, enveloping sound.

Giving the listener the ability to do so requires specific knowledge from the neurological path of our auditory system. Quickly recapping parts from section 3.1.3, Arons (2001) cites studies that show that multiple sources compete for the attention of the listener and simultaneous attention is only possible if the information rate is in accordance to the listeners ability to rapidly switch from one to another. Therefore, two messages with little information content can be dealt with simultaneously. Textural composition, for example, shows little information content in any specific, localized details. It should, therefore, be possible for the listener to focus on the *whole* sound while not being distracted by any particular section of the texture. Similarly to how Pierre Schaeffer based the success of a *emphmusique concrète* piece on its ability to place the listener into the mode he coined *reduced listening* (Chion 1994), the success of a spatial composition, especially with spatial sounds, is then clearly determined by how far the composer is able to give the listener a chance to reduce his spatial listening to the full surrounding soundscape. The original ideas on reduced listening given in Chion (1994) eventually lead to the conscious perception of the spectromorphology, as defined by Smalley (1997). As we already defined the perceptual morphology above, the conclusion is that the listener in his hearing exercise and the audible material has to be stripped down to the perception of the spatiomorphology. In effect, the spatial development needs to be abstracted for it to gain in attention.

Although (Smalley 2007) seems to fuse the temporal perception of a spatial texture into one singularity, the temporal development should not be neglected during the act of composing the piece. In Handel (2006) one can find a description of the already mentioned motion aftereffects. He gives the example of a falling waterfall, that, if one gazes at it for thirty seconds and deviates his view to the landscape aside, will cause the landscape to perceptually rise with the same speed the waterfall seemed to fall. The same happens when staring at a spiral for long enough, that anything else you look at thereafter, like your own hand, for example, will either grow or shrink, depending on the direction of the spiral. Although no studies have been made

⁴⁸Don Ihde, as quoted in (Harley 1994)

on the auditory effects of motion aftereffects, the neurons responsible for this tuning of the perceptual system to a constant movement are in charge for both the visual and the auditory stimuli. A spatial sound, therefore, with uniform movement around the audience, can have a similar, even contradictory effect on the perception of whichever sound may come thereafter. Furthermore, the order of the temporal development important, as, also considering Gestalt psychology, a spatial impression can be altered by a changed spatial occurrence of events: "At one level, motion perception simply is another kind of texture segmentation. There is a change in the visual field such that one region pops out and is perceived as beginning and ending at different locations in the field." (Handel 2006) A texture, in the visual as well as in the audible domain, can exhibit internal motion through alterations of internal structure, being the perceptual morphology, or spatialized spectromorphology of a sound. This notion is important later in chapter 4.2.3, when defining a gesture emerging from the tiny changes in a dense texture.

Composing the space deals with Gestalt principles of background and foreground. It also experiments with a sense of balance and the violation thereof. Deep sounds allude to weight, while their removal causes the aftereffects to heighten a region or the listener himself. Imagine, for example, a seesaw that would flip from one side to the other. Now have the balancing bar be of some sort flexible material that would cause the upward and downward movement from one side flow to the other (despite the intervening realism of physical laws). The imagination, though, has to be taken farther, away from the visual into the auditory realm. Possibly, this step can be difficult, if one has never experienced a comparable sensation. The association is trans-modal: an auditory sensation, possibly even barely noticed, translated in into a feeling of movement, a deformation and reshaping of space. The final step would be to see this seesaw as some imaginary object with the same flexible behavior but throughout all three dimensions. Then one would arrive at some notion of fluctuation in completely enveloping space. I personally find the thought of a two dimensional disc that waves easier than creating a three dimensional analog, probably because all waves that this metaphor alludes are most likely transversal. Relations between the sound and a transforming mass that is somewhere between solid and ethereal, volatile, but not fluid, are valid in this context. Virtual spaces can be gestalt spaces, in which the psychoacoustics of emergence and reification form spaces that are really not suggested. By employing the principles of foreground and background spaces can be opened up and closed by creating either. As background can be the apparent silence, the mere resonance of the room on one side, while, as mentioned above already, psychoacoustic tricks and decorrelation effects may be used to create imbalances that make a part of the texture stick out of the mass. In total, the background is always defined by that which overpowers the sensation in all dimensions, while the foreground is defined by that which is sparsely used. Multistability occurs, then, when both sensations are not clearly defined and the perception of either flips between both possibilities uncontrollably. Both the audible possibilities of such an effect and the compositional value of multistability are yet to be investigated.

Finally, the idea of spatial organization arrives at the notion of higher level organizational approaches. Referring back to Hagan (2008), for example, trajectories are existent somewhere underneath the level of the overall texture, but the overall composition methodology neglects

their existence by leaving them to be governed by stochastic processes. The control of the composer is located beyond the microstructure of the piece and is more concerned with the overall macrostructure. Artificially constructed spatial sounds are usually groups of individual point sources. It requires new modes of control that shape and modify these sources in a unified way. Spatial sound control is, therefore, a complex method that requires the organization of the sound in its detail by algorithmic means. Although perhaps violating against the distinction made by Smalley (1997), a composer can nevertheless choose to turn a texture into a gesture. The appeal of this approach are the many possibilities with which each virtual point source in the texture may be moved in relation to its grouped neighbors. The texture built in *real-time tape music III*, for example, could be changed by giving the trajectories a areal bias. New forms of spatialization call for transitions between these and the old methods, bridging gaps and constructing a musical language as whole. The more diverse, possibly hierarchical, these categories of spatial imagery become, the more the composer can jump between categories through several dimensions. Although a composer is somewhat expected to stay in his aesthetic frame, the overall language of spatial music will need to be able to fit all approaches under one coherent theory. A newly discovered method does not just open semantic connections within itself, but also within the relationships to the previously existing approaches.

Chapter 4

Spatial Sound Synthesis

”Die ersten Kompositionen elektronischer Musik, punktueller Musik überhaupt, waren äußerst homogen im Klanggemisch und in der Form. Alle musikalischen Elemente nahmen gleichberechtigt am Gestaltungsprozeß teil und erneuerten sich beständig von Ton zu Ton in allen Eigenschaften. Wenn nun alle Toneigenschaften sich beständig in gleichem Maße ändern, wenn nicht eine Eigenschaft für längere Zeit unverändert bleibt, dann eine andere Eigenschaft dominierend wird [...], wenn vielmehr Ton für Ton sich Höhe, Dauer, Farbe und Stärke ändern (Punkt für Punkt), so wird schließlich die Musik statisch: sie verändert sich äußerst schnell, man durchmißt in kürzester Zeit immer den ganzen Erlebnisbereich, und so gerät man in einen schwebenden Zustand: die Musik bleibt stehen.”¹ Stockhausen²

New languages need a new vocabulary. Many approaches to spatial music, in the past and present, have used, and still use sounds today that are of point source type. Furthermore, some sounds might cause connections to memories of possible causalities, which are world bound, have a mass and perhaps don't move much. The discussion in section 3.3.1 elaborately argued a certain validity of trajectory based spatialization techniques, but if a guitar is sent flying through a room, its spatial aspect is easily ridiculed, pushed aside, while the focus concentrates itself on the spectramorphological aspects; the timbre, the dynamics and the pitch of the instrument. If space is to become a musical language of its own, its foundation – the pillars – can not be constructed on the spatial movement of an alien vocabulary from different domains. Movement itself is only heard in the periodic movement of air particles, but as an abstract concept of flying objects it is not the movement we hear, but their sound coming from different areas in space that we connect to a trajectory. A non-sounding object that moves through air, like a hot knife through warm butter, is not heard, therefore, movement itself can not support a musical

¹”The first electronic compositions, or punctual music in general, were exceedingly homogenous in its tone mixture and in form. All musical elements equally took part in the arrangement process and renewed themselves consistently in all properties from tone to tone. If all tonal properties change consistently to the same degree, if not one property stays unchanged for a longer time and then another property starts to dominate [...], or rather, if height, duration, timbre and strength change from tone for tone (point for point), then, finally, will the music become static: it changes extremely fast, one experiences the same events in the shortest time, and, hence, one finds himself in a pending state: the music stands still.”

²As quoted in Zelli (2001)

language alone.

What is necessary for space to become a musical language are sounds that are inherently spatial. The relation between sounds and spaces contained in each other in both directions has been philosophized in section 3.2 to a good extent. Here, we want to look at creating sounds that do not just exist in space, but *are* the space. Sections 3.3.2 and 3.3.3 have, theoretically, touched upon this topic from both a real world and imaginary world approach. Sounds that are inherently spatial in nature could constitute those sounds where the individual components are difficult to discern and, together, make up an enveloping soundscape. Most ambiences, especially a noisy bar, or large events, like earthquakes, in which the low rumble of the quake itself in the most cases might be below the hearing threshold, but the agglomerate sound of everything shaking around you are prime examples for spatial sounds, especially of granular nature. As a composer of acousmatic music myself, I am more interested in unknown sounds, though. This chapter will look into practical solutions of constructing new spatial sounds through methods known as *spatial sound synthesis*. In this, the space and the spatialization will be used as part of the synthesis, making the space an inherent part of the sound and its audible perception. This means, that the aim of using spatial sound synthesis is to create sounds that are unthinkable without their space. Good summaries of different spatial sound synthesis techniques can be found in Barreiro (2010), Lynch & Sazdov (2011a). Moreover, these sounds have a necessary enveloping, even engulfing (Lynch & Sazdov 2011b) characteristic, while still being perceived as a single sound. Some have called this sensation a *texture* (Smalley 1997, Hagan 2008), being borrowed from the acoustic domain, describing density and width in terms of melodic, rhythmic, and harmonic material, but also, in a similar matter, describes the temporal reduction beyond the human noticeable. Based on this, it seems that it is best applied to spatial granular synthesis, for example, and its general use for all spatial sounds could be disputed.

One common characteristic of these experiments, at least in my own approach to these problems, is a technique that makes use of *virtual speakers*. Virtual speakers are a great method to abstract a sound system and overlay it onto another. One defines a point source in any virtual of the sound spatialization system and regard this as the actual speaker. In a perfect reproduction system, there should, theoretically, be no difference between a virtual speaker and accessing a physical speaker directly. For creative purposes, this technique gives way to higher order gestures that give the composer the ability to move the complete set of virtual speakers in a uniform manner. If the spatialization interface allows grouping of source, this can be done with a single command for all speakers. On the other hand, a composer might define transformations reordering the whole set of speakers. The idea is, that the spatial sound is defined by all virtual speakers together, being somewhat decorrelated by using a particular technique. The ability to look at the effective speakers abstractly and define them independently from the physical, stationary ones, now being able to move the speaker itself through space, opens up a complete new dimension of possibilities to the already existing spatial synthesis. For example, rapid panning modulation synthesis, which would normally pan the sound in a circle, can easily pan the sound along other trajectories by simply reordering the virtual speakers instead of somehow

changing the signal that pans the audio in a complicated way (see details in 4.2).³ A spatial synthesis technique can easily be rotated and juxtaposed against other spatial sounds, which would be difficult if accessing the physical speakers directly. Solely thinking about the possibility what would happen to the perception of a spatial sound, if it were to move into the distance is an interesting notion.⁴ Of course, on the other hand, no spatialization technology is perfect, and assessing the perceptual difference between virtual and physical speakers could constitute a future study.

An extensible framework for the handling of spatial gestures comprising multiple virtual sources is presented in [Schumacher & Bresson \(2010\)](#). The framework, name *OMPisma*, is written in OpenMusic, while the sound synthesis is done in Csound. While gestures can be easily implemented, its main aim is the generic handling of spatial sound synthesis methods. Paths are drawn by the composer in a graphical user interface. It supports several dispersion technologies in up to three dimensions, and includes an HRTF rendering possibility. Presented spatial synthesis examples include two feature analysis examples, where either the fundamental pitch information from a piano or the onsets of a tabla are computed and used to create statically spatialized clusters and rhythms. A more interesting example is a spatialized additive synthesis approach, in which they were able to create over 100 sine tones, each with their own frequency and spatial trajectory. Other sound source organization frameworks found in [Schumacher & Bresson \(2010\)](#) include the IRCAM *Spatialisateur*, focussing on perceptual parameters, the *ZKM Zirkonium*, allowing grouping of sources, which each group being moved independently ([Normandeau 2009](#)) or the extensive high-level control offering spatialization authoring tool *Holo-Edit* ([Petersa et al. 2009](#)).

One rather popular method, which was not fully investigated, due to the temporal limitations of this thesis, is spatialized granular synthesis. In a nutshell, it extends regular granular synthesis while also making use of spatial parameters for each grain. That is, next to the envelope, size and loudness parameters, among possibly others, the spatial position is calculated for each grain individually, and creates not just a stream but a cloud in two or three dimensional space. The organization of grains becomes the main issue here and a popular approach to the problem is the use of particle systems, where each particle, behaving according to specific physical or sociological rules, is tied to a grain of sound. The most used particle system to be found in literature is the boids algorithm by [Reynolds \(1987\)](#). Previous to this thesis, I had built my own implementation of the boids algorithm in a C external for Max/MSP, based on its principles and including a few experimental behaviors I could control via parameters. For this work, I mapped the two dimensional movement (x and y) of the individual boids to the surface of the sphere (angles θ and ϕ), using the length $l = \sqrt{x^2 + y^2}$ in the following equations:

$$\theta = a \sin\left(\frac{y}{l}\right), \quad \phi = a \sin\left(\frac{x}{l} / \cos(\theta)\right)$$

³Unfortunately, this perceptual effect of this idea could never be studied, due to the limited amount of time one has carrying out a master's thesis.

⁴Also here, unfortunately, due to no proper distance algorithm implemented at the time of the thesis, this effect has not been tried, yet.

In the few experiments done, the movement of a total of eight boids were mapped to a granular synthesizer built with the FTM/Gabor library for Max/MSP⁵. Changing the cohesion and separation behavior clearly influenced the size and width of the cloud. Experiments with a three dimensional version of the algorithm mapped also to the distance could unfortunately not be made. The movement was rather erratic and somewhat uncontrollable. Placing attractors into the listening space focussed the cloud almost immediately to this point (based on the speed and attraction radius of the boids). Extreme separation coefficients caused the cloud to separate itself into several individual streams, enveloping the listener. Interesting effects were made dispersing the boids around the whole listening space and using only very short and sparse grains, causing differently colored clicks to appear, reminding myself and other test subjects of the sense of being in a tent during a rainy day. The interested reader may be pointed to the previous research done by [Kim-boyle \(2005\)](#), who also heavily investigates particle systems, such as also mainly the boids algorithm, but also other physical particles, [McLeran et al. \(2008\)](#) using dictionary based methods for particle placement in space or ([Einbond & Schwarz 2010](#)), employing spatialization to corpus based concatenative synthesis. Finally, the master thesis of [Bates \(2004\)](#) gives an in-depth investigation of different method for spatialized granular synthesis, and can be recommended as a good starting point.

4.1 Timbral Spatialization

Timbral spatialization techniques are ways of dispersing the a sound in space by placing its frequential sub-bands in respectively different positions. The most common ways of doing so is by using an FFT and output each band individually, or by filtering the sound with several bandpass filters. In an early implementation, for example, [Torchia & Lippe \(2004\)](#) used FFT methods to subdivide the signal into different bands. They implemented a two dimensional interface with two [multislidder]-objects to draw on, each corresponding in the number of frequency bins used (64). The sound could then be spatialized in each bin individually at runtime. Furthermore, they discuss methods to algorithmically spatialize the audio using feature analysis, but more interestingly they propose a spatial cross-synthesis, where the analysis of one sound would govern the spatialization of another. They point to one downside, though, that the technique is only concerned with spatial energy distribution and silence will still have the original file playing, but in mono. A prolific practitioner in the field of spatial sound synthesis is [Kim-boyle](#). In his earlier paper [2006](#) he built on the work of [Torchia & Lippe](#), but used the boids algorithm to determine the spatial location of each bin. In a later paper ([Kim-Boyle 2008](#)) he would also include other particle system, like the `jit.p.vishnu`, `jit.p.shiva` and `jit.p.bounds`, exhibiting possibilities to spatialize the bins by simulated clouds of smoke or foam models. Moreover, his paper ([Kim-Boyle 2008](#)) discusses spectral delays, a technique by which a separate delay can be applied to each bin, effectively decorrelating the bins from one another.

In a later technique, [Barreiro \(2010\)](#) talk about his use of spectral spatialization algorithms in his compositions. His implementation is a fixed 8 channel FFT separation patch, but can

⁵http://ftm.ircam.fr/index.php/Main_Page

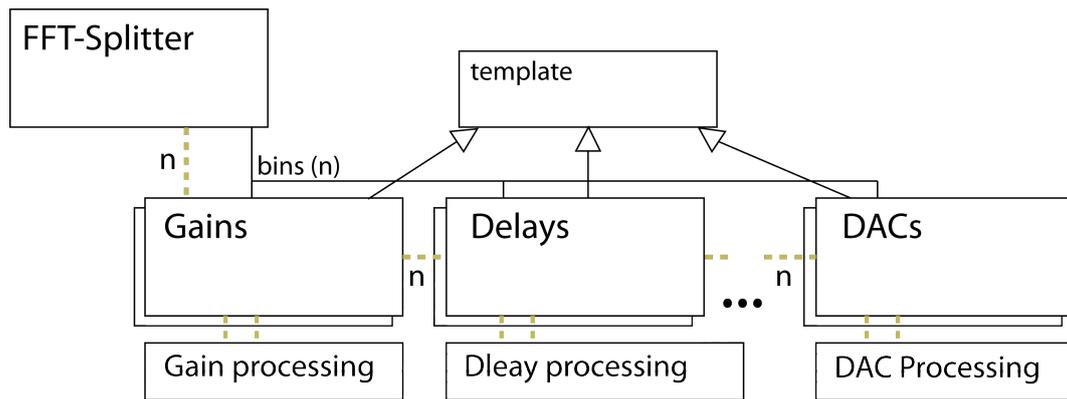
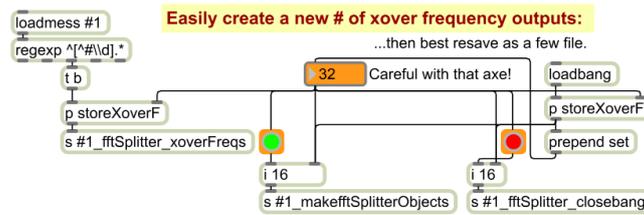


Figure 4.1: Diagram depicting the overall structure of the Spectral Splitter. The example shows a possibility using Gains and Delays, but, theoretically, any effect or processing should be fairly easily integratable, hence the three dots.

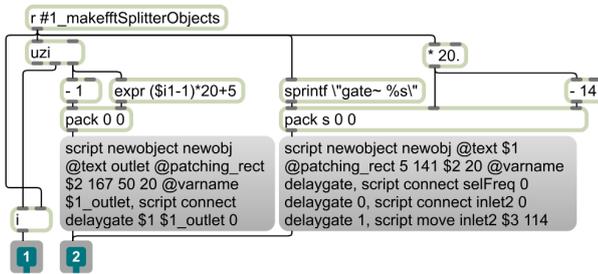
take a stereo input signal. The two input channels seem to be output in alternating channels. [Barreiro](#) reports a strong dependence on the input sound for the success of the technique: "In general, sounds with a broad spectral content tend to sound diffused, providing an enveloping sonic image. Sounds with energy concentrated on specific regions of the spectrum, on the other hand, usually sound more localised." ([Barreiro 2010](#)) Moreover, the windowing functions in his patch are flexible, allowing for overlapping channels in FFT bins, or gaps in the frequency spectrum. Furthermore, his implementation allows overall affecting gestures, being able to rotate the whole circle of virtual loudspeakers for example. In a similar fashion, [Gibson \(2009\)](#) built a FFT based delay with a dynamic grouping of bins in real-time. Although his approach does not include any notion of spatialization, it is still very interesting due to the possibility to decorrelate each bin from another. Unlike the method described in [Kim-Boyle \(2008\)](#), [Gibson](#) applies the delay outside the FFT, isolating the FFT separation from the effect processing. Decorrelation is very important in this context, because, unlike one might easily think that different bins spatialized in space would result in different frequency ranges in space, the effect is only slightly enveloping and hardly noticeable. Despite the fact that each set of frequencies of the source are coming from all different points in space possible, the brain is still able to fuse the sound into a cohesive whole. Experiments done in the studio indicate that the localization is based on the spectral mean of the overall sound. That this is a common misconception is just recently stated in a paper by [Kendall & Cabrera \(2011\)](#).

4.1.1 Implementation of an extensible Spectral Splitter

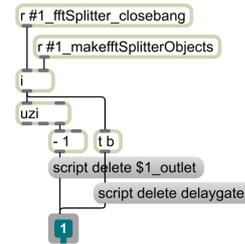
Spectral processing in space seems to be a rewarding approach to create enveloping sounds. Nevertheless, most approaches described above seem very static and inflexible in the possibilities to experiment with different creative ideas. To further facilitate the demand for experimentation, an extensible spectral splitter was built in Max/MSP. The focus was to lie on easy handling of a variable amount of sources so that the effects and controls would automatically adapt. A schematic outline of the design is given in figure 4.7. There was a significant limitation due to



(a) Interface to create new number of outputs



(b) Script to create new outputs.



(c) Script that deletes old outputs

Figure 4.2: Diagram depicting the overall structure of the Spectral Splitter. The example shows a possibility using Gains and Delays, but, theoretically, any effect or processing should be fairly easily integratable, hence the three dots.

the way Max/MSP handles its [pfft~]-objects, which is that no [fftout~]-object may effectively be added during runtime. Originally the idea was to implement the complete FFT-Splitter in one abstraction, but, due to this shortcoming, everything had to be designed orienting itself based on the fft-patch given. This means, that the user is not able to specify the number of frequency split sources via a parameter, which was the original and easier to use solution. For a differing number of outputs, the user, therefore, needs to create a new fft subpatch by hand, unfortunately. As arguments to the overall FFT Effect, one has to pass on a unique name for the whole system and the name of the FFT subpatch to be loaded. Based on the number of outlets contained in this subpatch, all other patches should adjust themselves accordingly. These other patches are all based on a template design which would make it easy to reorder or even include other, completely new effects. The beauty of this approach is that the effects only need to be prototyped for one channel as a Max/MSP abstraction and, depending on the number of outlets used with the specific Spectral Splitter, would automatically link into every distinct frequency range.

To create a new number of outputs, an interface (see figure 4.2a) is hidden in each of the original Spectral Splitter objects.^{6,7} Changing the number of outlets in the raw FFT patch is as simple as dragging the number in the number box up or down – or entering a number, of

⁶The original objects are most likely called *fftsplitter16~.maxpat* or *fftsplitter32~.maxpat*,... respectively.

⁷The reason why it is hidden on "locked" mode is that of error prevention. Once the [pfft~] abstraction is loaded into the main patch, one does not need this interface anymore, because it is useless anyway as of the instantiation of the [pfft~] when a dynamically added outlet is not registered as an actual outlet anyway. To prevent the user from altering the number of outputs after instantiation, it is hidden, since the abstraction can not be unlocked when loaded as such. Dynamically deleting outlets causes the outlet to still exist in the [pfft~]-object, but no [fftout~] can point to this outlet anymore. The whole patch would have to be restarted!

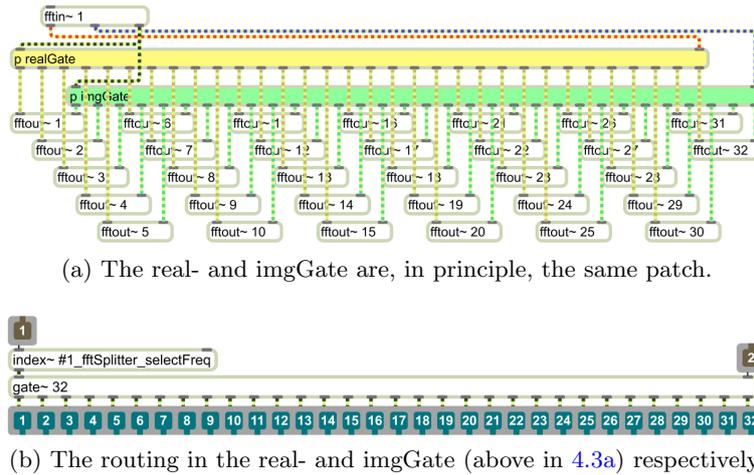


Figure 4.3: An example splitter with 32 FFT-outlets. Note that all the outlets and connections are done via scripting.

course.⁸ Upon release of the mouse, a message will first go to the delete script (4.2c) which creates a message intended for a [thispatcher]-object, deleting the amount of outlets based on the previous number. The new number of outlets is then stored in the [int]-object for preparation for the next change. Then, the current number of desired outlets is sent to the creation script (4.2b), again, building a message intended for a [thispatcher]-object, which not only takes care of the outlet creation and connection, but also reorders them for a somewhat structured look for better organization and programming style. The patcher [p storeXoverF] contains a simple message for storing the number of outlets upon saving Spectral Splitter under a new name. It is important to recall this number as, it is sent to all effect patches to create and connect the correct number of outlets (far left of figure 4.2a). This is done upon receiving an argument, i.e. when the object is instantiated, instead of in edit mode. The distinction is done using the [regexp]-object to check if the sent argument is a real argument or still the placeholder '#1'. The final structure of the Spectral Splitter would look something like figure 4.3. Figure 4.3a shows the connections to each [fftout~], while figure 4.3b demonstrates how each frequency bin is actually routed. A buffer identified by the unique name given to the top-level subpatch is filled with values that link each bin index to a outlet of the [gate~]. Determining the values for the buffer is shown in figure 4.4. This method was inspired by the solution made available by Gibson (2009). The idea is, that based on [multislider] of the the UI (center) it calculates the difference between each successive value. Putting it into relation to the bin width, it determines how many bins belong to each output of the [gate~], assuming an ascending order. The user is restricted to use the small slider on the top to keep the sliders in ascending order and prevent negative differences.

Once the [pfft~]-subpatch with the desired number of separated frequency ranges was created, it can be loaded into the main abstraction. Figure 4.5 shows an example of how this is done. The main high-level abstraction is a distinctive file on its own, since each effect needs to

⁸If not careful, a ridiculous amount of objects (IFFT processes!) may be created, slowing down Max/MSP and perhaps crashing it!

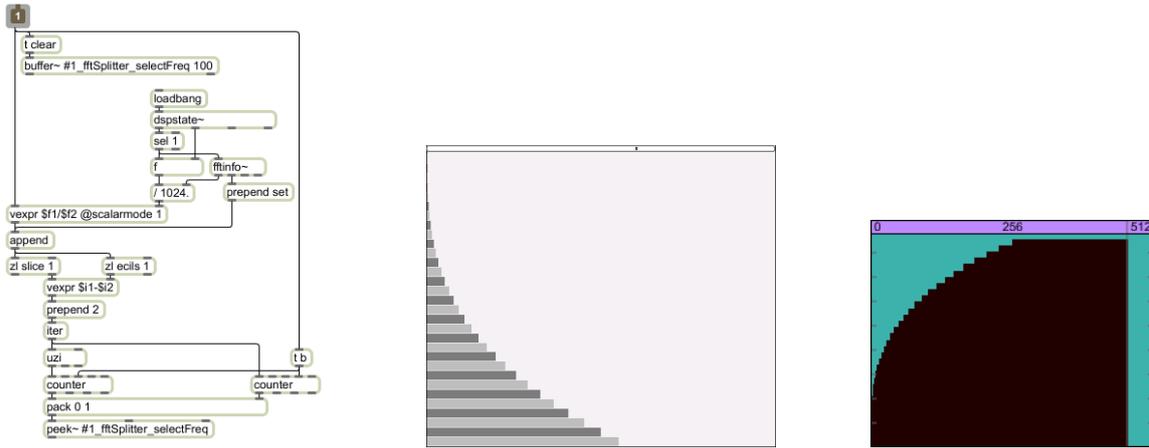


Figure 4.4: The subpatch responsible for filling the cross-over buffer controlling the [gate~] in figure 4.3b. Principle ideas came from a solution provided by Gibson (2009). Shown also is the user interface, and the resulting contents in the buffer for each of the 512 bin in a 1024 FFT.

be stored therein. For this reason, the object is called [3DspectralDelay]. Adding other effects will be dealt with further on. The first argument is the unique name (i.e. "bm.spec"), followed by the name of the Spectral Splitter file. Once created, the Spectral Splitter is loaded into the [pfftdim]-object, from which the number of outlets is extracted and sent to each of the effects (i.e. [p gain], [p delays] and [p sourceout]). Each of these subpatches work exactly the same, with the difference that each load a different abstraction as an effect.⁹ The order of events strictly must create the outlets first and then try to connect them to the object above. Each object receives a message from its higher located predecessor that its outlets are created, for this task to go smoothly. The connections are done by sending the messages to the [thispatcher] object in the same higher-level patch (see figure 4.5b, bottom object).

If we have a close look at the effect subpatches, which are all the same, basically, we

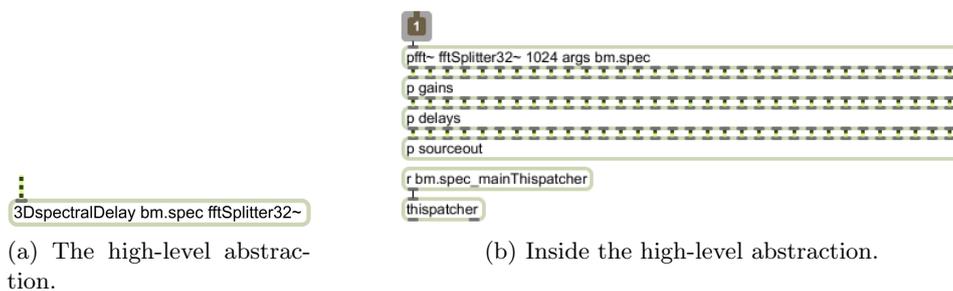
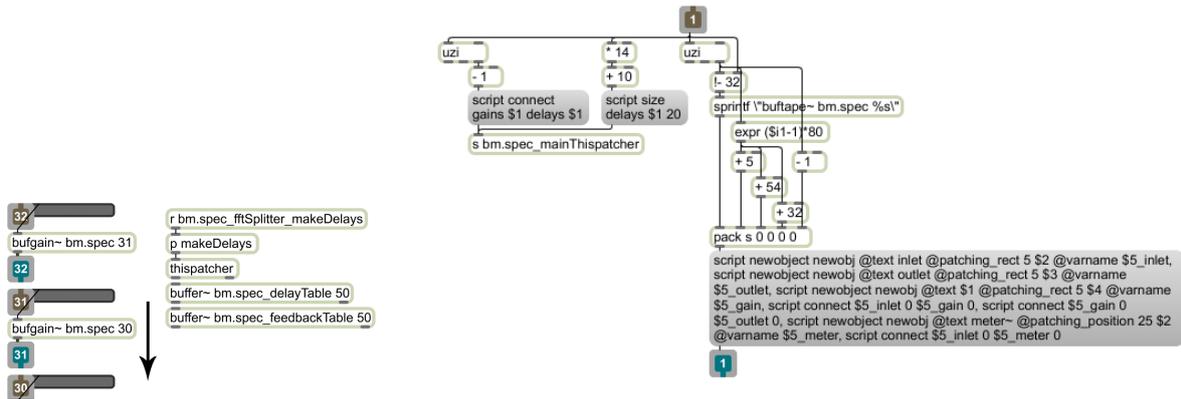


Figure 4.5: Showing figure 4.7 in an actual programmed version in Max/MSP. The high-level patch is called [3DspectralDelay] due to its functionality programmed inside (4.5b). Adding or changing functionalities on each respective bin is perfectly possible and the new patch would most likely be saved under a new name. *bm.spec* is also just an identifier in both figures and can be changed as a parameter in the main abstraction object at free will.

⁹[p sourceout] has a special position here, as it does not create any outlets. Whoever is experienced with Max/MSP scripting can change the functionalities to whatever they desire, really



(a) Inside the delay subpatch. Based on the number of FFT outlets, it creates one inlet, processing and outlet for each (descending downwards).

(b) A closer look inside the makeDelays subpatch.

Figure 4.6: The delay subpatch in detail.

can see the Max/MSP code for the delay example in figure 4.6. Upon receiving the number of outlets from the previous patch (the gains in this case) it creates a block consisting of an inlet, a processing abstraction and an outlet, for each outlet of the original Spectral Splitter. The buffers in this example serve as tables for all delay processors. The `[p makeDelays]` subpatch itself reveals the scripting done and the method to connect all the inlets to the previous effect (again, the gain) by sending it outside to the main `[thispatcher]`. If a user wants to create a new effect, they should best copy an existing one and replace the name of the effect in the `[sprintf]`-object with their new processing method. Then the scripting for the connections on the left of the patch in 4.6b should be changed. Unfortunately, this requires renaming the main effect patch’s scripting name to something fitting and exchanging it in the appropriate field (in this case, everywhere, where `delays` is specified). Furthermore, for the connections to be in the right order, the name of the patch containing the previous effect should be replaced as well (in this case, everywhere, where `gains` is written).

4.1.2 Conclusions on the Spectral Splitter

Continuing right from the end of the last section, it is clear that the adding of new effects is not the most user friendly yet. Ideally, a composer or sound designer should be able to craft their effect processing module as a prototype (or *abstraction* in the Max/MSP jargon) and specify its use in particular projects with just one command. Furthermore, coming from a usability direction again, I would like to see an update with Max/MSP that solves the issue that `[fftout~]`-objects can not be added at runtime in the `[pfft~]` environment. Ideally, the whole Spectral Splitter should just work by calling one and the same main abstraction, specifying a name, the number of split frequency ranges and its method, as well as all effect abstractions in order. The type of frequency splitting is important, because it makes a huge difference if you split the signal using an FFT or band limited filters. Although the solution using filters has not

been tried yet, it is very attractive, as it allows for more parameters to be changed, such as the center frequency and width of the band.

Furthermore, another important issue one has to deal with is the actual distribution of each frequency limited channel. Being a spatial synthesis method, clearly the spatial arrangement should be at the forefront of the research. Well, one clear setback, in the case of using FFTs to split the source into frequency limited bands, is that the impression is not at all what one might first suspect. Even though there is no cohesive version of the source sound coming from any one direction, the brain will still be able to fuse all the partial frequency bands together to one sound. This effect was just recently documented in (Kendall & Cabrera 2011). The sound will feel somewhat more enveloping, but, surprisingly, will sound almost as if it were coming from a point source. My assumption is that the different frequencies, being coherent within one another, function as something similar to early reflections, hence the subliminal spatial feeling. First experiments suggest that the associated location stands in relationship of the spectral centroid and its current position in space. Although, this is not perfectly clear, as the analyzed centroid was jumping around on screen much more frantically than the audible impression of the supposed location of the spread source. If the ear is insensible to such erratic movement in this context or if the localization cause is of a different nature is unclear. What is clear is that this slight envelopment does have a characteristic. That is, if the sources are all moved simultaneously, one can start perceiving the sound in a ghost-like manner. The impression of some morphing process that is suddenly happening is strong, even though the sound is still localizable. What is clear, though, is that this technique would benefit much from different decorrelation techniques, which is what this Spectral Panner was to investigate into and what it was built for.

As outlined in the start of this section 4.1, there have been many different approaches, such as drawing the spatialization (Torchia & Lippe 2004), chaotic spatialization, using the boids algorithm or other particle systems (Kim-boyle 2006, Kim-Boyle 2008) – but during this thesis, a more static and overall controlling method was used, such as in Barreiro (2010). Inspired by production techniques, the initial idea was to locate the lower frequencies in the center and spread the higher frequencies around, becoming more thinly spaced as we reach the higher hemispheres. The 32 sources were grouped into groups of $k_i = 1, 3, 5, 7, 9, 7$ sources. They were then distributed with the following formulas:

$$\theta_i = \sum_{j=0}^{k_i-1} \frac{j}{2a_i} 60b_i - 30b_i \quad , \quad \phi_i = \sum_{j=0}^{k_i-1} \left(\frac{\sin(\frac{\pi j}{2a_i})}{2} + 0.5 \right) 10i$$

with $i = [0, 5]$, and $a_i = 1, 1, 2, 3, 4, 3$ and $b_i = 0, 1, 2, 3, 3\frac{2}{3}, 4\frac{1}{3}$ being non-linear scaling factors. Further on, dynamic scaling factors were then multiplied to both these starting angles θ and ϕ . Wrapping the product around 360° would send the sources on non repetitive paths, allowing for ever differing, but symmetrical patterns to appear in the spatialization. This is one way the tone color changing property of the different spatial positions was noticed, even though the source was fused locally. It was as if the room was changing, probably due to the subliminal

enveloped feeling, but affecting the source sound directly – as if the source sound was part of the space.

4.2 Rapid Panning Modulation Synthesis

”Für Kreis-, Spiral- und andere periodische Bewegungsformen verwendete ich eine Rotationsmühle. Irgendein Mikrophon oder ein Gemisch von Mikrofonen konnte man auf den Eingang einer solchen Mühle legen. Die Mühle hatte zehn elektrische Ausgänge, die man mit beliebigen 10 der 50 Lautsprecherkanäle verbinden konnte, und wenn man mit der Hand den Steuerhebel wie eine Kaffeemühle links oder rechts herum drehte, bewegte sich der Klang entsprechend im Raum. Die höchste Geschwindigkeit betrug ungefähr fünf Perioden pro Sekunde.”¹⁰ (Stockhausen 1964)

The relationship between musical spatialization and the circle is integral. The circle is a basic geometrical shape, admired for its unreached perfection. Trochimczyk (2001) identifies a row of symbolic meanings, ranging from the social circle of friends, to the abundant use in many mystic and religious contexts. Canons are usually cyclical, and early composers using multiple ensembles preferred to place them in approximations of circle. Most of Stockhausen’s work consisted of circular arrangements and circular to spherical speaker setups in which the rotation of sound, its direction and speed, was a crucial aspect to his musical thinking. Sound spatialization technologies and dispersion software usually employs a spherical coordinate system for source determination, and aural perception is distinguished in angular and distance cues. We move our head, or turn our whole body around our own axis and the whole world turns in circles around us. The circle is a substantial part of our spatio-musical organization and, since the sounds tend to move in circles, also in its perception.

A sound can move in circles around the audience with varying degrees of speed. Using his rotation mill, a device to rotate a mono input over ten outputs, Stockhausen could have sounds fly through the German Pavilion in Osaka at around 5Hz (see quote above). The question is, is there a limit – technically, perceptually, physically? Sounds that are tied to a sounding object,

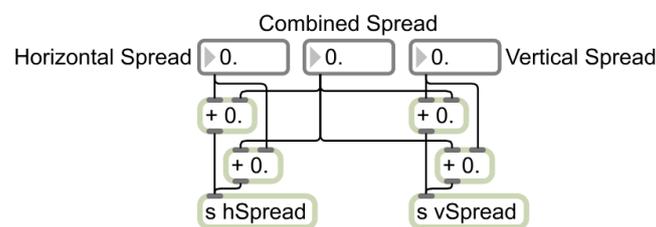


Figure 4.7: UI controls for the scaling factors, enabling smooth control in any angle or a combined gesture, without jumps when switching.

¹⁰”I used a rotation mill for circular, spiral and other periodic movement. You could connect any microphone or a mixture of microphones to its input. The mill had ten electric outputs which you could connect with 10 to 50 loudspeaker channels, and the sound would turn accordingly, if you turn the control stick like a coffee mill to the left or right with your hand. The fastest speed was about five periods per second.”

such as an organism, an instrument or a loudspeaker, for example, do possess a physical limit to the speed they can travel at. Moreover, Féron et al. (2009) have determined a perceptual limit of around 2.5 rotations per second, above which a rotation will not be perceivable as such anymore – half the speed that Stockhausen’s rotation mill was already able to achieve. But virtual sounds are not tied to any physical mass and clearly something will have to be perceived. Experimenting with this notion in the studio, I encountered a technical limitation. Since everything, audio and control messages, were sent over the network, the way one would spatialize and move a sound source through the VBAP space would normally be via OSC messages. These messages need to be created in Max/MSP on my own machine, which happens in the Max environment. The smallest rate at which Max operates is 1ms, which means that at a maximum rate of every millisecond the position of a source can be changed. The math to determine the maximum speed, the Nyquist frequency, so to speak, at which a sound can travel around the listener: If a sound ”traverses” half a circumference every millisecond, it would achieve 500 ”rotations” per second – although it would really only jump between two points in space, ignoring the jitter from the network and similar. Any bigger distance between that one millisecond threshold would have the source travel the opposite direction until it would come to a complete stop at exactly 1000Hz.

But, the sky is the limit, and in this fashion I looked for alternatives for enabling sounds traveling at ”unlimited” speeds. Panning in stereo at high speeds seems to be a known but unused technique. In the official code examples found in the Max5 environment¹¹, one can find a piece of code, spatializing audio on sample basis. This means, the source signal is panned between two speakers using a second signal as the panning source. This is where the notion of virtual loudspeakers comes most prominently into play: direct access to the physical speaker is impossible, due to the abstraction layer of the network interface. I decided to build a static set of sources equally spaced in the dome around the listening space, which I would regard as the speakers between which I wanted to pan to. Implementing an internal VBAP algorithm that would operate on signal basis was not viable in the time given, so I decided to take inspiration from the stereo panning example in the Max5 environment. For ease of treatment, the audio was not panned in a VBAP fashion but strictly in two distinct dimensions. The virtual speakers were to be setup in multiple circles around the audience, with the sound being panned within and between each circle. For this task to achieve the biggest success, I decided that the maximum amount of speakers would give the highest resolution and, hence, the highest quality in panning. Out of experience with the network interface at the time, I was limited to around fifty channels that could be sent over the network, before the dropouts became too noticeable.¹² The next question that arises is how to best distribute all speakers in the room, moreover, how many speakers should be assigned to each respective circle. It is clear that this two dimensional amplitude panning approach assumes the speakers to be at equal distance to

¹¹www.cycling74.com

¹²Note, that this notion is only applicable for live settings and experimentation, as it was the case here. Audio recording and engineering would not survive a slight – or even any click every five minutes, on average. Therefore, this solution was not ready to be captured on tape as yet, unfortunately. The engineers at Barcelona Media were working hard on the improvement of the interface at the time.

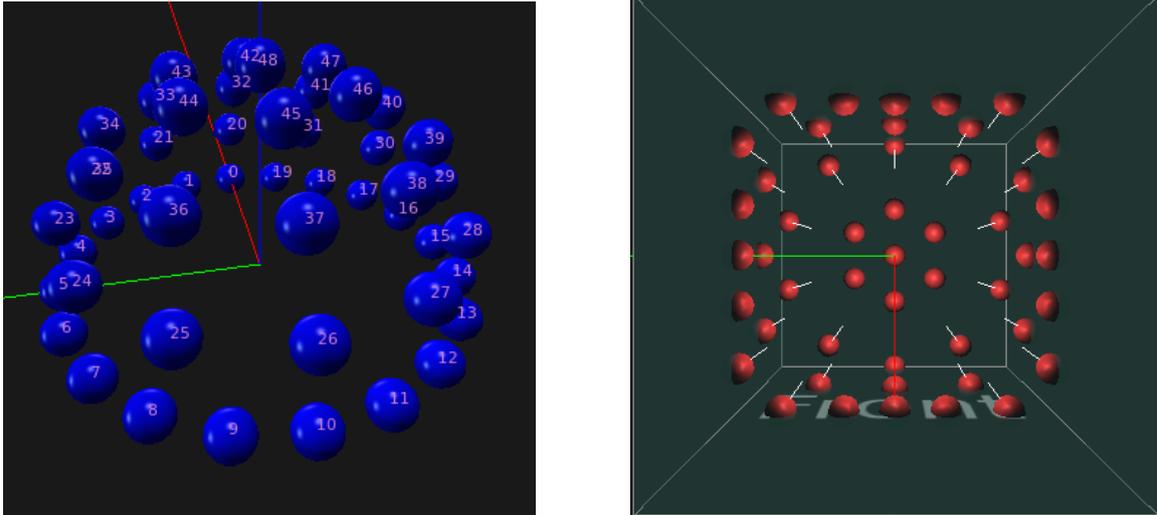


Figure 4.8: Two visualizations of the virtual speaker setup for the rapid panning modulation synthesizer.

each other. Researching into this problem made me realize that there is no full mathematical solution to this problem (Rusin 1998). The problem is the ill-defined property *evenly-distributed*, as well, as that solutions exist as recursive, approximating functions that would distribute the points in a stochastically equal distance to one another. Since equally spaced points on separate circles are needed for this purpose, the decision was made that this task was up to the composer. The solution in my case, using fifty loudspeakers, is shown in figure 4.12 and consists of five circles in 20, 12, 10, 6 and 4 loudspeakers.

4.2.1 Implementation of the Sphere Panning Spatial Synthesis Module

Similar to the Spectral Splitter (section 4.1), the Sphere Panning Spatial Synthesizer (*Spynth*) was aimed to allow easy usage and, in effect, speed up experimentation with this new synthesis technique. The core concept was inspired by the [pan2S] example found in the Max5 distribution. Essentially the audio was to be panned in all directions using two control signals, one for the azimuth and another for the elevation, using equal power panning between each successive speaker and circle respectively. Figure 4.9 shows a coarse outline of the structure and buildup of *Spynth*. The main abstraction is instantiated, giving it a unique name, a list of speakers per circle (passed in parentheses as a single *symbol*) and a possible offset (figure 4.10), if the first channel 0 is not the first channel used in *Spynth*. That said, *Spynth* assumes all channels needed are in successive order starting from the lowest circle and going around each circle respectively. The main abstraction takes three signals as inputs: the source audio to be panned and the two control signals, each within the range between $[0, 1]$.

Based on the number of elements in the list of speakers per circle, *Spynth* creates that

¹²Unfortunately, there exists no standard modeling language for Max/MSP and the diagrams are freely interpreted, taking UML as an inspiration.

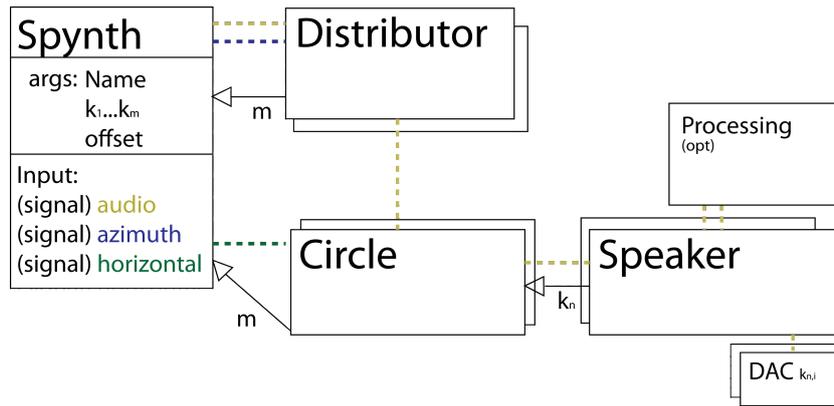


Figure 4.9: Diagram depicting the overall structure of the Sphere Panning Spatial Synthesizer *Spynth*. White arrows here denote *contained in*, as in *subpatch of*.¹³

number [circle] and [distributor] abstractions automatically (see figure 4.11).¹⁴ The [equalPowPan] patcher contains the code inspired by the [pan2S] Max-example to create a buffer with an equal power pan curve. This buffer is used in every panning operation, between every speaker and circle. [p levels] is a monitoring system, which automatically creates ordered [meter~] objects for visual feedback of the module. All the incoming audio signals are sent to a simple subpatch that routes the audio and control signals via [send~]-objects. The elevation control signal is merely multiplied by the total number of circles, which is sent from the [makers] subpatch. The [makers]-subpatch is responsible for the counting and creation of the distributors and circles, similar to the scripting techniques demonstrated in section 4.1.

The distributor is responsible for sending the source audio signal to the respective circle(s), based on the value of the elevation control signal (figure 4.12a). Each distributor takes care of sending the amount of source signal to its respectively assigned circle only. The actual selection of a sub-sectional control value is done in the [p scaleControlSignal] patcher (figure 4.12b). Based on the previous scaling of the elevation signal upon entering the Spynth module, the sub-section for each distributor is selected by its circle index and limiting the shifted elevation data to [0, 1], denoting silence at the boundaries. The actual circle abstraction (figure 4.12c) does not do much on its own, but creates the speakers for this circle and passes on a scaled version of the azimuth control signal to each.

The speaker (figure 4.13) works similar to the distributor. Instead of the elevated control signal, of course, it filters the azimuth control signal to window the source audio within its range.

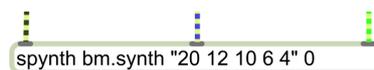


Figure 4.10: An application example of Spynth, being named *bm.synth* with the same setup of rings as described above, in section 4.2 (no offset). This is all there is to it, it is all set up and ready to rock.

¹⁴I did notice I have *distributor* spelled wrong in the abstraction name, but have had no time to correct this mistake. Since no user would need to interact with this abstraction, this issue is fairly irrelevant

Spynth

Sphere Panning Spatial Synthesizer

© Timothy Schmele
c/o Barcelona Media

Spynth can pan audio in 3D spherical surround to a set of (virtual) speakers using two signal vectors as azimuth and horizontal coordinate angles. The panning is done through simple 2D stereo panning. The speaker setup is assumed to be made of an arbitrary number of circles, each containing a desired number of speakers. One algorithm takes care of the panning in each circle (azimuth) while the other distributes the incoming audio to the correct circle in height (horizontal). The lowest circle around the audience needs to be defined first, the highest above last. All output need to be in succession from the lowest speakers to the highest linearly around each respective circle.

- The arguments the Splynth are:
1. Unique identifier of this panning object instance
 2. List of speakers in each circle. Place within quotation marks when writing the argument. Each number in the list stands of one circle respectively with its number of speakers.
 3. (optional) Total system channel offset (if the channel for the first speaker of the complete system does not start with 1)

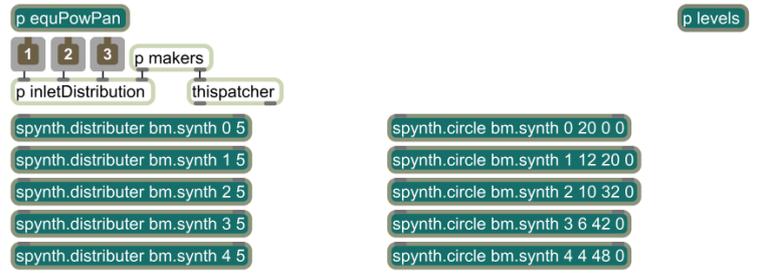


Figure 4.11: The main window of Splynth, here with 5 circles defined.

The other difference is, that unlike the elevation oscillating between each circle, each circle must pan the audio circular, meaning that the last and first speaker in each circle must match up. This is done using the modulo operators in figure 4.13b. Especially the [%~]-operator ensures, that the first speaker, when starting out below zero, will wrap around and start where the last speaker ended. The speaker are defined in steps of a half, overlapping each other fully, i.e. the left neighbor starts playing when the right has stopped. For this reason, the modulo is defined by half the total amount of speakers in the circle, since, due to the overlapping, the length shortened to half.

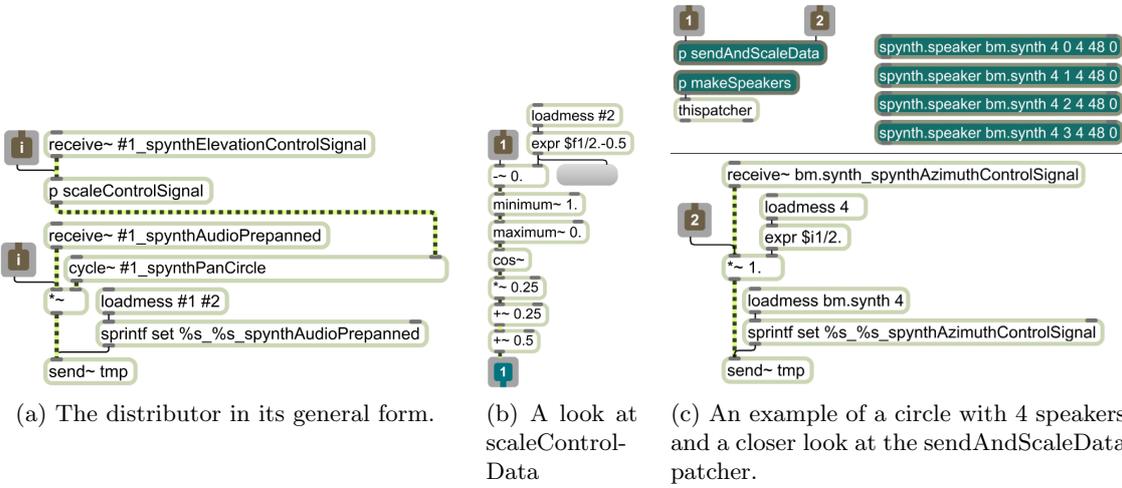


Figure 4.12: View of the abstract distributor and an example of a circle.

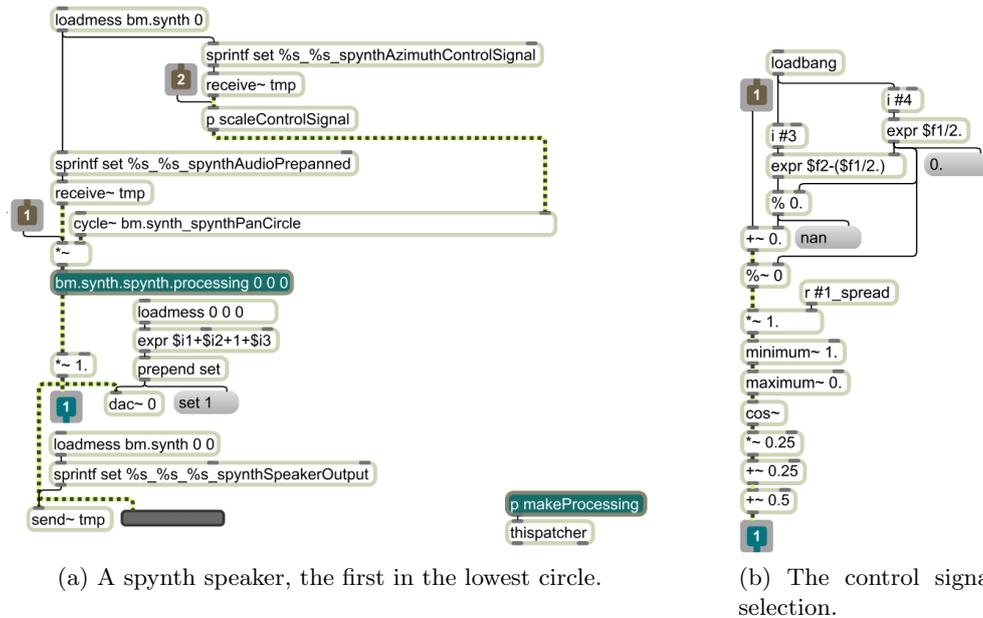


Figure 4.13: The speaker subpatch in detail.

The speaker exhibits some other functions. Such as the ability to tap into every speaker individually using the `%s-%s-%s_spynthSpeakerOutput`, where the first string is defined as the global name of the Spynth instance, the second the number of the circle and the third the number of the speaker relative to circle (first speaker in each circle starts with 0). This way, for example, the gain display is done in the main patch. One very interesting feature is the ability to include a processing patch into every speaker. The most convenient thing is, that the script `[p makeProcessing]` actually looks for a patch named `[%s.spynth.processing]`, where the first string needs to be replaced with the global identifier of the Spynth instance, in the Max search patch and places the abstraction into the signal flow automatically. If the object can not be found, it will automatically connect the multiplication objects with one another, skipping a possible processing patch. The arguments passed are the circle number, the speaker number within the circle, and the global offset, to relate the speaker to the whole system. Using a high resolution field of virtual speakers, one can create quite intricate effects by filtering and shaping the incoming audio. Finally, the diagram in figure 4.13b shows a value named `#1_spread` that can be multiplied to the control signal. This modifier enables the user to spread the control regions of each speaker within the azimuth control signal. When done with too much overlapping, one destroys the synthesis effect and the original audio is heard enveloped around the audience. If reduced too short, one eventually eliminates all audio all together. In-between is a region of exploration, in which some interesting effects can be achieved.

4.2.2 Using Spynth

Spynth enables its user to pan audio within the defined speaker array at blazing speeds in a complete sphere. The manner in which it does so can be visualized just as the globe is organized.

The elevated panning is attracted by the poles at each high and low extreme, while the azimuth panning evolves in a perfect circle. If using low speeds slow enough to be detected as a rotation, no real synthesis happens. Once the speed is increased somewhere beyond 5Hz, the sound becomes spatially blurred and enveloping around the body. Once you come close to 20Hz you start hearing a pitch arising. This can be done in two dimensions, further complicating the timbre as both panning directions also modulate each other. A notion of the original signal will always remain, especially in areas where the panning accumulates, such as the top speaker for example. This actually poses a problem, because the unmodulated signal coming directly from up above distracts and almost overpowers the rest of the soundscape created. For this, different solutions that do not use a speaker on top in the center have been tried. While this reconstruction on top can be somewhat eliminated, you sacrifice the complete height in the synthesis technique.

The initial synthesis technique is comparable to that of amplitude modulation.¹⁵ Of course, this method creates a completely enveloping sound around the audience, to start. The resulting sound is *highly* spatial, by which I mean, it can *not* live without its space. The most pleasing insight was the following, that, due to the equal amplitude panning, when panning all sources into the one speaker, the result is a perfect copy of the original source sound. That is, any synthesis effect magically disappears. The synthesis method, therefore, needs its space to uncover its potential. One further difference, though, is that with amplitude (and ring) modulation, the volume drops to zero many times and starts oscillating, while with Spynth, the modulation happens by panning the audio around in frequencies, but, moreover, in every speaker individually while the overall amplitude remains intact, due to the equal power amplitude panning. If looked at a speaker alone, the sounds being sent to it are short bursts separated only a few samples apart, depending on the speed panned. Figure 4.14 shows a comparison between an amplitude modulated signal and a stereo pair panning modulated signal. As expected, the short bursts instead of smooth, periodic modulation of the AM synthesis, causes many more high frequencies to appear. What is quite interesting though, is that the spectra for both channels in the stereo panning are not nearly the same. One channel seems to lose high frequency content quite fast. Exact causes for this effect have not been found, but it does explain the high enveloping factor of the synthesis. Unlike the spectral splitting technique done in section 4.1, this synthesis method does not fuse as a perceptual cohesive whole, but remains distinct and decorrelated in every point, creating a truly spatial sound.

The ability to include a processing patch into every speaker and control each processing patch individually adds another dimension of unimaginable possibilities. To start, due to the high energy high frequency content created, it allows many different filtering techniques to be applied. With normal amplitude modulation, for example, the filtering techniques are quite limited (to three frequencies, when using a sine tone as source, to be exact). Other ways of decorrelating the individual speaker from the whole sound mass could be to introduce feedback lines or pitch shifting effects. One interesting idea that could be looked into in the future

¹⁵More so to amplitude modulation than ring modulation, because the original signal remains somewhat intact.

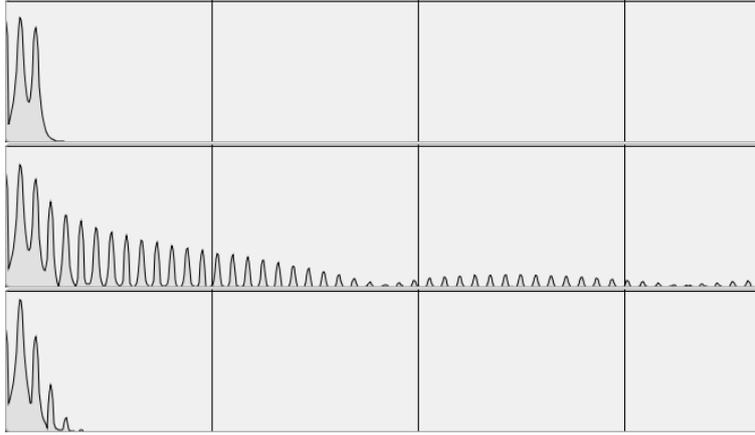


Figure 4.14: Comparison of an amplitude modulated signal (above) to a stereo panning modulated signal. Interestingly enough, the spectrum in the stereo pair (two lower spectra) are not the same.

is a way, that uses a second Spynth to control the processing patches of a first. Because each speaker sends its signal over a dedicated [senddim] connection, this can be programmed into the processing patch of another Spynth to be received by each corresponding speaker respectively filtering the sound in the same way that it is produced. Details on these ideas are still to be sought out.

Other ideas are to change the path that the panning goes along. Some basic, different methods have been tried with varying to minor success. For one, using noise as a control signal in any varying degree, will always result in audible noise with the corresponding varying amplitude. Having a signal pan itself does not necessarily result in a heightened listening experience, either. While it sounds somewhat interesting, it distorts the signal more than anything else. Another method that was also tried instead of circling around the audience, was changing the sound direction in front of the audience. The question was, if this would make the complete sound appear to be coming more from the front with a remaining enveloping sensation. But, this was not the case. It did, however, shape timbral differences in the higher frequency regions characteristically. Possibilities that have been tried so far are, for example, to rearrange the speakers in constellations that allow for different trajectories to appear. The interesting notion here is, that nothing fancy has to be done to the control signals, as the order of the virtual loudspeakers governs the trajectory as well. Some interesting trajectories that should be tried are, for example, the lissajous curves, described by [Chowning \(1971\)](#). According to his paper, these curves would sound exceptionally good, when used in audible trajectories, at least. Therefore, it is very tempting to implement these for the Spynth module.

Finally, I want to comment on why the distance parameter has been completely ignored in this approach so far. For one, distance perception is still too little understood for it to be simulated properly in an amplitude panned speaker system. But, if one were to implement the air absorption and the doppler effect correctly, what would modulate the sound in that case would be many more factors than just our own physiology. Moreover, to enable a rapid

oscillation technique in distance would exponentially increase the amount of channels needed to create a three dimensional grid. At least with the current workflow, this is not possible. Not that this approach is uninteresting, but I would view distance as a mediator between different textures for now. Considering the aesthetics of textural composition described in [Hagan \(2008\)](#), which would very well apply to the soundscapes created here, a way to facilitate the distinction of several textures at the same time, would be by blending them in and out of one another by sending them into the distance and having them "grow" in the process, as all virtual speaker would be panned away from the audience in a star like manner. This synthesis technique is just in its infant shoes, and still requires lots of research to grasp the full compositional potential.

4.2.3 Gestural Synthesis

One of the greatest insights this synthesis module enabled was a technique I labeled with *gestural synthesis*. Here the synthesis does not focus on the sound, but the spatial gesture it produces. This can be best demonstrated when using a periodic impulse, a click, as a sound source. There are basically three frequencies with which the composer can play around: the frequencies of the control signals and that of the repetitive click. Solely by achieving specific constellations between the three of them can achieve trajectories, held together like pots on a white page in a Gestalt psychology test image. The trajectories created are quite chaotic due to the high frequencies used and it requires a good amount of experimentation to build a good library of them. Listeners have reported that these synthesized trajectories, even though they are distinct clicks with sometimes large gaps in-between them, sound much smoother compared to the source panned effect. This might be a result of natural continuity and flow that the different frequencies together create, whereas much effort needs to be invested to have a source move along a complex, curved line.

This new method of synthesis is radical in the sense that it, if looked at in isolation, does not consider the actual sound, but rather suggests movement in space by causing certain events to happen in cohesive, but progressing spaces. Furthermore, it is a new form of synthesis exclusively applied to spatial music, or music that is partially concerned with its expression through spatiality. Instead of creating new timbres or spectromorphological constructs exclusively in the frequency domain, gestural synthesis' main aim is to create patterns and spatial constellations in three dimensional space around the audience. To a degree, this does inseparably impact the overall spectromorphology as well, but the tight entanglement between the spectro- and spatiomorphology has already been discussed in section [3.3.3](#). Hence, gestural synthesis touches upon the overall perceptual morphology, giving the possibility to change a sound in its entirety, including its spatial aspects, in a complete spatial gesture.

But, clicks, even though they might constitute compositional potential, are just ways of demonstrating this new effect the most clearly. Together with a fellow student, we started constructing a sketch for a possible complete composition. By combining his field of study, the sonification of MRI brain scans, we teamed up together with the Spynth to test its potential.¹⁶

¹⁶It should be clear, that these MRI scans are prerecorded data and no live playing was involved in this activity

Using a short, layering like sequence played by a guitar we controlled band- and low-pass filters on every virtual speaker in the array. The idea was to directly sonify the brain by using its data together with the right coordinates in space, and send them to the approximately closest virtual speaker in the same coordinates. The brain activity controlled both the gain, center frequency and quality value of the filters. The first result was already astonishing and meditative at the same time. The constant rotational feel together with the soothing guitar sound being turned into a wind like quality would lay a certain paradoxical calmness into the room. The filters, based on the strength of the brain activity, would then break off from the background in some places and create a distinct auditory stream. It was a constant play against the background, that seemed to soak everything back up again. Furthermore, the sequence of distinct and sometimes scarce break-offs from the overall enveloping texture would create figures and trajectories at, seemingly, its own will. The brain was truly playing the sound, which was the space. Furthermore, it seemed to create an imagery of its own through gestaltist figures and allusions, perhaps even speaking out what the owner of the brain was thinking at the time, speaking in a new language – the language of the space.

Chapter 5

Conclusions

”Die Kategorien diagonal durchkreuzen: das ist, glaube ich, das Allerwesentlichste. Deshalb haben wir auch manchmal Probleme mit Phänomenen, die das Bewußtsein des Menschen sehr schnell erweitern und erneuern, weil es einfach zu schnell geht. Es gibt ein Trägheitsgesetz des Psychischen genau wie der physikalischen Materie. Und dagegen ist nicht immer anzukommen, oder nur sehr schwer. Es liegt ganz an der Psyche der einzelnen Menschen, wie weit sie fähig sind, solche Bewußtseinsenerweiterungen, die durch die neuen Erlebnisse kommen, möglichst schnell mitzuvollziehen.”¹ – (Stockhausen 1964)

This thesis approached the discussion on the use of space in a musical context from three different viewpoints: For one, it traced the use of space throughout the musical history from the earliest assumed examples to contemporary trends and sophisticated existing theories on a possible framework on how space can be used truly musically. It was found that the physical space, although used more frequently, even in acoustic music, than one might expect at first, has mainly played a submissive role as an effect or, as it is a common view today, as a way to enhance the listening experience. Furthermore, the few composers that concern themselves with the use of space as a form of musical expression seem to have approached the problem in very individual ways. Although, especially in contemporary trends, the notion of space has reached sophisticated and emancipated levels, a complete framework, a school of spatial music does not exist so far. It is assumed that with the increased interest in spatial audio today, an established language and notion of space in a musical context will not be a long time in coming.

The importance here is the distinction of using space as a means to merely articulate the musical content and enhance the listening experience, using it as an effect, as opposed to making use of spatial developments in the music as a language of its own, without which the musical piece might not be able to carry itself and its meaning anymore. The best analogy to reduce a three dimensional spatial piece into a stereo format that could be given in this context

¹”Crossing through the categories diagonally: that is, I think, the most important thing. That’s why we also sometimes have problems with the phenomenon which quickly changes and extends the consciousness of people, because it happens too quickly. There is a law of inertia of the psychic similar to that of the physical material. And you can’t resist it, or only with difficulty. It is completely up to the psyche of the individual person how capable they are to quickly realize such consciousness expansions which come from new experiences.”

is the example of a melodic piece, which has to be reduced to only a subset of pitches. Three dimensional music, which uses spaces as a central aspect of its concept, should not be able to be reproduced in any format that does not exhibit all three dimensions to some degree of accuracy. Any allowance of a spatial piece to be reduced into a format of lower dimensional resolution than the one composed should indicate the carelessness with which the space is considered, indicating a failure, if intended, of its use as a true musical language.

Next, the thesis investigates the perception of space, from the lower level auditory spatial cues, all through the the perception of auditory scenes in ecological psychoacoustics, to a more metaphysical discussion on the perception of space itself. While the previous auditory cues and discussion of ecological psychoacoustics represents the literature synthesis undergone and acts as a pointer to limitations and consideration when composing spatial music, the discussion on the metaphysical notion of space points to a personal composition methodology which was taken into account in the later sections. In particular, the experience of space, the learning of space and the cross-modal perception of space lead to a notion of space in sounds that goes beyond the mere spatial position or the room qualities encoded into it. The idea that sounds themselves hold and allude to physical spaces merely through their spectromorphology is a conclusion that arose out of the experimentation with spatial synthesis and composition and was interrelated with similar conclusions drawn by recent publications on the matter.

The thesis goes on to discuss and define certain compositional trends in the isolated use of physical space, classifying three general approaches. The first is the popular use of sound moving around the space as sonic objects, termed trajectorial approaches. Sound trajectories are considered somewhat cliché, but can also not be neglected due to them being a fundamental building block spatial music generation. Within trajectorial composition methodologies, the effort is made to move away from the cliché by coming up with ways to avoid being termed a theatrical approach to the matter. Here, the previous discussion on the notion of space in the spectromorphology plays an important role. Furthermore, the notion of rhythm implied by spatial motion and displacement can not be neglected, but the perception of the such also requires a cultural integration, due to studies that show rather poor and discouraging results. Flaws and possible misinterpretations were also identified in these studies, which also indicate that the spectromorphology and the spatial movement can act counterproductively to one another. One suggestion, therefore, is for composers to listen closely to the sounds they spatialize and carefully decide on their movement based on their spectral development. Nevertheless, the insights made in the course of this section [3.3.1](#) lead to an analogy to the language of contemporary dance. While also still lacking a unified language, it is suggested that the developments in this field should be closely observed to possibly draw conclusions in the case of moving sounds. If the analogy succeeds in the respective composition, the theatrical element of moving sounds should be diminished just as contemporary dance is not necessarily associated with theater as well.

The second trend and strong schools of spatial considerations are those occupied with the use of real spaces themselves. These movements are both the soundscape artists, recording and experimenting with spaces and ambiances on tape, and the site-specific installations, that take

the space in which the installation is placed into account. While both these trends were not investigated practically during the course of this thesis, their theoretical discussion leads to the following third approach in spatial composition. Textural composition, virtual spaces and spatial sounds forms a relatively new way of composing spatially, especially due to increased computational power and the ability to handle large amounts of musical information in real time. Coming from soundscape composition, these approaches create imaginary, unheard and enveloping sound worlds which pose a completely new way of spatial compositional thinking. Especially the inclusion of Gestalt theory in the identification of musical meaning of these spatial sounds and their fluctuations in different areas around the listening space has been found to be a good discussion ground of these musical approaches. As the texture builds a canvas, a background, emerging changes in different spots can lead to many illusions and even the impression of motion, leading to the idea of gestural synthesis in the last chapter.

The final, practical chapter, discusses the construction of mainly two spatial synthesis tools. The first tool represents a new approach to timbral spatialization, giving the composer the possibility to decorrelate an arbitrary grouping of FFT bins via self-defined methods. The example given is a spectral delay, but the architecture allows a relatively fast change to other methods, such as possibly temporal or other coloring decorrelation methods to be employed, or even to be used in succession with one another. The second synthesis method is a relatively new way of creating enveloping sounds by panning the audio at high frequencies in all angles around the listener. Although I do not believe that this method has never been tried before, any mention of this technique in literature could not be found, other than the examples given in the Max5 distribution, upon which the fundamental functionality was based on. The module built for this spatial synthesis technique was also created with a high flexibility in mind, allowing for a greatest possible degree of experimentation. Compositional sketches were done and the method seems fruitful based on the experiences made and described. It especially led to a new notion of synthesis technique, which is not so much concerned with the sound, primarily, but rather with the space itself: using different frequencies in both the horizontal and vertical panning, as well as controlling the possible frequencies of the source sound, one can reach specific interlockings that result in coherent events along a trajectory, formed in the mind by several Gestalt principles of grouping. The implications could be far-reaching: having the new compositional methodology that is mainly concerned with space as a musical language, a synthesis technique that creates spatial gestures and spatial sounds simultaneously could lead to a supporting sub-framework for the greater good of establishing space as a musical parameter.

Future work will definitely concentrate on the investigation and formalization of the possibilities for this new technique. Furthermore, more developmental work on these tools will have to be done. The most promising and most widely used spatialized granular synthesis technique had to be left out of the scope of this thesis. Its impacts on the spatial compositional methodology worked out in this thesis and still needs to be investigated. But, finally, musical works have to be created using all this theoretical knowledge and the tools built, as they form the final product best capable of demonstrating these concepts in a finished form.

5.1 Redefining the Musical Conception of a Whole Culture

Computational power has opened completely new doors for the imaginative composer today. Still, composers, in their drive to seek new approaches to electronic music, hit new walls as they combine processes stacking up on top of each other in complexity. But most significantly, it is not the software that limits composers in spatial composition and performance, but the expensive hardware. This problem also cripples spatial music, as composers are forced to squash their ideas into a stereo panorama for radio, cd or even applications to contests or convocations, for example. Stockhausen, for example, already argued the necessity of standardized spaces:

”Bafänden sich in jeder größeren Stadt solche Räume, so bekäme auch das gemeinschaftliche Hören in Musikhallen, im Gegensatz zum Radiohören, wieder einen neuen Sinn. Die bisher übliche Konzertpraxis würde – was das Hören elektronischer Raum-Musik betrifft – von einer Form abgelöst, die dem Besuch von Bildergalerien entspräche. Es gäbe permanente Programme, die periodisch wechselten, und man könnte zu jeder Tageszeit das elektronische Programm hören.”²
(Stockhausen 1959)

For spatial music to become an accepted, and, foremost, respected musical parameter, it needs to be exposed to the public and the audience on a regular level, meaning that electroacoustic concerts exhibiting space as a musical parameter need to become a standard point in the program of the major concert halls. Bringing up such new and special notion as a musical language, it can run the risk of being a sole academic stream of thought, an elitist movement which passes by the general public. This, though, should not be in the best interest of the spatial music community. For space to be absorbed as a musical language, it needs to be embedded in the cultural thinking, as an average concept of music. As Schenker writes:

”[...] the secret of balance in music lies in the permanent awareness of levels of transformation and of the movement of the surface structure towards the initial generative structure, or of the reverse movement. This awareness is always on the composer’s mind, without it, every surface structure would degenerate into chaos.”³

Another problem three dimensional spatial music faces is the preservation and its delivery to the home. Solutions to the first point can lie in adaptable technologies such as Ambisonics, giving the possibility to both record the three dimensional scene in one standardized encoding and playback the piece over an arbitrary set of speakers. The short discussion on Ambisonics in section 2.2 shows though, that first order ambisonics do display issues in resolution and simply increasing the order also comes with its own deficits. The psychoacoustic discussion in section 3.1 also shows, that an HRTF encoded binaural recording would not work for every individual

²”If there were such rooms in each larger city then communal listening in music halls, as opposed to listening to the radio, would take on a new significance. The heretofore usual concert experience - in terms of listening to electronic space-music - would be separated in a way that it would become similar to visiting a picture gallery. There would be permanent programs which would change periodically and one could listen to the electronic program at any time of the day or night.”

³As quoted in Bigand & Poulin-Charronnat (2006)

and, although viable, does not display an optimal solution for the archiving and distribution of spatial pieces. Compared to three dimensional speaker systems, binaurally encoded recordings of spatial works are still the most plausible solution when it comes to bringing the piece to a home stereo. But binaural recordings as well as multichannel speaker setup in a home setting have the problem of being fragile and risky if the spatial information is essential for the piece's success.

Three dimensional speaker setups are, furthermore, unthinkable at home, as they become very complex, expensive and obtrusive. Thinking about the fact that most home stereo owners do not even take care to place their stereo speakers in a correct position, trusting the average private person to sit in the correct position of an engulfing twenty-plus speaker system can be laughed off with an ironic smile. As a composer of spatial music, I need to be sure that the listener can perceive the spatial developments of the piece as correctly as a post-digital artist can be sure that most decent stereo systems can reproduce all possible frequencies in a more or less correct way.⁴ While the hi-fi quality of a stereo system is a given, though, the correct spatial placement of the speakers and the listener himself is not. Therefore, if the home listener will perceive the spatiality correctly is outside of the composers reach and could possibly lead to misinterpretations. Smalley (2007) also addresses this issue, calling the malformed relation of the personal space to the speaker system a *vantage point shift*. Depending on the impact the absence or distortion of spatial cues in the composition might have – especially with spatiality as means of musical expression being such an unknown and unusual factor in music listening – this should be a factor to be considered when thinking about distributing the music through any sort of media.

Instead, the composer needs to motivate and encourage the audience to come to the concert hall. Here, the spatial image can be tested and controlled, and the audience can be placed within the speaker array where they can be sure to perceive the music correctly as to the composers intent. Spatial music, as with academic and concert hall music in general, is not necessarily thought of for mass consumption. As long as there exists enough possibilities to perform the music live, there should be no real need to provide the music for home use. Instead, what spatial music probably needs even more than any other contemporary musical movement, is a newfound willingness, motivation and desire to visit a performance outside the house. There needs to be a general shift the acceptance of music, that leads away from *iTunes* and instant downloads; mindless mass consumption; running in the background during dinner and back to the social aspect of consciously experiencing a concert of music outside the house. It seems that the idea of each piece of music having its time, space and purpose needs to be awoken in our general society. People need to become aware, therefore, that you can *not* obtain all music for your home stereo and that there are certain occasions that require the listener to come into the concert hall. Concerts like those with the BEAST or the Game of Life foundation simply can not be experienced unless we revive the social event of going to a concert to new heights of popularity.

⁴Although not the target group for glitch artists, with the emerging popularity of laptop and mobile phone speakers, we currently see a regressing trend in the hi-fi quality of audio reproduction systems

That said, artists, for the most part, have not gone out to drive the existing technologies to their limits. A true methodology in spatial composition perhaps also does not exist, because composers seem to be perfectly happy sitting in the comfort zone of using the possibilities that the engineers provide (Otondo 2008). New ground in spatial composition has hardly been touched, because what seems to be popular is the mere panning around of sounds, cross fading and reconstruction of enveloping soundscapes. Based on the survey done in Otondo (2008), composers hardly seek to tease the technology in terms of spatial reproduction, by bringing it and human psychoacoustics to its limits, instead preferring to remaining in what is known in the field. Xenakis, for example, went out to shape the musical thinking and improve its perception by exactly providing challenging spatial compositions:

”Xenakis’s aim, for example, was precisely to play with the diversity of the senses, and not to create correspondences in their expression. When Xenakis remarks: ”Man is intelligent enough to follow two discourses at once,” his aim is to integrate in his installations as much intelligence, difference and variation as possible. The audience has to contribute actively to the construction of the sense of these art works; the spectator himself has to effect the operation of synthesizing the poly-temporality of the proposed spectacle. Therefore, instead of focusing the spectator’s attention by simply playing with his reflexes or his corporality, or hypnotizing him with sequences of familiar images, Xenakis’s abstract and multi-layered Polytopes try to open the audience’s mind to diversity and simultaneity. This way, these electronic poems express the idea of an intelligent space, long before it became a fashionable concept in contemporary architectural theory.” (Sterken 2001)

Many of the perceptual studies cited in this thesis have shown a short term improvement in sound localization over the course of their respective experiments in both peripheral localization and distance. This shows that the senses can be focussed to outperform themselves when concentrated. That the higher level neurological mechanisms to decode the spatial auditory scene are shapeable and trainable has been discussed in detail as well. The more important, unanswered question in this thesis that remains, though, is the question of long term improvement in the recognition of new musical concepts. Or, more colloquially, can composers of spatial music shape society and can we train our culture to be more aware of their spatial hearing abilities? The results found with the early-blind’s spatial hearing abilities and the cross-modal hypothesis both suggest a strong influence on neurological development of these abilities. Considering musical developments from the Middle Ages to Modern times shows, that society does absorb the avant-garde composer’s constant search for new languages and decodes the ever more complex ways of expression over time. One just needs to look at the intricate language of Igor Stravinsky’s *Le sacre du printemps* (1913), and its extreme history: the riots it first caused during its premiere due to its intense rhythmic orientation and use of dissonance and recognition it now enjoys, is known as one of most celebrates works in history.⁵ Although these

⁵Although this is somewhat of an urban fact, I may refer to Classic Cat (2011), with a reference given therein, stating that Leonard Bernstein once said one particular section in *Le sacre du printemps* had “[...] the

developments mostly appeared in melodic and harmonic domains, other musical progressions as radical as *musique concrète* was able to embed itself in our musical culture as well.⁶ The spatial parameter is another radical break from traditional musical conceptions, and goes hand in hand with *musique concrète*, complicating its acceptance in the general public even more.

”When a culture provides consistent exposure to a class of sounds, perception is reasonably consistent among listeners within that culture. Perception does not require the sound to have any relevance to life; a spoken sequence of random numbers can be perceived as linguistic objects, a sequence of musical notes can be perceived as a melody, and a sound source can be localized. Perception is predominantly a property of cultural exposure.” (Blessner & Salter 2007)

Composers of spatial music therefore need to find a way to communicate their intentions through constant exposure. As already argued in the main text of this thesis, the awareness of space as a musical parameter gain for both the audience, being shown new modes of listening and spatial awareness in their every day life, and the composers, given the possibility to move into more complex stratum with a general improvement in the perception of music:

”The deeply meaningful sense of space that is aroused when listening to electroacoustic music has its roots in a lifetime of embodied spatial experience.” (Kendall 2010)

”Thanks to both factors (predisposition and intensive exposure), non-musicians become *experienced listeners* that do not strongly differ from musically trained listeners in numerous experimental tasks.” (Bigand & Poulin-Charronnat 2006)

If we, for example, arrive at a natural, almost unconscious notion of musical, physical space in acoustic music, we have reached the far goal of total cultural integration. Electronic music would then have all the freedom to express itself in this domain without the fear of being misunderstood.

Finally, the discussion on the spatial auditory perception of blind people in section 3.1 has led to one idea for future work on a greater scale. One question that arose was, what would spatial acousmatic music sound like composed by a blind person? Currently to the knowledge acquired in the course of this thesis, no blind composer that has had the chance to compose for 3D Audio is known. Although some research presented in this thesis might suggest – contrary to popular belief perhaps – that blind people actually do not have a significantly improved spatial hearing (comparable to that than seeing people, as it was the case in the frontal range with the early blind in Collignon et al. (2007)), other research found does support this idea, especially in the rear areas of our hearing or when the subject lost his sight at a later stage in his life, having had the chance to refine his spatial hearing through an early stage of cross-modal comparison.

best dissonances anyone ever thought up, and the best asymmetries and polytonalities and polyrhythms and whatever else you care to name.”

⁶It has, at least, established itself as a widely used stylistic device in the music industry. One just needs to look at early bands like The Beatles or Pink Floyd.

Future work in this direction would, then, be to build a software tool that would specifically allow for blind people to compose acousmatic music and, more specifically in the interest of the thesis, would allow them to make use of the spatialization possibilities available today with similar ease as seeing composers do. This is not as much an experiment, as it simply is the personal interest in the final compositions created. If this tool allows for the electroacoustic musically interested blind person to engage in this musical activity, for a blind composer of electronic music to emerge, then it could be interesting to hear and learn from how he perceives the spatiality of the music evolving. Despite this, if these individuals' spatial hearing would be measurably worse or better than that of a hearing person, it would be clear that he, being more reliant on his ears (especially, being interested in electroacoustic music, if he is a musician himself), would perceive the space in the completely different way. It could, therefore, not just be a tool for seeing spatial composers to learn from the spatial notions of blind composers, it could also form a method for the blind composer to communicate his world, as he perceives it, through new means by creating his sound world directly for the audience.

”Wenn ich zum Beispiel jetzt das Licht ausmache und eine Vierspurvorführung eines bestimmten Abschnittes von *Kontakte* mache und dann sage: ’Merken Sie sich diesen Ton, wie der sich räumlich verhält; er geht ganz weit weg und kommt dann wieder zurück nach vorn’, dann sagen Sie: ’Ja, das ist doch eine Illusion, der Ton ist ja nicht weit weggegangen.’ Daran merken Sie etwas sehr Interessantes. Die Tatsache, daß Sie sich nicht vorstellen können, daß diese Wand sich bewegt hat - mit anderen Worten: daß die Wände des Raumes, in dem Sie sind, sich offenbar für ihre Augen nicht bewegen, verleitet Sie zu der Reaktion, zu sagen: ’Ja, was ich akustisch erlebe ist eine Illusion, und nur das, was ich optisch prüfen kann, ist die Wahrheit.’ Und das hat unerhörte Konsequenzen; denn Sie wissen, daß heutzutage nur das, was geschrieben ist, als glaubwürdige Wahrheit gilt, das, was *unterschrieben* ist. Das alles hat mit unserer visuellen Erziehung, mit der Schreiberziehung zu tun. Wir sind keine akustischen Menschen mehr. Wir sind im Grunde taub. Unser Wahrheitsbegriff basiert nur auf der Wahrnehmung der Augen.

Ich kann zum Beispiel als Musiker sagen: Für mich ist das die Realität des Klanges, was ich erlebe, wenn ich die Augen zumache; die Feststellung, daß die Mauern sich nicht bewegen können, ist eine Illusion; der Klang hat sich 150 Meter weit entfernt und ist dann zurückgekommen. Da sehen Sie, Sie könnten das Ganze auf den Kopf stellen: Es sei eine Illusion, zu sagen, die Mauer habe sich nicht bewegt. Warum? Weil wir das Visuelle als das Absolute für unsere Wahrnehmungskriterien bezeichnen und nicht das Akustische. Wir leben in einer rein visuellen Gesellschaft und Tradition.”⁷ (Stockhausen 1964)

⁷”If I were to turn the light off, for example, and make a 4-track performance of a certain section of *Kontakte* and then say: ’Pay attention to how this sound behaves in space; it goes very far away and then comes back to the front’, then you say: ’Yes, that is an illusion, the sound did not really move far away.’ You notice something interesting in that. The fact that you cannot imagine that this wall did not move - in other words: the walls of the room that you are in, which, according to your eyes, obviously did not move, induce you to the reaction to say: ’Yes, what I experience acoustically is an illusion and only that which I can verify optically is reality.’ And that has outrageous consequences; because you know that nowadays only that which is written is considered a credible if it has been signed. That all has to do with our visual upbringing, with writing education. We are no longer acoustical people. Basically, we are deaf. Our concept of reality is based solely upon the perception of our eyes.

As a musician, I can say, for example: ’For me, the reality of a sound that I experience is when I close my eyes; the ascertainment that the wall cannot move is an illusion; the sound moved away 150 meters and then came back. You see, you can turn everything upside-down: It is an illusion to say that the wall did not move. Why? Because we identify the visual, not the acoustical, as the absolute for our perception criteria. We are living in a purely visual society and tradition. ”

Bibliography

- Adair, S., Alcorn, M. & Corrigan, C. (2008), A study into the perception of envelopment in electroacoustic music, *in* 'Proceedings of the 2008 International Computer Music Conference', Belfast, UK.
- Andreeva, I. G. (2004), The influence of spectral range signal on threshold duration of perception of sound source radial motion, *in* 'XV Session of the Russian Acoustical Society I.G.Andreeva', Nizhny Novgorod, Russia, pp. 500–502.
- Arons, B. (2001), 'A Review of The Cocktail Party Effect', *Journal of the American Voice I/O Society* **35–50**(July), 701–705.
- Ashmead, D. H., Wall, R. S., Ebinger, K. A., Eaton, S. B., Snook-Hill, M.-M. & Yang, X. (1998), 'Spatial hearing in children with visual disabilities', *Perception* **27**(1), 105–122.
- Austin, L. & Smalley, D. (2000), 'Sound Diffusion in Composition and Performance: An Interview with Denis Smalley', *Computer Music Journal* **24**(2), 10–21.
- Baalman, M. A. J. (2010), 'Spatial Composition Techniques and Sound Spatialisation Technologies', *Organised Sound* **15**(03), 209–218.
- Barreiro, D. L. (2010), 'Considerations on the Handling of Space in Multichannel Electroacoustic Works', *Organised Sound* **15**(03), 290–296.
- Barrett, N. (2004), Ambisonics spatialisation and spatial ontology in an acousmatic context, *in* 'Proc. of the 8th Electroacoustic Music Studios conference', Shanghai, China, available online: http://www.natashabarrett.org/EMS_Barrett2010.pdf (Accessed in August 2011).
- Barrett, N. (2010), Kernel expansion: A three-dimensional ambisonics composition addressing connected technical, practical and aesthetical issues, *in* 'Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics', Paris, France, available online: http://ambisonics10.ircam.fr/drupal/files/proceedings/presentations/017_10.pdf (Accessed May 2011), pp. 6–9.
- Bates, E. (2004), Composing, Perceiving and Developing Real-time Granulation in Surround Sound., Master's thesis, Trinity College of Dublin.

- Bates, E. (2009), *The Composition and Performance of Spatial Music*, PhD thesis, Trinity College Dublin.
- Belendiuk, K. & Butler, R. A. (1975), ‘Monaural localization of low-pass noise bands in the horizontal plane’, *Journal of the Acoustical Society of America* **58**(3), 701–705.
- Beranek, L. L. (1992), ‘Music, Acoustics, and Architecture’, *Bulletin of the American Academy of Arts and Sciences* **45**(8), 25–46.
- Beranek, L. L. (2004), *Concert halls and opera houses: music, acoustics, and architecture*, Springer Verlag, New York, NY, USA.
- Beyst, S. (2003a), ‘Luigi Nono: Il Prometeo: A Revolutionary’s Swansong’.
URL: <http://d-sites.net/english/nono.htm> (Accessed in August 2011)
- Beyst, S. (2003b), ‘Musical space and its inhabitants: an inquiry into three kinds of audible space’.
URL: <http://d-sites.net/english/musicalspace.htm> (Accessed in August 2011)
- Bigand, E. & Poulin-Charronnat, B. (2006), ‘Are we ”experienced listeners”? A review of the musical capacities that do not depend on formal musical training.’, *Cognition* **100**(1), 100–130.
- Blauert, J. (1997), *Spatial hearing: the psychophysics of human sound localization*, 4th edn, The MIT Press, Cambridge, MA, USA.
- Blessner, B. & Salter, L.-R. (2007), *Spaces Speak, are you listening?*, MIT Press, Cambridge, MA, USA.
- Brant, H. (1955), ‘The uses of antiphonal distribution and polyphony of tempi in composing’, *American Composer’s Alliance Bulletin* **4**(3), 13–15.
- Bregman, A. S. (1990), *Auditory Scene Analysis: The Perceptual Organization of Sound*, The MIT Press, Cambridge, MA, USA.
- Bregman, A. S. (2004), *International Encyclopedia of the Social and Behavioral Sciences*, Pergamon (Elsevier), Amsterdam, Netherlands, chapter Auditory scene analysis, pp. 940–948.
- Brungart, D. (1993), Distance simulation in virtual audio displays, in ‘Proceedings of the IEEE National Aerospace and Electronics Conference’, Dayton, OH , USA, pp. 612–617.
- Bryant, D. (1981), ‘The *corti spezzati* of St. Mark’s: myth and reality’, *Early Music History* **1**, 165–186.
- Byrne, D. (2010), ‘How architecture helped music evolve’.
URL: (Video file) http://www.ted.com/talks/david_byrne_how_architecture_helped_music_evolve.html (Accessed in August 2011)

- Cabrera, D., Nguyen, A. & Choi, Y. (2004), Auditory versus visual spatial impression: a study of two auditoria, *in* 'Int. Conf. on Auditory Display', Sydney, Australia.
- Callanan, M. J. (2010), 'Sonification of You: Arseny (Arsenij) Mikhailovich Avraamov'.
URL: <http://www.music.columbia.edu/masterpieces/notes/stockhausen/GesangHistoryandAnalysis.pdf> (Accessed in June 2011)
- Cardew, C. (1961), 'Report on 'Carré': Part 2', *The Musical Times* **102**(1425), 698–700.
- Cascone, K. (2000), 'The Aesthetics of Failure: "Post-Digital" Tendencies in Contemporary Computer Music', *Computer Music Journal* **24**(4), 12–18.
- Cheveigné, A. D. (2003), Espace et Son, *in* 'Colloque espaces de l'Homme', Paris, France.
- Chion, M. (1994), The three listening modes, *in* 'Audio/Vision: Sound on Screen', Columbia University Press, New York, NY, USA, available online: <http://helios.hampshire.edu/~hacu123/papers/chion.html> (Accessed in August 2011).
- Chowning, J. M. (1971), 'The simulation of moving sound sources', *J. Audio Eng. Soc* **19**, 2–6.
- Classic Cat (2011), 'Le sacre du printemps'.
URL: http://www.classiccat.net/stravinsky_i/_tros.info.php (Accessed in August 2011)
- Collignon, O., Lassonde, M., Lepore, F., Bastien, D. & Veraart, C. (2007), 'Functional Cerebral Reorganization for Auditory Spatial Processing and Auditory Substitution of Vision in Early Blind Subjects', *Cerebral Cortex* **17**(February), 457–465.
- Darwin, C. J. (2008), *Auditory perception of sound sources*, Springer Verlag, Heidelberg, Germany, chapter Spatial Hearing and Perceiving Sources.
- de la Motte-Haber, H. (1986), 'Zum Raum wird hier die Zeit', *Österreichische Musikzeitschrift* (June).
- de Nooijer, A. (2009), Investigations on the intersection of music and architecture: the sources and works of Iannis Xenakis, Master's thesis, University of Technology in Delft.
- Demers, J. (2010), *Listening Through the Noise*, Oxford University Press, New York, NY, USA.
- Einbond, A. & Schwarz, D. (2010), Spatializing timbre with corpus-based concatenative synthesis, *in* 'Proceedings of the International Computer Music Conference', New York, NY USA.
- Féron, F.-X., Frissen, I. & Guastavino, C. (2009), Upper limits of auditory motion perception: The case of rotating sounds, *in* 'Proceedings of the 15th International Conference on Auditory Display', Copenhagen, Denmark.

- Fowler, M. (2010), ‘The Ephemeral Architecture of Stockhausen’s Pole für 2’, *Organised Sound* **15**(03), 185–197.
- Garity, W. E. & Hawkins, J. N. A. (1998), ‘Fantasound’.
URL: <http://www.widescreenmuseum.com/sound/fantasound1.htm> (Accessed in August 2011)
- German Federal Cultural Foundation (n.d.), ‘Motion bank by william forsythe’.
URL: http://www.kulturstiftung-des-bundes.de/cms/en/programme/tanzerbe/motion_bank.html (Accessed in June 2011)
- Gibson, J. (2009), ‘Spectral Delay as a Compositional Resource’, *eContact!* **4**(11), 271–289.
URL: http://cec.sonus.ca/econtact/11_4/gibson_spectraldelay.html (Accessed April 2011)
- Goßmann, J. & Dombois, F. (2003), The Spatial Sound Lab at Fraunhofer IMK.VE, in ‘Proceedings of the 2003 International Conference on Auditory Display’, Boston, MA, USA.
- Grisey, G. (1987), ‘Tempus ex Machina: A composer’s reflections on musical time’, *Contemporary Music Review* **2**, 239–275.
- Guerrasio, F. & Palme, C. D. (2008), The impossibility of absolute spatial determinations: from Luigi Nono and Salvatore Sciarrino’s musical aesthetics to Giordano Bruno’s innumerable worlds., in ‘Proceedings of the fourth Conference on Interdisciplinary Musicology’, Thessaloniki, Greece.
- Guski, R. (1992), ‘Acoustic Tau: An Easy Analogue to Visual Tau?’, *Ecological Psychology* **4**(3), 189–197.
- Hagan, K. (2008), Textural Composition and its Space, in ‘Sound and Music Computing Conference: sound in space-space in sound’, Berlin, Germany.
- Handel, S. (2006), *Perceptual Coherence: Hearing and Seeing*, Oxford University Press, Inc., Oxford, NY, USA.
- Harley, J. (2004), *Xenakis: his life in music*, Routledge, New York, NY USA.
- Harley, M. A. (1994), Space and Spatialization in Contemporary Music: History and Analysis, Ideas and Implementations, PhD thesis, McGill University, PDF Reprint under Maja Trochimzyck, Moonrise Press, Los Angeles, California, 2011.
- Harley, M. A. (1997), ‘An American in Space: Henry Brant’s ”Spatial Music”’, *American Music* **15**(1), 70–92.
- Hofmann, B. (2006), Spatial aspects in xenakis’ instrumental works, in ‘Definitive Proceedings of the International Symposium Iannis Xenakis’, Athens, Grece.

- Hofmann, B. (2011), Exploring the Sound Tapestry: The role of space in Nomos Gamma, in 'Xenakis International Symposium', London, England, available online: <http://www.gold.ac.uk/media/06-2-boris-hofmann.pdf> (Accessed in August 2011).
- Hollerweger, F. (2006), Periphonic Sound Spatialization in Multi-User Virtual Environments, Master's thesis, Graz University of Music and Dramatic Arts.
- Howard, D. M. & Angus, J. A. (2009), *Acoustics and Psychoacoustics*, 4th edn, Focal Press, Burlington, MA, USA.
- Ives, C. (1933), *American Composers on American Music*, Stanford University Press, Stanford, CA, USA, chapter Music and its Future.
- Kane, B. (2007), 'L'Objet Sonore Maintenant: Pierre Schaeffer, sound objects and the phenomenological reduction', *Organised Sound* **12**(01), 15–24.
- Kearney, G., Gorzel, M., Boland, F. & Rice, H. (2010), Depth perception in interactive virtual acoustic environments using higher order ambisonic soundfields, in 'Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics', Dublin, Ireland.
- Kendall, G. S. (1995), 'The Decorrelation of Audio Signals and Its Impact on Spatial Imagery', *Computer Music Journal* **19**(4), 71.
- Kendall, G. S. (2010), 'Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music', *Organised Sound* **15**(03), 228–238.
- Kendall, G. S. & Cabrera, A. (2011), Why things don't work: What you need to know about spatial audio, in 'International Computer Music Conference', Huddersfield, England.
- Kerber, S., Wittek, H., Fastl, H. & Theile, G. (2004), Experimental investigations into the distance perception of nearby sound sources : Real vs WFS virtual nearby sources, in 'Fortschritte der Akustik, DAGA', Strassburg, France, pp. 1041–1042.
- Kim-boyle, D. (2005), Sound spatialization with particle systems, in 'Proceedings of the 8th International Conference on Digital Audio Effects', Madrid, Spain, pp. 65–68.
- Kim-boyle, D. (2006), Spectral and Granular Spatialization with Boids, in 'International Computer Music Conference Proceedings', New Orleans, USA.
- Kim-Boyle, D. (2008), Spectral Spatialization-An overview, in 'Proceedings of the International Computer Music Conference', Belfast, Northern Ireland.
- Koffka, K. (1955), *Principles of Gestalt Psychology*, Routledge, London, England.
- Kopčo, N., Čeljuska, D., Puszta, M., Raček, M. & Sarnovský, M. (2004), Effect of spectral content and learning on auditory distance perception, in 'Proc. 2nd Slovak-Hungarian Joint Symposium on Applied Machine Intelligence', Herlany, Slovakia.

- Lakatos, S. & Shepard, R. N. (1997), ‘Constraints common to apparent motion in visual, tactile, and auditory space’, *Journal of Experimental Psychology: Human Perception and Performance* **23**(4), 1050–1060.
- Lauke, K. (2010), ‘The placement of sound in artistic contexts’, *Body, Space & Technology Journal* **09**(1).
- Lessard, N., Paré, M., Lepore, F. & Lassonde, M. (1998), ‘Early-blind human subjects localize sound sources better than sighted subjects’, *Nature* **395**, 278–280.
- Lombardo, V., Arghinenti, A., Nunnari, F., Valle, A., Vogel, H., Fitch, J., Dobson, R., Padget, J., Tazelaar, K., Weinzierl, S. & Others (2005), The virtual electronic poem (vep) project, in ‘Proceedings of the 2005 International Computer Music Conference’, Citeseer, pp. 451–454.
- Lombardo, V., Valle, A., Fitch, J., Tazelaar, K. & Weinzierl, S. (2009), ‘A Virtual-Reality Reconstruction of Poème Électronique Based on Philological Research’, *Computer Music Journal* **33**(2), 24–47.
- Loomis, J., Klatzky, R. & Golledge, R. (1999), Auditory distance perception in real, virtual, and mixed environments, in ‘Mixed Reality: Merging Real and Virtual Worlds’, Tokyo, Japan, pp. 201–214.
- Lossius, T. & Baltazar, P. (2009), Dbap - distance-based amplitude panning, in ‘Proceeding of the International Computer Music Conference’, number 1, Montreal, Canada.
- Lynch, H. & Sazdov, R. (2011a), An ecologically valid experiment for the comparison of established spatial techniques, in ‘Proceedings of the ICMC 2011’, Huddersfield, UK.
- Lynch, H. & Sazdov, R. (2011b), An investigation into compositional techniques utilized for the three-dimensional spatialization of electroacoustic music, in ‘Sound and Music Computing Conference: sound in space-space in sound’, New York, NY, USA.
- Maconie, R. (2005), *Other Planets: The Music of Karlheinz Stockhausen*, Scarecrow Press, Lanham, MD, USA.
- Macpherson, E. A. (1995), A Review of Auditory Perceptual Theories and the Prospects for an Ecological Account, Master’s thesis, University of Wisconsin-Madison.
- Malinina, E. S. & Andreeva, I. G. (2008), Evaluation of models of sound source movement in vertical plane by younger and elder listeners, in ‘XX Session of the Russian Acoustical Society’, Moscow, Russia, pp. 680–683.
- Marentakis, G., Malloch, J., Peters, N., Marshall, M., Wanderlay, M., West, S. S. & Montreal, H. A. (2008), Influence of performance gestures on the identification of spatial sound trajectories in a concert hall, in ‘Proceedings of the 14th International Conference on Auditory Display’, Paris, France.

- McLeran, A., Roads, C., Sturm, B. & Shynk, J. (2008), Granular sound spatialization using dictionary-based methods, *in* ‘Proceedings of the 5th Sound and Music Computing Conference’, number 1, Berlin, Germany.
- Middlebrooks, J. C. & Green, D. M. (1991), ‘Sound localization by human listeners.’, *Annual review of psychology* **42**, 135–59.
- Miller, P. (2009), Stockhausen and the Serial Shaping of Space, PhD thesis, University of Rochester.
- Morgan, R. P. (1980), ‘Musical Time/Musical Space’, *Critical Inquiry* **6**(3), 527–538.
- Neuhoff, J. G. (2004), *Ecological psychoacoustics*, Elsevier Academic Press, San Diego, CA, USA, chapter Interacting perceptual dimensions.
- Normandeau, R. (2009), ‘Timbre Spatialisation: The medium is the space’, *Organised Sound* **14**(03), 277–285.
- Normandeau, R. (2010), ‘The visitors and the residents’, *Musica/Tecnologia* **4**, 59–62.
- Oliveros, P. (2003), Acoustic and virtual space as a dynamic element of music, *in* J. Malloy, ed., ‘Women, Art, and Technology’, MIT Press, chapter 14, pp. 212–223.
- Otondo, F. (2008), ‘Contemporary trends in the use of space in electroacoustic music’, *Organised Sound* **13**(01), 77–81.
- Payri, B. (2010), Limitations in the Recognition of Sound Trajectories as Musical Patterns, *in* ‘Proc. of the 7th SMC Conference’, Barcelona, Spain, available online: <http://smcnetwork.org/files/proceedings/2010/11.pdf> (Accessed April 2011).
- Peters, N. (2010), Sweet [re]production: Developing sound spatialization tools for musical applications with emphasis on sweet spot and off-center perception, PhD thesis, McGill University.
- Petersa, N., Lossiusb, T. & Schacherc, J. (2009), A stratified approach for sound spatialization, *in* ‘Proceedings of the 6th Sound and Music Conference’, Porto, Portugal.
- Pierce, J. S. (1959), ‘Visual and Auditory Space in Baroque Rome’, *Journal of Aesthetics and Art Criticism* **18**(1), 56–57.
- Potard, G. & Burnett, I. (2004), Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays, *in* ‘Proc. of the 7th Int. Conf. on Digital Audio Effects’, pp. 5–8.
- Pulkki, V. (2001), Spatial Sound Generation And Perception By Amplitude Panning Techniques, Phd thesis, Helsinki University of Technology.
- Ramakrishnan, C. (2009), ‘Zirkonium: Non-invasive software for sound spatialisation’, *Organised Sound* **14**(03), 268.

- Reyleigh, L. (1875), 'On Our Perception of the Direction of a Source of Sound', *Proceedings of the Musical Association* **2**, 75–84.
- Reynolds, C. W. (1987), 'Flocks, herds and schools: A distributed behavioral model', *ACM SIGGRAPH Computer Graphics* **21**(4), 25–34.
- Reynolds, R. (1978), 'Thoughts on Sound Movement and Meaning', *Perspectives of New Music* **16**(2), 181–190.
- Roads, C. (1985), John Chowning On Composition, in 'Composers and the Computer', William Kaufmann, Inc., Los Altos, CA, USA, pp. 17–25.
- Robjohns, H. (2001), 'You Are Surrounded', *Sound On Sound* (August), available online: <http://www.soundonsound.com/sos/Aug01/articles/surroundsound1.asp> (Accessed in August 2011).
- Rusin, D. (1998), 'Topics on sphere distributions'.
URL: <http://www.math.niu.edu/~rusin/known-math/95/sphere.faq> (Accessed in May 2011)
- Santana, H. (2001), ex tempore Terretêktorh : Space and Timbre , Timbre and Space, in M. Solomos, ed., 'Présences de Iannis Xenakis', Vol. IX, Centre de documentation de la musique contemporaine, Paris, France.
- Schafer, R. M. (2007), 'Acoustic Space', *Circuit: musiques contemporaines* **17**(3), 83–86.
- Schmele, T. (2010), Inharmonic Source Study II: Hypercolor, Master's thesis, University of Limerick.
- Schumacher, M. & Bresson, J. (2010), 'Spatial Sound Synthesis in Computer-Aided Composition', *Organised Sound* **15**(03), 271–289.
- Sengpiel, E. (n.d.), 'Forum für Mikrofonaufnahmetechnik und Tonstudioteknik'.
URL: <http://www.sengpielaudio.com/> (Accessed in March 2011)
- Shinn-Cunningham, B. G. (2000), Distance cues for virtual auditory space, in 'Proceedings of the IEEE 2000 International Symposium on Multimedia Information Processing', Sydney, Australia, pp. 227–230.
- 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**(01), 35–58.
- Smalley, J. (2000), 'Gesang der Jünglinge: History and Analysis'.
URL: <http://www.music.columbia.edu/masterpieces/notes/stockhausen/GesangHistoryandAnalysis.pdf> (Accessed in June 2011)

- Smith, J. & Smith, D. (2006), 'The Patria Design Project'.
URL: <http://www.patria.org/pdp/ORDER/OVERVIEW.HTM> (Accessed in August 2011)
- Solomon, J. W. (2007), *Spatialization in Music: The Analysis and Interpretation of Spatial Gestures*, PhD thesis, University of Georgia.
- Speigle, J. M. & Loomis, J. M. (1993), Auditory Distance Perception by Translating Observers, *in* 'Proceedings of the IEEE Symposium on Research Frontiers in Virtual Reality', San Jose, CA, USA, pp. 92–99.
- Stefani, E. & Lauke, K. (2010), 'Music, Space and Theatre: Site-specific approaches to multi-channel spatialisation', *Organised Sound* **15**(03), 251–259.
- Sterken, S. (2001), 'Towards a Space-Time Art: Iannis Xenakis's Polytopes', *Perspectives of New Music* **39**(2), 262–273.
- Stockhausen, K. (1959), 'Musik im Raum', *Die Reihe* **5**, 67–82.
- Stockhausen, K. (1964), Die vier Kriterien der Elektronischen Musik, *in* 'Selbstdarstellung: Künstler über sich ed. Wulf Herzogenrath', Düsseldorf, Germany, transcribed lecture at the Folkwang-Museum, Essen, Germany, available online: http://www.elektropolis.de/ssb_story_stockhausen.htm (Accessed in August 2011).
- The Forsythe Company (n.d.), 'Motion Bank – a context for moving ideas'.
URL: <http://motionbank.org> (Accessed in June 2011)
- Theile, G. (2004), Wave field synthesis – a promising spatial audio rendering concept, *in* 'Proc. of the 7th Int. Conference on Digital Audio Effects', Neaples, Italy, pp. 125–132.
- Torchia, R. H. & Lippe, C. (2004), Techniques for Multi-Channel Real-Time Spatial Distribution Using Frequency-Domain Processing, *in* 'Proceedings of the 2004 conference on New interfaces for musical expression', Singapore, Singapore.
- Trevor, W. (1986), *The Language of Electroacoustic Music*, Palgrave Macmillan, Basingstoke, UK, chapter Sound Symbols and Lanscapes.
- Trochimczyk, M. (2001), 'From circles to nets: On the signification of spatial sound imagery in new music', *Computer Music Journal* **25**(4), 39–56.
- Vaggione, H. (2001), Composing Musical Spaces By Means of Decorrelation of Audio Signals, *in* 'Addendum of the COST G-6 Conference on Digital Audio Effects', Limerick, Ireland.
- Valiquet, P. (2009), There will always be a you: Experiences with Game of Life's mobile Wave Field Synthesis system, *in* 'Sound, Sight, Space and Play Postgraduate Conference 2009', Leicester, UK, pp. 25–36.

- Vartanyan, I. A. & Andreeva, I. G. (2007), ‘A psychological study of auditory illusions of approach and withdrawal in the context of the perceptual environment’, *The Spanish Journal of Psychology* **10**(2), 266–276.
- Verron, C., Aramaki, M., Kronland-martinet, R. & Pallone, G. (2009), Spatialized Synthesis of Noisy Environmental Sounds, in ‘Computer Music Modeling and Retrieval’, pp. 392–407.
- Voss, P., Lassonde, M., Gougoux, F., Fortin, M., Guillemot, J.-P. & Lepore, F. (2004), ‘Early- and Late-Onset Blind Individuals Show Supra-Normal Auditory Abilities in Far-Space’, *Current Biology* **14**, 1734–1738.
- Waller, D. (1999), ‘Factors affecting the perception of interobject distances in virtual environments’, *Presence: Teleoperators and Virtual Environments* **8**(6), 657–670.
- Weale, R. (1999), ‘Discovering how accessible electroacoustic music can be: The intention/reception project’, *Organized Sound* **11**(2), 189.
- Wilson, S. & Harrison, J. (2010), ‘Rethinking the BEAST: Recent developments in multichannel composition at Birmingham ElectroAcoustic Sound Theatre’, *Organised Sound* **15**(03), 239–250.
- Wyatt, S. A. (1999), Investigative Studies on Sound Diffusion/Projection at the University of Illinois: a report on an explorative collaboration, in ‘Soudning Out 5’, Montréal, Canada, available online: <http://cec.sonus.ca/econtact/Diffusion/Investigative.htm> (Accessed in August 2011).
- Xenakis, I. (1967), *Eonta*, Boosey & Hawkes. (Sheet Music).
- Yildizeli, H. (2010), ‘Spatial music – electronic sound: A history of spatial music’.
URL: <http://knol.google.com/k/spatial-music> (Accessed in June 2011)
- Zahorik, P. (2002), ‘Assessing auditory distance perception using virtual acoustics’, *The Journal of the Acoustical Society of America* **111**(4), 1832–1846.
- Zainea, L. (n.d.), ‘Musical space as a metaphorical inside/outside’.
URL: <http://www.zainea.com/morphspace.pdf> (Accessed in June 2011)
- Zelli, B. (2001), *Reale und virtuelle Räume in der Computermusik*, PhD thesis, Technische Universität Berlin.
- Zelli, B. (2009), Spatialization as a Musical Concept, in ‘IEEE International Symposium on Mixed and Augmented Reality 2009 Arts, Media and Humanities Proceedings’, Orlando, FL, USA, pp. 35–38.
- Zelli, B. (2010), *Musique Acousmatique and Imaginary Spaces*, in ‘Soudning Out 5’, Bournemouth, UK, available online: http://www.bijanzelli.com/MusiqueAcousmatiqueandImaginarySpaces_econtact13.3.pdf (Accessed in May 2011).

Zvonar, R. (2006), 'A history of spatial music', *eContact!* **7**(4).

URL: http://cec.concordia.ca/econtact/Multichannel/spatial_music.html
(Accessed in June 2011)

Zwiers, M. P., Van Opstal, A. J. & Cruysberg, J. R. M. (2001), 'A Spatial Hearing Deficit in Early-Blind Humans', *The Journal of Neuroscience* **21**.