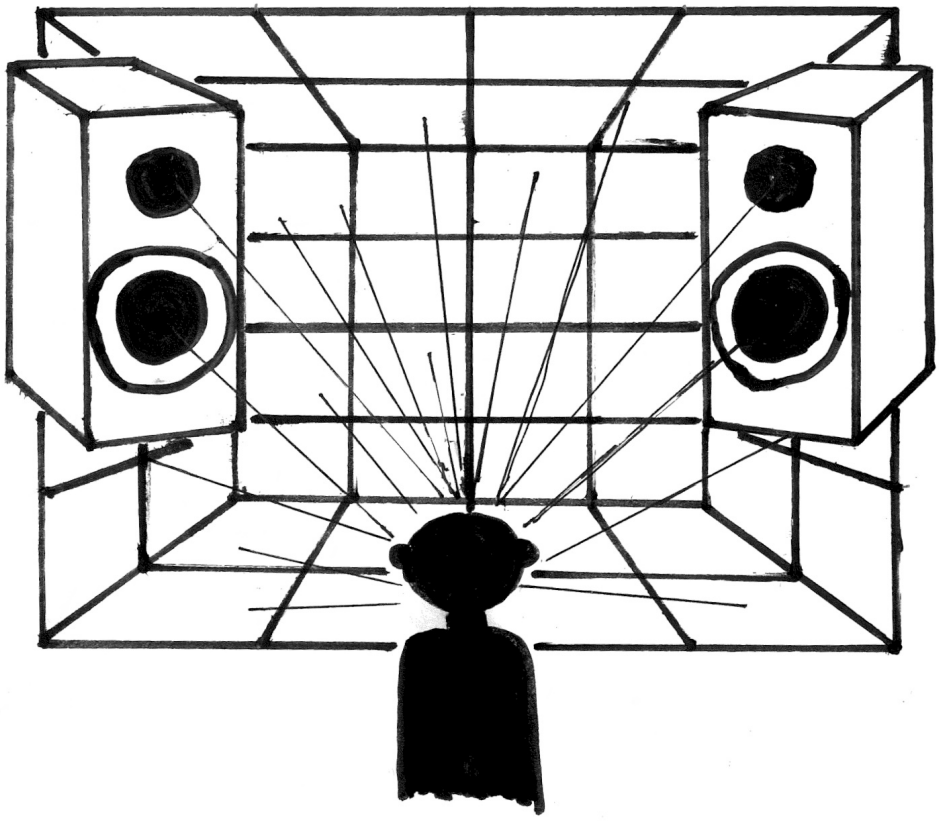
Spaces of Sound

Meanings of Spatiality in Recorded Sound

Emil Kraugerud



Master's thesis in musicology

Department of Musicology

UNIVERSITY OF OSLO

Spring 2016

Cover illustration: “Listener in sound-space”, Emil Kraugerud.

Abstract

This thesis concerns the importance of forming the virtual space in recorded popular music. Several researchers have addressed the spatial aspects of recorded music, with a focus on how such aspects contribute to the sound of a recording. However, little attention has been given to the ways in which the spatial shaping of recorded sound affects the interpretation of meaning in a recording. In this thesis I address this gap by introducing a new concept called *the sound-space*. The sound-space is a production-analytic model that comprises both the four-dimensional spatiotemporal organization of previous models for sound analysis, and the ways in which that organization evokes or emphasizes meanings in the recorded music. Through the discussions in this thesis I will attempt to get a grasp on the ways in which spatiality can be used as a basis for investigating how meaning is affected in record production, and how spatiality is possibly experienced through previous corporeal experiences with different spaces.

In order to demonstrate how the sound-space can be applied as a tool in record production I have also recorded and mixed a song and compared it to previous a version that was produced before the work on this thesis started. In the production of the new version I based my decisions regarding recording and mixing techniques in the sound-space model. Through a comparative analysis of the two productions I elucidate the effects of this application of sound-space as a record production tool for the new version, in particular emphasising the ways in which it affects the narrative(s) of the song, and the corporeal engagement of the listener, that is, how the listener is becoming corporeally “involved” in the sound-space of the recording.

Acknowledgements

This thesis could not have been completed without the help and support of a number of people. First of all, I owe profound gratitude to my two excellent supervisors, Anne Danielsen and Eirik Askerøi, for thorough critical reading and constructive feedback. Both have invested a lot of time in this project, and their engagement for my work has been a true source of inspiration. Thanks to Kevin Steinman for copyediting this thesis and thus considerably improving my language. Thanks also to Professor Allan F. Moore for clarifying answers regarding the "diagonal mix". A draft of this thesis was presented at the 10th Art of Record Production conference at Drexel University, Philadelphia in November 2016. I want to thank the organizers of the conference as well as other participants for valuable feedback and inspiring conversations. I also want to thank current and former co-students for rewarding conversations and opinions. Special thanks go to Bjørnar Sandvik for being a valuable sparring partner as well as a great travel companion to the ARP conference. I am very grateful to my good friends and co-musicians in Dingus for letting me use our music in this thesis, and for trusting me with the responsibility of recording and mixing it. I have also benefited a lot from recording a range of fantastic musicians over the past two years. Lastly, I want to express my profound appreciation to friends and family for their invaluable support.

Oslo, April 24, 2016

Emil Kraugerud

Contents

1 Introduction

When I first started taking courses in popular music and record production I quickly became aware of the potential of spatial approaches to popular music. Ragnhild Brøvig-Hanssen and Anne Danielsen's paper "The Naturalised and the Surreal" (2013) introduced me to some of the many ways in which a recorded space can reveal itself to the listener. As a guitar player I was at the time very much into using delay effects for expanding the possibilities of the instrument, something that contributed to my discovery of Peter Doyle's book *Echo & Reverb* (2005), in which he investigates some of the cultural significations of spatiality in pre-stereo music. I brought these ideas into my work in record production, where I was particularly interested in how different types of timbre can evoke specific time periods, for example the ways in which the sonic characteristics of an old plate reverb or tape echo can reference specific eras.

A recording consists of a diversity of sounds, separated by their different timbres and spatialities. These timbres and spatialities are affected by, among many things, how musicians use their instruments, how the instruments are recorded, and how the recorded tracks are manipulated in post-production. In record production, we often hear producers and engineers talk about doing what is best for the song, whether it is guitar sound, drum tuning, microphone choice, or something else. This also applies to other features more directly linked to spatiality, for example panning, reverb and delay. But why is it the best for the song? Does it emphasize rhythmic or harmonic elements of the song, or maybe it is something that underlines the contents of the lyrics? Is it conceivable that particular types of spatiality refer to something outside of the song, and perhaps also something entirely extra-musical? And what about the listeners, do they notice the results of such decisions? Denis Smalley suggests a reason for the lack of listeners' focus on spatiality in music: "in my experience we are not that used to listening out for spatial attributes, for spatial forms, and space-form, partly because there is so much else to listen out for" (Smalley 2007: 35). So, what is the point of discussing spatiality in music when only a minority of people are able to hear it? One of the wonders of music, I think, and this applies to all musics, is that you can listen to the same piece over and over, and every time discover something new. Thus, spatiality can be regarded as a hidden treasure in recorded music, giving the listener a rewarding experience when she discovers it.

1.1 Aims and research questions

Based on existing approaches to spatiality, it could be interesting to investigate if such an approach can be used as a basis for further investigation of why recorded popular music is so meaningful to so many people. Spatiality is something that everyone relates to at all times, but not necessarily on a conscious level. It is embedded in our lives to such an extent that we don't need to think about it. This may explain why it is also often ignored when listening to recordings. In a recording, the forming of spatiality is affected by such factors as frequency spectrum, panning, reverb and delay effects, relative volume, distortion, and much more. Recording and mixing engineers are thus in great capability of forming the spatiality in the ways they, or the producer, want to. Against this backdrop, I will base the discussions in this thesis on the following research questions: How can a new production-analytic model elucidate the effects of spatiality in a record production? And, how can such a model contribute to an understanding of the ways in which producers and engineers work with space to create interesting productions? Answering these questions will also include a discussion of how listeners' experiences with body movements in different spaces are relevant to their experience of recorded music.

1.2 Theory and method

1.2.1 Theoretical background

My research can be placed in the rather young theoretical field of record production studies, which in one way is closely connected to the study of popular music. Popular music studies are also a rather recent field in musicology, emerging in the 1980s as a reaction to the devaluation and neglect of popular music in traditional musicology. Several early examples of popular music studies have either focused on extra-musical factors in popular music culture, or they have analyzed the music in terms of "the canonical 'masterpieces' of the Western tradition" (Moore 2001: 9f). Richard Middleton addresses this problem in *Studying Popular Music* (1990), where he argues for a "new" musicology tailored for popular music. He claims that traditional musicology had a vocabulary unable to describe essential parameters of popular music, and thus devaluating or even neglecting it. Also, he argues that terms for music analysis are commonly ideologically loaded, and that these terms "involve selective, and often unconsciously formulated conceptions of what music is" (Middleton 1990: 104). Middleton's point is that popular styles must be analyzed on their own terms.

This means that popular music analysis must apply a different vocabulary and different terms from the analysis of Western art music. Middleton thus concludes that a new musicology of popular music should address popular music on its own terms.

Another contribution to the studies of popular music is Allan F. Moore's *Rock: The Primary Text* (1993, revised 2001). With the subtitle "Developing a Musicology of Rock", Moore addresses the problems of previous studies of rock, in which the focus was more on rock culture than the music itself. As suggested by the title, Moore looks to the primary text in rock, which is the music. He addresses the lacking focus on the music itself by proposing a new analytical model, more or less specific to rock, and by addressing the importance of what he calls "texture". The analytical model he proposes stratifies sound-sources into four layers, based on a notion that "[t]he stream of sounds a listener hears is composed of rhythm *and* harmony *and* melody *and* instrumental timbre *and* lyrics", and "[t]hese basic elements are distinguishable one from another in the abstract, and on reflection […], but they conspire to produce the music we hear" (Moore 2001: 33, emphasis in original). Thus, music, according to Moore, cannot be heard as only melody or rhythm or harmony etc., but should be studied as a whole. The texture of rock music, Moore writes, "refers to the presence of and relationships between identifiable strands of sound in a music" (ibid.: 121). Moore sees this in light of technology, particularly the way in which sounds have "been layered through multi-tracking and manipulated through 'black boxes'" (ibid.). Moore thus emphasizes the sound of recorded music as vital to rock, as he deals partly with the analysis of record production.

The study of record production is also at the heart of the annual Art of Record Production conference (ARP), its related journal and the Association for the Study of the Art of Record Production. In the introduction to *The Art of Record Production* (2012), Simon Frith and Simon Zagorski-Thomas offer their thoughts on ARP. What distinguishes ARP from other approaches to record production is its cross-disciplinary dialogue between academics and practitioners. Scholars participating at the ARP conference represent a variety of academic fields, including musicology, sociology, anthropology, cultural theory, ethnomusicology, psychology, history, electrical engineering, psychoacoustics, literary theory and history of science (Frith & Zagorski-Thomas 2012: 2). The study of record production is thus, like popular music, multi-disciplinary and not a discipline in itself (ibid.). The inclusion of practitioners means that there is also hands-on knowledge of the field, meaning that ARP should be able to grasp several aspects of record production. As Frith and Zagorski-Thomas point out, "[i]n the studio technical decisions are aesthetic, aesthetic

decisions are technical, and all such decisions are musical" (ibid.: 3). In other words, the different approaches are necessary for an adequate study of record production.

It seems like the study of popular music is increasingly connected to the study of record production. This has to do with the central role of recorded sound in popular music, and the increasing recognition of this factor. It also means that many contributions to the study of popular music can be regarded as studies of record production as well, thus blurring the line between the two. This thesis, though, should be easily categorized as a study of record production, as it focuses on the processes of production rather than the music resulting from these processes. Nevertheless, it will largely draw upon a theoretical framework from popular music studies.

1.2.2 Some reflections on method

This thesis can be divided into two main parts: one larger part in which a theory of recorded spatiality is presented, and one shorter part that concerns practical work and analyses. The theoretical part will concern the proposition of a new approach to thinking of the recorded space. As suggested in the foregoing section, I will draw on theories from the fields of popular music studies and record production studies.

In the practically oriented part of this thesis, I will be analyzing two recordings produced, engineered and mixed by myself. The two recordings are of the same band and the same song. I also played electric guitars on both recordings. The recordings took place at different points of time, over a year apart. In the meantime I had started the work on the ideas presented in this thesis. Although I had certain thoughts on the use of references to other recorded spaces during the production of the first recording, it was only on the second recording I employed thoughts on spatiality as a basis for most decisions during recording and mixing. Thus, by analyzing and comparing the two recordings, I will try to distinguish some differences between them. In the case of the second recording, I will also look briefly into my initial intentions of emphasizing different transtextual meanings before the production, and how these intentions are translated to the finished production. Thus, I will investigate how these intended meanings come forth when compared to the first recording.

An approach like this can remind of what is called practice-led research, where the researcher takes a central part in the investigated process, instead of coming in as an outsider. The investigated process is followed systematically through reflection and analysis. A method like this has some downsides, as it may involve a high level of subjectivity.

Accordingly, my conclusions will necessarily be somewhat saturated by my preconceptions about the productions. On the other hand, being involved to such an extent in the process offers unique insight which could not be gained through other research methods. Thus, my approach has its pros and cons, but I believe it can provide some interesting thoughts on transtextual relations embedded in the recorded space. A further discussion on the practice-led research in this thesis follows in the first part of chapter 4.

1.3 Delimitations

The subject of spatiality in recorded music is a broad one. It opens up for a wide spectrum of approaches. Certain limitations to what is investigated in this thesis are thus necessary, some of which I will discuss in this section.

As implied in the introduction to this thesis, my main concern here is recorded music, and more specifically, music that is recorded and mixed for stereo playback. Although there are obvious correlations between live and recorded spatiality, these two realms have widely differing bases. Whereas live spatiality necessarily is based on a connection with visual spatiality, recorded spatiality is not bound by such rules, and can thus be totally "unrealistic". On the other hand, recorded spatiality has been formed by the context in which it has evolved, namely as an imitation of live spatiality, and thus it does in many cases not fulfill its potential for "unreality".

The most obvious reason for focusing on stereo playback, leaving out mono and various multi-channel systems (such as 5.1 and 7.1 surround), is that stereo is, and has been for a long time, the most widely used playback format. Furthermore, as I see it, this is not to change in the near future. As Théberge, Devine and Everrett write in the introduction to *Living Stereo* (2015), "[s]tereo is a living part of sound culture" (Théberge, Devine, & Everrett 2015: 1). The increased use of streaming services enables people to listen to anything anywhere, mostly through headphones. At the same time, there is at the moment a revival of vinyl records, restricting listeners to stereo and mono playback. In the case of home playback systems, a stereo system requires much less space than a surround system, and it is easier to set up. As Eric "Mixerman" Sarafin comments, "it's difficult enough to find a household with an actual established center position in a proper stereo field – we're going to erect four speakers in perfect symmetry how?" (Mixerman 2010: 66).

Furthermore, I believe that a stereo speaker setup is capable of simulating all the dimensions of height, depth and width. The depth and width dimensions of a stereo system

are different from those of a surround system in that the sound emerges from in front of the listener instead of around the listener. This is not unlike going to a concert, and it may therefore be easier to get used to for the listener than surround playback. The exception would be headphone listening, where the sound comes from the sides of the listener's head and may even be perceived as being *inside* his head. Based on my own experiences with headphone listening, however, listeners, at least to some extent, tend to adjust their perception of spatiality in music as if the sound was emerging in front of them, as with stereo speaker setups. There is thus some correlation between the two different kinds of stereo playback.

One could of course elaborate more on different variations of playback, for example different digital and analog formats (CDs, mp3s, wav-files, vinyl records, cassette tapes, etc.), and how these affect the spatiality of a sound. However, although these formats probably have differing effect on spatiality, this effect is in most cases so small that most listeners would ignore it. Moreover, such considerations would (mostly) be the concern of mastering engineers, and thus be outside of the roles of producers and mixing engineers, which is my concern in this thesis.

Dealing with recorded spatiality, a certain focus on how the recorded space is perceived by listeners seems only natural. However, it is also tempting to dive deeper into the realm of music cognition, for example by investigating how the brain reacts differently to different kinds of spatiality. Although highly interesting, such studies are beyond the scope of this thesis. I will however touch on the subject of how virtual recorded spaces can act as references for real spaces, and how corporeal experiences of real spaces can be active in the perception of recorded spaces. This may be seen as a kind of embodied cognition. I will keep this section relatively brief, aware that I will be leaving out a big part of a large field of research. I believe in any case that it is important to address this issue, as it points out something about how recorded spatiality is perceived.

1.4 Terminology

Some basic terms are important for understanding some of the concepts that will be used later on in the thesis. This section is devoted to their explanation, as well as an introductory explanation of the term sound-space.

1.4.1 Sound

The term *sound*, is, and has been, used in several different ways, and can therefore be difficult to comprehend. Since the latter half of the 20th century, several definitions have been proposed, providing different approaches. Per Erik Brolinson and Holger Larsen's book *Rock…* (1981) provides an early attempt at finding a definition of sound. Before suggesting their own definition, they refer to some previous contributions to the discourse, amongst others Wolfgang Sandner and Dave Laing. Sandner suggests that sound describes the total musical impression, as well as certain indirectly related factors such as visual impressions of amplifiers etc. (Sandner referred to in Brolinson & Larsen 1981: 181). As Brolinson and Larsen point out, this definition appears rather vague (ibid.). It says nothing about the importance of the different components for "the total musical impression". Furthermore, they find it unreasonable to include visual factors (ibid.). Laing, on the other hand, suggests that sound "includes the role of each instrument, and the particular roles of the voices" (Laing cited in ibid.). Laing is thus more directed to the individual elements of a sound than Sandner, but Laing's focus tends to leave out the overall aspect of the sound of a song. Brolinson and Larsen arrive at a middleground between Sandner's and Laing's approaches, describing sound as referring to the fundamental character of all musical elements the way it appears in a very short time segment of the music, while also being characteristic of a larger continuous section (ibid.). The authors add that this definition excludes any unrepeated events within a song, but still includes the possibility of the sound varying throughout it (ibid.).

Further on, they suggest a "rhythmic and timbral structuring of the accompanying layer" (ibid.: 183) as definitive of a "disco sound", and vocal timbre, melody and phrasing as "active parameters to characterize a 'soul sound'" (ibid.: 184). This implies that the very short time segment in which one can recognize a sound must have a certain length, since a rhythmic structure or a melody needs some time to evolve for the listener to be able to recognize it.

Eirik Askerøi offers a different view on the sound term. He suggests that, "[a] song's sound […] should be seen not only as the sum of its sonic events but also as yet another way to categorise it" (Askerøi 2013: 26). He emphasizes this by referring to Paul Théberge's claim that "individual 'sounds' have come to carry the same commercial and aesthetic weight as the melody and the lyric of pop songs" (Théberge cited in ibid.). In other words, both an overall sound and its individual sounds might carry the same "categorizing role" as melody and lyrics. Askerøi does not seem to conflate sound with genre or musical style, but he

applies the term as another way of determining the genre of a song, based on the notion that different genres have different sounds related to them. In popular music styles, it might even be easier to distinguish a genre based on sound than melody and lyrics.

I will base my further discussions in this thesis on the assumption that there are two main types of sound by which recorded popular music can be categorized: a *global* sound and a *local* sound. This distinction is based on Lelio Camilleri's notion of global and local *sonicprints*, where global sonicprints are the "sonic signatures" of an overall sonic image, whereas local sonicprints refer to particular sounds that characterize a recording (Camilleri 2010: 200). My use of the term global sound also relates to Brolinson and Larsen's definition of sound, and is meant to suggest that the sum of individual parameters (rhythm, timbre, harmony, melody, production aesthetics, etc.) in a song constitutes an overall sound. In most cases, the global sound is closely related to genre. A local sound, on the other hand, is constituted by individual parameters (timbre, playing style, etc.) in for example the sound of an instrument. A local sound can be related to genre, but it does not have to. Different local sounds in a song have direct influence on the song's global sound. In the context of this thesis, parameters like rhythm, harmony and melody will not be in focus. I will rather concentrate on how local sounds include, or constitute, spatiality, and how the composition of such local spatialities affects a global spatiality.

1.4.2 Recorded space

In this thesis I will employ an understanding of the *recorded space* as a virtual space, based on an illusion of three spatial dimensions.[1] Brøvig-Hanssen and Danielsen draw on Denis Smalley's theory of *source bonding* and James J. Gibson's theory of *ecological perception* to suggest that such spaces are interpreted by the listener by unconsciously comparing them to previous experiences with real spaces and how sounds behave in them (Brøvig-Hanssen & Danielsen 2013: 71). That does not, however, mean that recorded spaces have to resemble real spaces, even though this is often the case – at least to some extent. A recorded space can be fundamentally different from any possible real spaces, or it can be composed of multiple "physically possible" spaces in a way that would never occur in reality. It is the reference to real spaces that makes such spaces appear as unrealistic (ibid.: 75).

[1] As I will discuss further in chapter 3, the spatial dimensions are based on illusions to differing degrees. For example, the width dimension is mostly based on the physical location of the two speakers (or earbuds) of a stereo system, and the difference between the sound emerging from each speaker. In contrast, the height dimension is based on the culturally embedded notion that bright sounds are high and dark sounds are low.

It must be mentioned that any recorded space will be fundamentally unreal, as they are affected by such factors as microphone placement and choice, microphone preamps and other types of signal processing, as well as playback equipment. Many recorded spaces are nevertheless perceived as natural (and maybe even realistic) by the listener, which might be because such spaces have become naturalized (Brøvig-Hanssen & Danielsen 2013: 75). I will return to the discussion of realistic and unrealistic recorded spaces in chapter 3.

1.4.3 Sound-box, sound-room, sound-space, and space-form

In the definition of *sound-box* in *Rock: The Primary Text*, Moore comments that the sound-box might seem similar to what producers and engineers call "the mix", but whereas the mix takes the view of the producer, the sound-box is meant to privilege "the listening, rather than the production, process" (Moore 2001: 121). Why, then, do I find it adequate to base a study of record production in a term relating primarily to record listening?

A central aspect of the concept of sound-space is the idea of the producer or engineer as a listener, a listener who interacts with the music through reflection and association. Basing this concept on Danielsen's *sound-room* ("lydrom"; 1993) and Moore's sound-box, instead of "mix", I mean to emphasize this. It sheds light on the active and critical listening that is a vital part of record production, and on which technical and aesthetic decisions are based. Additionally, the terms "sound-box" and "sound-room" tend to draw attention to the spatial aspects of production to a greater extent than "mix". Still, I will refer to the technical and creative processes of forming the aesthetics of previously recorded tracks as "the mix". "Sound-box" or "sound-room" could be proper terms for referring to the spatial effects of these processes.

When I nonetheless will apply the term "sound-space" for this purpose, it is because of the associations connected to each term. Regarding the sound-box, a box is not something in which one usually experiences sounds, if so it would most likely be perceived in a negative way ("a boxy sound"). The sound-box must therefore be regarded as a very abstract metaphor, as it has little to do with normal listening conditions. The notion of a sound-room meets this expectation, as it relates to the illusion that sounds in a recording propagate in a room. However, the term, when strictly interpreted, excludes spatial environments that are not reminiscent of rooms. "Lydrom" can also be translated to "sound-*space*", which resonates better with Danielsen's use. Sound-space allows for interpretation to a greater extent than sound-room – a space can be any variation of either an open or an enclosed space (including

a box or room). It is thus more linkable to experiences of real spaces than the sound-box. It is also not restricted to three-dimensional spaces. The *Merriam-Webster* dictionary defines space as "a limited extent in one, two, or three dimensions" (*Merriam-Webster* 2016). The term sound-space thus opens for two-dimensional spaces such as monophonic recordings. In order to distinguish between Danielsen's concept and my own, I will use the term sound-room when discussing her concept.

Another relevant term for investigating the spatial properties of a recording is Smalley's *space-form* (2007). Smalley applies the term to look at the spatial organization of sounds in acousmatic music: "Space-form in acousmatic music is an aesthetically created 'environment' which structures transmodal perceptual contingencies through source-bondings and spectromorphological relations" (Smalley 2007: 40).[2] Although Smalley focuses on acousmatic music, the concept of space-form also applies to acousmatic listening (listening without a visible sound-source, like record listening) in general.

The terms space-form, sound-box, and sound-room all relate to how recorded or acousmatically performed sound is organized in terms of space. What they do not investigate to any greater extent are the effects of different forms of spatial organization on the listener. The exception would be Brøvig-Hanssen and Danielsen's use of sound-room and their investigation of the naturalization of surreal virtual spaces. What I seek to investigate with the sound-space, however, are the ways in which the forming of the recorded space by producers and engineers acts as a carrier of meaning, in particular its ability of referring to other musics, other spaces, and other factors.

1.5 Outline of the thesis

This thesis can be divided into two main parts, where part 1 comprises chapter 2 and 3, and part 2 consists of chapter 4. Part 1 will concern the discussion of different theoretical concepts related to the spatiality of recorded sound and how we make sense of it. In chapter 2 I will discuss the basis of understanding the recorded space as a textural three-dimensional space, starting from the analytic concepts of sound-box and sound-room. I will further look at the possible adaptation of such concepts to the realm of record production, by discussing already existing models that take a record production approach to spatiality, and comparing

[2] *Spectromorphology* refers to the ways in which sound spectra change and are shaped through time (Smalley 1997: 107).

those to the sound-box and sound-room models. Relevant to this is the question of how an understanding of spatiality can contribute to an actual production.

This will be further investigated in chapter 3, where I will turn to the discussion of how listeners make sense of recorded spaces. This includes taking into account the listener's associations with different types of spatiality. I will focus on the different ways in which we interpret recorded spaces based on our previous experiences with spatial situations, including other recorded spaces as well as actual spaces. After a brief discussion of how humans perceive spatiality through hearing, as well as a further elaboration on the different dimensions of the recorded space, I will discuss some theoretical concepts that are relevant to our perception of space. These discussions will form the basis for a production-analytic tool that may be useful not only for creating "tidy" recordings, but also for making interesting productions in which the listener can be involved.

Based on the theories to be discussed in part 1, I have recorded and mixed a song that will be the subject of chapter 4. I will analyze the outcome of that production in relation to my intentions during the recording and mixing processes, and I will compare it to a previous recording of the same song. The aim for the analysis will be to detect any differences between the two versions, based on the discussed theories, and to see if there is correlation between my intentions during the production of the new version and the perceived meanings of its final recorded space. Additionally, it may be interesting to see if there are unintended effects of the spatiality in the older version of the song, that is, if the forming of the space can be associated with anything related to the song.

2 Approaching space in record production

2.1 Sound-box and sound-room as analytical models

In his effort to address the so-called "music itself" in *Rock: The Primary Text* (1993, revised 2001), Moore introduces the sound-box model to grasp what he calls the *texture* in rock music. Texture, according to Moore, "refers to the presence of and relationships between identifiable strands of sound in a music" (Moore 2001: 121). He further describes the sound-box as "a 'virtual textural space', envisaged as an empty cube of finite dimensions, changing with respect to real time (almost like an abstract, three-dimensional television screen)" (ibid.). As pointed out above, the sound-box, according to Moore, differs from what in record production is called "the mix" in that it takes a listener's point of view rather than a producer's or engineer's (ibid.). Thus, Moore's approach focuses on the perceived effects of production in finished recordings. Moore further describes the dimensions of the sound-box:

> All rock has strands at different vertical locations, where this represents their register. Most rock also attempts a sense of musical 'depth' (the illusory sense that some sounds originate at a greater distance than others), giving a sense of textural foreground, middleground and background. Much rock also has a sense of horizontal location, provided by the construction of the stereo image (ibid.).

Moore further suggests a notion of density as a parameter of the sound-box model. As he remarks, the types and degrees of density filling the sound-box are the most important features of it (ibid.). This includes the presence or lack of "holes" in this space, denoting "potential areas left unused" (ibid.). Holes can appear in all three spatial dimensions, as results of relatively few or "narrow"-sounding sound-sources, with wide spacing between stereo positions and/or frequency spectra. Thus, the sound-box appears to be a good model for analyzing texture in stereophonic popular music sound.

With the concept of sound-room ("lydrom"), Danielsen (1993) also takes a three-dimensional approach to analyzing popular music sound. She applies the term as an alternative to the Norwegian word "lydbilde", which can be translated to "sound image" or just "sound". The preference of "lydrom" rather than "lydbilde" is meant to emphasize the presence of depth in music (Danielsen 1993: 51). Similarities to the sound-box can be seen to some extent in Danielsen's description of the sound-room's dimensions. She explains that the placement along the foreground-background axis is a result of a synthesis of the parameters

reverb and volume, of close/far and loud/quiet, and that the left-right axis is the placement in the stereo image, whereas high-low is connected to frequencies and pitch (ibid.).

Thus, Danielsen's and Moore's analytical models seem quite similar in their descriptions. Both describe the texture of a song's sounds, according to the place from which they are perceived to emanate in a virtual three-dimensional space, privileging a listener's point of view. Both models are also based in the four dimensions of time, width, depth, and height. Still differences between the two models will come more clearly into focus as we explore Moore's and Danielsen's use of their concepts in analysis.

2.1.1 Sound-box

After introducing the sound-box, Moore further applies the model to trace changes of texture in rock styles. Starting with what he calls the "pre-stereo" textures in the Beatles' "She Loves You", Moore follows a development to a "normative mix" (Moore 2001: 121ff). He remarks that his first three examples, Beatles' "The Long and Winding Road" and "She Loves You", and Amen Corner's "Bend Me Shape Me", "tend to define their spaces in terms of blocks of sound rather than lines" (ibid.: 122). What Moore implies here is that sound-sources are located in three defined horizontal positions (left, center and right), instead of being placed between these positions. According to Moore, Yes' "Roundabout" is based more on individual lines, with a dense texture (ibid.).

In Def Leppard's "Love Bites", he observes what he calls the normative placement of sound-sources in the sound-box:

> Snare drum and tomtoms central with cymbals to either side, bass drum central but lower, voice and bass respectively towards the top and bottom, but on a slight diagonal through the snare drum, guitars and other instruments at some point between the stereo centre and extremes (ibid.: 123).

This type of configuration has been normative in popular music since the 1970s, hence Moore's label for it. In "Love Bites", as he comments, the heavy reverb in the song tends to fill holes that otherwise probably would appear in the sound-box (ibid.). In addition to the reverb used, the dense texture of the song also is the result of doubled, distorted guitars and close vocal harmonies. Thus, the full stereo spectrum is utilized.

Together with Ruth Dockwray, Moore takes a similar path in "Configuring the Sound-box 1965-1972" (2010). The article focuses on the crucial period from the beginning of stereo recording for popular music, up to the point when a normative mix configuration

was established. Within this period Dockwray and Moore recognize four different mix configurations, based on the placement of lead vocals, snare drum and bass guitar along the horizontal plane (Dockwray & Moore 2010: 185). These will be discussed in greater detail in chapter 3.

Central to Dockwray and Moore's analyses are a series of illustrations of the sound-box in each of their example songs. These illustrations are based on empty, rectangular cubes of finite dimensions, in which "sounds were plotted" by the authors during auditory analysis (ibid.: 184). This type of illustration is also used in Moore's more recent book, *Song Means: Analysing and Interpreting Recorded Popular Sound* (2012). Moore affords here an updated definition of sound-box, describing it as "a heuristic model of the way sound-source location works in recordings" (Moore 2012: 31). In this description Moore emphasizes the importance of the development of stereo for a full manifestation of the sound-box. He further points out that this does not exclude texture from pre-stereo (mono) recordings (ibid.). Still, as mono recordings lack a sense of width, I find the sound-box model to be limited in analyzing such sounds, as Moore chooses to focus on "the possibilities of full stereo and its three spatial dimensions" (ibid.). Doyle has a similar view on this. In the introduction to *Echo & Reverb* (2005), he points out the sound-box's limited usefulness to his own studies of pre-stereo recordings, as well as criticizing a viewing of mono sound reproduction as a limitation (Doyle 2005: 25). In Moore's use of sound-box such a view is implicit, as the target of the sound-box, its "full manifestation", seems to be accomplished by utilizing all three spatial dimensions.

In *Song Means*, Moore addresses *prominence* as a dimension of the sound-box (ibid.). Although he describes it in much the same way as he described depth in *Rock*, I would argue that this suggests a change of focus, from how far or near sounds appear, to how much they tend to "pop" out of the mix. A shift like this might be related to the density of many modern mixes, where vital sounds, such as vocals or other lead parts, need to pop out to be able to promote their message. Still, I think prominence and depth should be treated, at least partly, as two separate parameters, as they are affected by different factors. By this I mean that a sound-source can be perceived as very near without necessarily being much more prominent than other sound-sources. Prominence, I think, is also affected by the clarity of sounds. This might also relate to Doyle's discussion of the perception of depth in recordings. Doyle breaks down Moore's description of sound-box depth to cover reverberation amount and the relative amplitude of sounds (Doyle 2005: 26). However, for Doyle's own studies, this definition is

not applicable, as he deals with recordings "in which the loudest element is also the most reverberant" (ibid.).

Maybe this is where the difference between depth and prominence becomes evident. An example, also used by Doyle, is Elvis Presley's voice in "Baby, Let's Play House". The repeated echo on the voice in some ways makes it sound farther away than the dry-sounding (at least when compared to the vocals) electric guitar, even though the voice is the loudest sound. William Moylan addresses the case of prominence, writing that "*[p]rominence* is perhaps what is conspicuous, or perhaps it is the thing that draws us to hold it within the center of our attention" (Moylan 2015: 320, emphasis in original). The distinct sound of the echo effect on Presley's voice makes it more conspicuous than the more traditional sound of the electric guitar, thus drawing more attention. Moylan further claims that "[w]hat is most prominent is not necessarily most important or most significant" (ibid.). In Presley's case, the voice is both the most prominent and most important sound, and the echo, in combination with the relative amplitude of the voice, helps to draw attention to the vocal sound and the lyrics. Thus, a sound can appear as prominent and far away at the same time. To describe such relationships in the context of the sound-box model, it could be useful to incorporate "internal" sound-boxes, which I will come back to in the discussion of Danielsen's sound-room.

Further on in *Song Means*, Moore offers an update of his sound-box analyses from *Rock*, following the development of stereo sound from 1964 to the late 1980s. He uses several of the same examples as in *Rock*, but with certain updates. In his analysis of Yes' "Roundabout" he now suggests that the density of the song "makes the texture appear as blocks of sound rather than simply points" (Moore 2012: 40). Except from mentioning the normative positioning of the drum kit – central in the stereo field but spread wide – Moore's focus in this short analysis is the relationship between Steve Howe's acoustic guitar and the rest of the band:

> At the very beginning of the track, the guitar is very much to the fore, with cleanly audible chordal harmonics (sounds that are naturally very soft). In this aural context, the entry of the band (44") should be much louder than it actually is – clearly the guitar is brought to our attention through being "unnaturally" loud. Later in the track, the guitar reprises its early material, but now with an organ backing (4'58"). Again, one would expect the organ to drown the guitar, but the balance is very much in the opposite direction (ibid.).

It must also be mentioned that the difference in depth between organ and acoustic guitar in the addressed part is affected by two factors. Firstly, the organ is played in a high register and

without the use of its lower drawbars (resulting in a reduction of bass frequencies), which makes it less proximate in comparison to the full-sounding guitar. Secondly, there is also a significant amount of reverb added to the organ, whereas the guitar is kept rather dry. Similarities to the case of prominence in Elvis' "Baby, Let's Play House" are apparent here. Now the guitar is unnaturally loud compared to the organ, just as Elvis' voice was unnaturally loud compared to the electric guitar.

Nevertheless, these "unnatural" proximity differences do not necessarily appear as surprising to modern ears. In her doctoral dissertation, Brøvig-Hanssen (2013) addresses what she calls *opaque* and *transparent mediation*, with mediation functioning as "a comprehensive generic term for the technological mediation that happens within the production process of music – that is, the processes of recording, editing, and treating sounds with various signal-processing effects" (Brøvig-Hanssen 2013: 14). Opaque and transparent are terms used to define the degree of mediation detected by the listener: "transparent mediation implies that the listener ignores the mediation" (ibid.: 17). Brøvig-Hanssen further describes the gradual naturalization from opaque to transparent. An example she applies is that of crooning and the mediation of the human voice through microphones:

> [I]t was at first regarded as a profoundly opaque mediation, since the intimate voice had never before been able to penetrate in a concert hall. This familiar-made-unfamiliar vocal sound did not correspond to the singer's spatial location either; though the vocalist sang from a stage in a concert hall, far away from the listener, it sounded as if he or she was sitting right next to the listener. Of course, as listeners grew accustomed to such live performances, the microphone-staged voice gradually came to stand for the musical voice itself (ibid.: 18f).

Similarly, the sound of the relatively loud guitar in "Roundabout" might have appeared as opaque to listeners in the early 1970s, but our ears have grown accustomed to such amplitude relationships since then.

The next example in Moore's demonstration of the sound-box is Fleetwood Mac's "Little Lies". As he comments, the song "demonstrates the degree of control of textural space available to producers by the late 1980s" (Moore 2012: 40). The many different sounds are, according to Moore, carefully separated both in register and across the stereo field, and thus there is "no sense of being swamped" (ibid.). Additionally, Moore reveals differing distances of sounds here as well. The claim that sounds are "carefully separated" suggests the producer's or engineer's active role in populating the sound-box, and he thereby reveals some kind of spatial thinking in the recording and mixing process. I will elaborate further on this later in this chapter.

2.1.2 Sound-room

As suggested in the introduction to this chapter, Danielsen approaches the spatiality of recorded music in a rather different manner than Moore. As mentioned above, she applies the sound-room as a three-dimensional equivalent to *sound image* or *sound*. Her analyses based in the sound-room model reveal further differences. In her master's thesis Danielsen introduces the sound-room to grasp what she calls the "fragmented soundscape" (my translation; Danielsen 1993: 51). She explains that sound is a term used to gather a work into a sonic unity, whereas the sound-room is an attempt to differentiate, through focusing on the processes within a sound (ibid.: 52). Thus, as Danielsen explains, the sound-room is an analytical tool that allows for envisaging different processes within a production, through revealing change and any incoherence across time and space (ibid.). This description, in addition to revealing a specific aim for the sound-room, suggests that Danielsen has a more technical approach than Moore, as she seeks to focus on processes within production.

Danielsen's use of sound-room also corresponds with her focus on so-called fragmented soundscapes. She applies this term to describe Prince's then recent music, which consisted of, according to Danielsen, a chaotic and collage-influenced soundscape as a background for a clearer melodic foreground (ibid.: 49). To describe the lack of internal coherence in the song "Push", Danielsen cites Stan Hawkins:

> As this song progresses various musical references enter constantly modifying the stylistic flavour: long, sustained brass chords, orchestral string phrases, scratch samples, percussion sounds are a few of the devices employed. Gradually the song incorporates more rap and hiphop ideas as it leads into a free, jam-like session climax (Hawkins cited in ibid.: 50).

According to Danielsen, the different parts of this collage do not fit to each other, being located in separated parts of the soundscape (ibid.: 50). It is in this context the sound-room becomes useful to Danielsen. The possibilities due to technological development – multi-track recording, digital effects processing and digital sampling – have allowed producers to create such fragmented soundscapes within a song (ibid.: 53). The use of sound-room to analyze the outcomes of technological developments are explored further in Brøvig-Hanssen and Danielsen's "The Naturalised and the Surreal" (2013), which I will discuss below.

As previously mentioned, the appropriation of "internal" or "local" sound-rooms (or sound-boxes) within the global sound-room of a song, can be useful for example when discussing songs in which a highly reverberant sound appears at the fore of the sound-room. In their study of surreal sound-rooms and the naturalization of these, Brøvig-Hanssen and

Danielsen touch upon both the case of internal sound-rooms and drastic changes to the sound-room during a song. They also describe the relationship between sounds in real spaces and sounds in a virtual space (a song), drawing on Smalley's theory of source bonding and Gibson's theory of ecological perception: "we might conclude that in order to make sense of the virtual space projected by a given popular music sound, we unconsciously compare it to previous experiences with *actual* spatial environments, such as a stage or an enclosed room" (Brøvig-Hanssen & Danielsen 2013: 71, emphasis in original). The examples used in Brøvig-Hanssen and Danielsen's study include songs more or less relying on the possibilities of digital audio editing. Starting from three of Smalley's *global spatial styles*, they analyze three songs, each an example of a spatial style. The first example, Prince's "Kiss", incorporates a *single spatial setting*, that is, the song is "set in a single type of space of which the listener is aware at the outset" (Smalley cited in ibid.: 76). This type of global spatial setting can describe a number of songs in which there are no major spatial changes. What is noteworthy of "Kiss" is rather what is done within this one spatial setting. Due to the use of gated digital reverb, "the sound suggests a space that could not possibly conform to the physical laws of sound reflection" (ibid.). The authors explain that "the hyper-presence and lack of depth imply a small space with almost no reverberation, but the high intensity and voluminous sound imply a larger resonant one" (ibid.: 77). Additionally, the vocals and drums are placed in a much smaller proximate space, thus forming an internal sound-room within the global spatial setting (ibid.: 76).

In the Suede song "Filmstar", Brøvig-Hanssen and Danielsen reveal *multiple spatial settings*, meaning that "there is a profound contrast between the verse of the song, which suggests a small, narrow space, and the chorus, which suggests a much larger, broader spatial environment" (ibid.: 77). In other words, there are sudden changes of the sound-room's outer dimensions in the shifts between verses and choruses. According to the authors these changes are probably made possible by "the new possibilities for storing and programming the settings of the digital mixing board", including complete alteration of the settings of the digital reverb used in the mix (ibid.).

The last example analyzed by Brøvig-Hanssen and Danielsen is Portishead's "Half Day Closing", in which they detect what is called *spatial simultaneity*: "Each individual sound occupies a subordinate space within a song's all-encompassing spatial environment, or, as Smalley puts it, the 'holistic' space of the music comprises 'zoned spaces'" (ibid.: 78). The enclosed vocal and drum space within a larger environment in Prince's "Kiss" might also invoke spatial simultaneity, but the Portishead song has a far more complex spatial structure

in this context. According to Brøvig-Hanssen and Danielsen, each sound-source occupies its own different-sounding zoned space. The bass sounds rather dry, invoking a small, absorbent room; a manipulated sound, probably produced by a violin and an electric guitar, sounds like it is played in a long, narrow cylinder, like a tunnel; the voice sounds like it is sung through a megaphone, but in stereo (probably the effect of a Leslie rotary speaker), and it has a rather long reverb; the drums are panned hard left and have a bottled-up quality to them, invoking a small room with hard surface; the electric guitar on the hard right (from the second chorus) has a similar, but warmer-sounding reverb (Brøvig-Hanssen & Danielsen 2013: 78). Additionally, all these sounds seem to be set in a larger "holistic" sound-room through the use of a hardly distinguishable main reverb. Thus, local (zoned) sound-rooms may be said to be situated within the global (holistic) sound-room of this song.

What becomes evident in Brøvig-Hanssen and Danielsen's study, in light of Moore's sound-box analyses, is their treatment of the outer dimensions of the sound-room. Although Moore, in his definition of sound-box in *Rock*, allows for the finite dimensions of the sound-box to change over time (Moore 2001: 121), he does not incorporate this aspect in his analyses. This is also reflected in his illustrations, all of which use the same size for the sound-box. As Brøvig-Hanssen and Danielsen explain, Moore's sound-box is "a music-analytical tool that can be used as a matrix to map the placement of the different elements of a mix" (Brøvig-Hanssen & Danielsen 2013: 72). They, in contrast, vary the dimensions of each sound-room, even within songs, to reflect the perceived overall depth, height and depth of a song. This is apparent in their description of the changing guitar sound in Suede's "Filmstar": "This abrupt transition both reduces the density of the sound and widens the sound box considerably" (ibid.: 77). The authors also highlight the spatial qualities of single sound-sources, whereas Moore focuses on the timbral qualities of such spatially distinct sounds. Again, these different approaches relate to what kinds of music each concept is developed to cover. Moore deals primarily with rock music from the 1960s, 1970s and 1980s, a time when the possibilities for changing parameters within a song were limited, and Brøvig-Hanssen and Danielsen deal with music that is more or less based on such changes.

Moore's and Danielsen's approaches to the spatiality of recorded popular music seem thus to differ, not so much in their definitions, but in their treatment of the dimensions (and the interacting of the dimensions) when applied in analysis. Moore's focus is mostly on relatively static sound-boxes limited by analog technology, whereas Danielsen's analyses of digitally produced music require a different approach. Thus, Moore's sound-box seems to be

particularly suitable for analyzing the global texture of a song, and especially its density. Danielsen's sound-room, on the other hand, seems to shed light on the spatial differences within a song.

Still, most relevant for this thesis are their different understandings of the size of the recorded space. For Moore, this seems to be an invariable measure, which makes sense, since the distance between the two speakers of a stereo playback system also does not vary. But, as Danielsen's approach suggests, a song does not always have to occupy all of the stereo width, and sounds can even be manipulated to appear from beyond the speaker array. In practice, the size of recorded spaces is thus not at all invariant.

2.1.3 Other spatial approaches

Having investigated Danielsen's and Moore's different approaches to the three-dimensionality of recorded sound, I find it appropriate to present in brief some other approaches to a spatial thinking of popular music. One of these is Lelio Camilleri's term *sonic space*, which Moore addresses in *Song Means*. Camilleri argues for three different spaces instead of one. The types of space he addresses are *localized space*, *spectral space*, and *morphological space* (Moore 2012: 37). Moore describes these terms briefly as such:

> "Localized space" is, effectively, the sort of place I have already been describing [sound-box]. "Spectral space" is to do with timbre […] – it is here that we recognize the degree of saturation within a particular part of the soundbox. What Camilleri terms "morphological space" is the sensation of change we experience as timbres subtly alter, a factor also of the register in which different instruments are playing (Camilleri referred to in ibid.).

Unlike Moore, Camilleri defines depth in near and far, or foreground and background, leaving out a sense of middleground (Camilleri 2010: 201). He also applies the term *sonicprint*, to which I will return later in this chapter.

Another approach addressed by Moore is Francis Rumsey's idea of saturation. Of particular interest here is his distinction between distance and depth. Moore explains this: "*distance* is the distance between the 'front' of a sound-source and a listener, while *depth* acknowledges that there can be a 'back' to that sound-source, and the difference between front and back delivers depth" (Moore 2012: 37). Again, the examples of Elvis Presley's "Baby, Let's Play House" and Yes' "Roundabout" comes to mind, as Rumsey's distinction between distance and depth can be useful to describe Elvis' voice, and the balance between Rick Wakeman's organ and Steve Howe's guitar, in the context of a sound-box.

Albin Zak III, in his production-related model, adds a *narrative* dimension, which relates to changing locations within the sound-box. These locations, though, do not move as much as technology allows (Moore 2012: 38). As Moore comments,

> [t]here is perhaps something rather disorienting about sounds moving in ways of unquestionable significance, which presumably frequently cuts across the aesthetic that is driving a particular recording. That is particularly the case if the sound source that is moving is physically static (that of a piano, for instance) (ibid.).

Zak accounts thus for what can be called "aesthetic limitations", which means that some things do not "fit in the mix". This is a question of making a credible mix, an issue I will discuss further in the next chapter.

2.1.4 Disadvantages

Having discussed the sound-room and the sound-box models, as well as some alternative approaches, I will now move on to some of the disadvantages of the concepts. In his master's thesis, Askerøi (2005) addresses the sound-box's problems with representing timbre. He comments that the model describes little about the timbral properties of sound-sources, and nothing about the ways in which timbre is formed through performance practice (Askerøi 2005: 10).[3] According to Askerøi, this makes the sound-box overlook at least two dimensions that are essential to understand sounds (ibid.). Moore partly makes up for this by using pictures of instruments in his sound-box illustrations (Moore 2012; Dockwray & Moore 2010). These illustrations are still not an adequate representation of timbre in my opinion, as for example a specific electric guitar model could sound many different ways, depending on what guitar amplifier and effect pedals are used, and how the parameters on these are set. Thus, Moore mainly treats timbre as separate from the sound-box, while reserving the use of sound-box to describe musical texture.

A better representation of timbre within a spatial model is the one found in Camilleri's sonic space, or more specifically, his description of spectral space. As previously mentioned, the spectral space is where we recognize the amount of saturation, or spectral content, in a sound. Camilleri also takes into account the blurred connection between real physical spaces and timbre: "If localisation is, in a certain way, a real sensation, then spectral space is metaphorical since there is no such physical space, even though we can experience the sense of saturation or emptiness due to the spectral content of the sounds used" (Camilleri

[3] This also applies to Danielsen's sound-room.

2010: 202). Thus, while aware of the problems with incorporating timbre in a spatial model, he still finds a way to implement this with his concept of spectral space.

Another disadvantage of both the sound-box and the sound-room is their difficulty of representing mono recordings. I have already briefly mentioned Doyle's view on the sound-box model. He further argues that Moore implies a

> specificity to stereo and/or multi-tracked recordings. Moore derives his sound-box as a tool for the analysis of progressive rock music; that is, that body of recorded music which had its beginnings, for Moore's purposes, in the Beatles' *Sgt. Pepper's Lonely Hearts Club Band* of 1967. By that time virtually all album tracks and many singles were recorded and issued in stereo, and Moore's construct inevitably privileges stereo spatiality (Doyle 2005: 25).

Such specificity to stereo recordings is also inherent in Danielsen's use of sound-room, as she, like Moore, deals with music that was recorded after stereo established itself as the norm. Camilleri, though, takes also this problem into account by distinguishing between a mono space and a stereo space. Referring to Doyle, he comments that the mono format has "developed its own spatial organisation" (Camilleri 2010: 201). When illustrating stereo and mono spaces, Camilleri defines depth in terms of a foreground and a background. In the stereo space, width is defined in terms of left, center and right (ibid.).

Although I think such a distinction between mono and stereo spaces can be adequate in some cases, especially when analyzing pre-stereo music, it must be considered that most music today is produced for stereo playback. Even if songs have all sound-sources located in the center of the stereo spectrum, there is often a stereo reverb applied to the mix, which in effect makes the centrally placed sound-sources a part of a stereo space. As my concentration in this thesis will fall on stereo recordings, I find Moore's and Danielsen's omission of mono spaces unproblematic for my approach to the spatiality of popular music production. In a historic-analytic context, though, Camilleri's approach could be adequate.

Sound-box and sound-room prove to be rather limited models for analysis. As I have discussed, the models differ in focus, although both deal with the spatial aspects of recorded popular music. This difference results from their being developed for analyses of different types of music. Consideration of spatial aspects in the two models seems also to be somewhat limited. Both timbral aspects and mono recordings are unrepresented. Yet, other spatial approaches find ways to incorporate these aspects, as well as other aspects (like that of narrativity).

Although Danielsen's and Moore's concepts have certain limitations, they shed light on important aspects of popular music, and they can be good bases for a spatial approach to record production, particularly when combined with aspects from the other spatial models discussed above. In the following subchapter I will suggest further opportunities for applying spatial models, such as sound-box and sound-room, to the context of record production.

2.2 Raising awareness of spatiality in record production

Many engineers work in terms of spatiality in some way, in particular when working with effects like echo, reverb, and stereo panning. Producer and mixing engineer Eric Sarafin, also known as "Mixerman", takes such an approach. A substantial part of his book *Zen and the Art of Mixing* (2010) is based on his suggestion of five basic planes of space: panning, frequency, balance, reflectivity, and contrast (dynamics) (Mixerman 2010: 65). Most of Mixerman's mixing philosophy relates thus to spatial thinking. Roey Izhaki also incorporates spatial thinking to some extent, but he does not base his mixing in it in the same way as Mixerman does. In *Mixing Audio: Concepts, Practices and Tools* (2012), a more instructional book than *Zen and the Art of Mixing*, Izhaki takes a spatial approach only when discussing delays, reverbs and panning. Much of this is based in the *sound stage* model, derived from Moylan (2002, revised 2015). Moylan, too, reserves his model mostly for work in the dimensions of depth and width, but uses it in a broader sense than Izhaki. For Moylan, the sound stage model pairs the visual clues of a group playing on stage with what the listener hears, whereas Izhaki adopts the term only in direct relation to the use of different mixing tools (reverbs, delays, panning). Mixerman seems to take the broadest approach to spatiality of the three, applying space in nearly every, if not all aspects of mixing.

What all these approaches have in common is their more or less purely technical approach. All of them are focused on getting a "great mix". They, and Mixerman in particular, discuss how to create mixes that are interesting for the listener, but none of them addresses the listener's experience of the music to any greater extent. In my opinion, there is room for an approach also considering the effect of a production on listeners. This might include, for example, the listener's different associations to the spatial aspects of a mix. I believe producers and engineers also take this into account, more or less consciously, when crafting a mix.

2.2.1 Producing space

As mentioned, there are already some models that take a production approach to the spatiality of recorded sound. Moylan's sound stage is probably the most recognized. Sound stage is a two-dimensional model, in which the listener is placed in front of a perceived stage floor with a perceived width and depth (Moylan 2015: 52). Thus, "[t]he sound stage encompasses the area within which all sound sources are perceived as being located" (ibid.). This makes for a tool for recognizing stereo location, depth, and proximity.

To grasp the diversity of sounds across the width of the sound stage, Moylan applies the term *phantom images*. Phantom images, according to Moylan, "are sound sources that are perceived to be sounding at locations where a physical source does not exist", for example between the two stereo speakers, and even 15º beyond the speaker array (ibid.: 53). The term also explains that sounds can have an illusion of width (ibid.). He further addresses the case of timbral balance, which can be related to the vertical dimension of the recorded space. Timbral balance affects the overall sound of the mix, but can also affect to what extent sounds will blend, an effect that can be desired or undesired (ibid.: 325). What I find particularly interesting is Moylan's conjunction of timbral balance with horizontal location. He comments that this gives the recordist "another level of control over the balance of the mix", meaning that sounds with similar frequency content can be separated from each other, thus adding clarity to the mix (ibid.). Another aspect of this, I would argue, is that the distribution of frequency content across the width dimension should be balanced, at least to some extent, so as to avoid an unpleasant listening experience. That is, unless the effect of unbalance is desired.

Moylan also takes *environmental characteristics* into account (ibid.: 64). He relates this to sonic environments that are suitable for the sound stage, more than associations to real physical spaces. It seems to me that Moylan takes an "internal" approach, in which his concern is what suits the mix. The opposite would be an "external", listener-oriented approach, concerning how sonic environments affect the listener. As pointed out above, I believe producers and engineers, either consciously or not, have some idea of how the virtual spaces they create will affect the listener. Thus, an external approach would, in my opinion, be more adequate for addressing all aspects of how producers and engineers work, as they also take some sort of listener position when working with a production.

One thing I find particularly appealing with the term sound stage, though, is how it makes associations to the on-stage movements of musicians. A common example is a

guitarist moving to the front of the stage when playing a solo. In a recording, this could be simulated by increasing the volume of the guitar in relation to the rest of the mix. Analogies like this one are not as apparent in a sound-box or a sound-room as with a sound stage. Thus, combining the approaches to open up both for the spatial properties of virtual spaces – either natural- or surreal-sounding – and for associations to performative movements and locations, seems relevant to the use of a spatial model as a production tool.

Mixerman also takes a spatial approach to production, but in a more inclusive way than Moylan. As previously mentioned, he partly bases his mixing philosophy in five basic planes of space, corresponding to the three-dimensionality of the sound-box and sound-room models, in addition to the reflectivity plane, which corresponds to the time dimension. A simple description of the balance and reflectivity planes would connect them to relative amplitude and reverberation respectively. In Mixerman's description, though, the difference is that balance relates to the sense of front to back, and reflectivity relates to the sense of far to near (Mixerman 2010: 65). This can thus be linked to the previous discussion of prominence. According to Mixerman, balance is "the holy grail of the five spatial planes", and it "has to do with how loud [an instrument] is in comparison with all the other parts that are playing at a given moment" (ibid.: 85). Thus, balance for Mixerman seems to be related to an instrument's ability to stand out of (or drown in) the mix, whereas reflectivity only relates to distance. But, as he points out, the illusion of reflectivity, and thus distance, is accomplished within all the three spatial dimensions – width, height, and depth – as well as a fourth dimension: time (ibid.: 79). This means that the perception of distance is not necessarily achieved just by adding reverb and/or delay, but can also be affected by the sound's timbral content, its relative amplitude, or its perceived width. For example, the farther away a sound-source is, the more difficult it is to detect exactly where it comes from. In order to create a sense of depth, it could therefore be helpful to let the sound occupy a certain part of the stereo space, thus making what Moylan calls a *spread phantom image* (see explanation in chapter 3). On the other hand, a sound-source that is under a meter away from you will take up more space in width and height than sound-sources that are further away. Thus, all dimensions have some role in creating an illusion of distance. Also, Mixerman's definitions of balance and reflectivity allow an instrument to sound both loud and far away at the same time.

The last of Mixerman's planes of space that I will address here is that of contrast. He relates contrast to the dynamics of a mix, and he comments that it "relates not to what is happening at a given moment, but to how the mix works over the course of time" (ibid.: 76).

Mixerman provides several examples of how the illusion of dynamics can be accomplished by making contrasts in a mix. This is necessary because a pop/rock mix needs to have a dynamic range of about 4 dB or less for all parts of the song to be audible in different situations (for example in a car) (Mixerman 2010: 77). An example of how an illusion of dynamics can be created is to vary the density of the mix between verse and chorus. A variation on this is done in Red Hot Chili Peppers' "Californication" (1999). Whenever vocalist Anthony Kiedis sings, all sounds are panned to the center. When he does not sing, a synth-organ occupies the sides of the stereo image. This tends to break up the closed-ness of the mono space, thus giving a sense of expansion that tends to soften these parts of the song. Such variations of density also relate to Brøvig-Hanssen and Danielsen's discussion of the multiple spatial settings in Suede's "Filmstar", in which the sound alternates between a small, narrow space and a large, broad space (Brøvig-Hanssen & Danielsen 2013: 77).

The different approaches provided here to working in terms of space cover different parts of the production. Mixerman takes the most extensive approach, covering most aspects of mixing. He is mostly concerned with what is best for the song. Moylan takes quite a similar approach, but covers only the dimensions of width and depth, thus leaving height and time (and contrast) outside of the spatial domain. In a way Moylan also has a bit more focus on the listener, as his sound stage model actually puts the listener in front of the speakers. The least listener-oriented approach discussed here is Izhaki's, which mostly provides instructions for creating a good mix.

2.2.2 Cultural meanings of space

Having discussed some ways in which producers and engineers can work in terms of space, I will now elaborate on their intentions in doing so. I have already touched upon this important subject, in addition to the purely technical aspects of creating virtual spaces. As mentioned in the foregoing chapter, Zak adds a narrative dimension to the discussion of spatiality. Although not dealing directly with spatial models, like Moore, Danielsen, and Moylan, the narrative dimension can be useful to understanding the meaning of spatial aspects of music. An example is Zak's comment that the stereo spectrum "can be employed quite literally as a dramatic stage on which a character's movements are choreographed and woven into the dramatic scheme" (Zak 2001: 147). In addition to this movement-oriented view, I think a more abstract approach can be of interest. This includes the effect on the listener of appropriated environmental characteristics – what is associated with certain kinds of space.

Such associations can of course be individual, but there are also lots of associations that have been incorporated in popular culture through repeated use. An example is the use of reverb on a movie character's voice, often in combination with a slightly blurred picture, to signify that he or she is dreaming. Thus, spatial aspects of recording and mixing should be seen in connection with how they affect the listener. To this I will now turn.

The cultural meaning of echo, reverb and the lack thereof

Such culturally appropriated sonic environments are the subject of *Echo & Reverb* (2005), in which Doyle addresses different aspects of the use of depth-related effects in pre-stereo popular music. In the second chapter of the book, Doyle traces the notion of echo as something otherworldly and haunting back to the Greco-Roman myth of Echo and Narcissus:

> In Ovid's account, Echo is a beautiful nymph […]. She is, however, inclined to ceaseless chatter and always seeks to have the last word in any exchange. Zeus […] has Echo detain his wife Hera […] in ceaseless small talk, while he adulterously cavorts with nymphs. Hera comes to suspect Echo's role in the deception and decrees that the nymph may henceforth speak only when spoken to, and then may only repeat the last few syllables uttered to her by others.
>
> Later Echo observes and falls in love with the beautiful and vain Narcissus, and follows his footsteps, longing to address him, but is obliged to wait until he chooses to address her. He eventually calls on her to show herself, but when she does he spurns her. Heartbroken, she hides herself in the recesses of the woods, living in caves and among mountain cliffs. Over time her flesh wastes away, her bones change into rocks until nothing is left of her but her voice. She remains however ever ready to reply to anyone who calls her (Doyle 2005: 40).

Aside from providing a sense of distance to sounds, echo can offer a kind of mysticism, for example when used with several repeats in psychedelic rock music. In the example of Elvis Presley's "Blue Moon of Kentucky", Doyle suggests that the echo on Presley's voice signifies "authority, heroic anti-authoritarianism and the distant (or subterranean)" (ibid.: 185). He continues by claiming that the echo "serves to double the singer's presence, as though to indicate he is being shadowed by another voice" (ibid.: 186). This is also not far from the case in another Presley song, "Mystery Train". While the slapback echo seems to intensify the presence of the voice, Doyle suggests that

> [t]he reverberant voice and bass, the delay on the guitar – all serve to place the music in the distance, to create an effect of relentless approach, reinforced in the second verse, "Train train, coming round the bend," and the third verse, "coming down the line." It is coming, toward us. The singing voice appears to be both *on* the train and *with* us hearing its approach (ibid.: 190).

Thus, the use of echo tends to emphasize the lyrical content, through appropriation of external meaning.

Another example of this is the use of a very short room reverb on some of the songs on the Talking Heads' album *More Songs About Buildings and Food…* (1978). The reverb is applied to most of the instruments, and it provides an illusion that the band is playing in a relatively small, unprepared room. This matches the band's ideal of separating themselves from the larger-sounding bands of the time (Byrne 2012: 39f). Also, the short reverb tends to keep a separation between the sound-sources, underlining the often confusing contents of the lyrics.

This leads me to another aspect of echo and reverberation in music: the lack of it. Tom Lubin claims that "[r]everb brings sounds to life. It gives the impression of power, adds drama, size and depth" (Lubin cited in Doyle 2005: 28). On the other hand, a lack of reverb can provide closeness, intimacy, and thus, a different kind of power. Claiming that reverb brings sounds to life seems thus too simple; an intimate, reverb-less voice can be just as lively. Doyle addresses the case of Les Paul's pioneering microphone techniques in the late 1940s. According to Doyle, Paul placed the microphone only inches away from singer Mary Ford's mouth (ibid.: 149). This was a highly uncommon technique at the time, and it required the singer to sing much softer to avoid overloading the microphone. The result, according to Doyle, is that "Ford's voice suggests a relaxed, breathy intimacy. The listener and the singer are 'placed' now in intimate proximity" (ibid.). By contemporary standards, the proximity of Ford's voice is not as radical as it must have been back then, which is also a result of Paul's tape recording techniques – he doubled Ford's voice, so she in practice sang harmonies with herself. The added harmonies make it more unlikely that Ford is singing right in front of the listener.

A more recent example of the effect of lack of reverb is the vocal sound on Lucinda William's album *Car Wheels on a Gravel Road* (1998). Williams' voice is affected by both the lack of reverb and her soft singing close to the microphone. The voice is also compressed, which adds to the proximity effect, as it brings forth breathing and other "vocal noises". All this adds to an illusion that the voice is unmediated – that there is nothing between Williams and the listener. Thus, Williams' message seems untainted when it reaches the listener. In her music, the lack of reverb and the proximity acts to emphasize the roughness of her voice and her style of singing. It sheds light on her "imperfections" rather than hiding them away. The result is a voice that can be interpreted as honest and authentic. It is not larger *than* life, but large *as* life. Also, this vocal sound suggests the lyrics as the most important part of the song,

as the voice takes so much space. It is not just loud – it is big – thus making it clear that Williams has something to say.

Sonic markers and sonicprints

As a way of analyzing the possible meanings embedded in recorded sound, Askerøi (2013) introduces the term *sonic markers*. Sonic markers are "musical codes that have been historically grounded through a specific context, and that, through their appropriation, serve a range of narrative purposes in recorded music" (Askerøi 2013: 16). The context through which a sonic marker is grounded can be one of for example time, place, politics, or musical style. This context decides the narrative purpose(s) of the sonic marker, and thereby its role in the music it is appropriated in. As part of a spatial model, sonic markers, through their role as compositional tools, can contribute to clarify the relationship of recorded spaces to other spaces. This issue will be discussed at greater length in the next chapter.

Another approach to such "sound signatures" is Camilleri's term *sonicprints*, described as "sonic fingerprints" (Camilleri 2010: 199). Referring to Theodore Gracyk, Camilleri writes that "music in recorded format is *autographic* like paintings and sculptures, in the sense that there is no possibility that an exact replica could be produced" (ibid.: 200). This means that each recording has its own sonic signature, or sonicprint, and that an artist or band trying to recreate the sonicprint of a record only will be able to create their interpretation of it, resulting in a new sonicprint. Camilleri distinguishes between two types of sonicprints: global sonicprints and local sonicprints. Global sonicprints, he explains, represent "a persistent overall sonic image in a piece or a collection of pieces, an album" (ibid.: 210). On the other hand, local sonicprints "are particular sounds, of an instrument or another sound source, which characterise that recording" (ibid.). It is apparent from Camilleri's descriptions that sonicprints deal with only the sound of a recording and not how this sound relates to its sociocultural environment, like sonic markers do.

Opaque and transparent mediation

The inclusion of sonic markers as a parameter thus allows for a spatial model to also apply to musical sounds with socio-cultural mooring. Another possible way of widening the scope of spatial thinking in record production is the distinction between opaquely and transparently mediated sounds, as discussed by Brøvig-Hanssen (2013). I touched upon this subject in the foregoing subchapter, but I would like to add a few more notes on it in the context of production. As previously mentioned, the binary set up by "opaque" and "transparent"

demarcates detected versus ignored mediations. Which sounds belong to each category depends on the listener's experience with sounds.

Mixerman claims that a good mix is one that the listener does not notice: "We don't ever want the listener to notice the mix" (Mixerman 2010: 62). The song should be in focus, and thus the mix should be more or less transparent. Does this mean that an opaque mix is incapable of being a good mix? Again, I think this depends on the norms of the genre, as well as the listener's experience with that genre. It must also be emphasized that Mixerman only writes about the mixing part of a production. Opacity within the recorded space can also be affected in the arrangement or recording stages. Moreover, the citation above is from Mixerman's discussion of prosody – the consistency between lyrics, music and production. He provides several examples of how a mix should emphasize contrasts in a production, so as to create illusions of dynamics. If employed to the extreme, this can be interpreted in the direction of the mix being used to emphasize opaque relationships in a recording. In that case, the mix itself is not necessarily opaque, but the production contains opaque balances that are brought forth by the mix.

I have discussed here how single elements in the recorded space can be transparently or opaquely mediated, but what is also relevant for the use of a spatial model in production is whether the balance between elements in a recorded space appears as transparent or opaque. As previously mentioned, Moore discusses the "unnatural" amplitude relationship between the acoustic guitar and the rest of the band in Yes' "Roundabout". Given today's norms for popular music production, a more adequate example of unnatural balance in the recorded space could be the placement of lead vocals on the extreme left or right side, or vocals that are placed too far back. Such locations of sound-sources would probably appear to listeners as opaque, just as the balance between guitar and band in "Roundabout" might have done in the 1970's. Thus, at least if one is to follow Mixerman's advice, one should be careful with such mix decisions. On the other hand, a perceived unbalance in the recorded space can, in some cases, help to draw attention to certain aspects of the song. An example of this is the Talking Heads song "The Good Thing" from *More Songs About Buildings and Food*... (1978). The lyrics seem rather ambiguous, which makes it difficult to pinpoint any clear message. In the choruses, this is reflected by the location of the lead vocal. David Byrne's voice keeps appearing at different locations across the horizontal plane, making it seem unpredictable. However, this changing of location is done with care, so that it does not appear to the listener as a radical mix choice. In other words, the mix matches Mixerman's idea of

prosody, while at the same time pushing the boundaries for what the listener notices as mix choices.

By adding new dimensions like the ones provided here to a spatial model for production, one can account for how producers and engineers intend the virtual spaces they create to affect the listener. The approaches I have discussed overlap to some degree, but they also offer different frameworks by which the production of popular music sound can be understood. Doyle argues for cultural meanings of certain early popular music sounds, whereas Askerøi's sonic markers tell something about what happens when such sounds are reused in new music. Brøvig-Hanssen's discussion of opaque and transparent mediation describes what happens when these sounds are used extensively over time. This means that an extended spatial model, based on the concepts discussed above, can help us to understand what possible meanings sounds added to (or formed in) a production can have. Accordingly, the recorded space understood as a totality can act as a carrier of meaning on its own.

Over the course of this chapter I have tried to widen the spatial approach to the musicology of record production. The aim of this approach is to raise awareness of spatial thinking in the process of producing, recording, and mixing popular music songs, but also to discuss the meanings that such practices may produce on the part of the listener. As discussed, there are already several models taking a spatial approach to popular music and record production. The suggested approach is unique precisely because of its extended scope and the ways in which it encompasses all aspects from production to reception – both of which I will discuss further in the next chapter. Accordingly, in addition to addressing more traditional aspects directly related to the process of mixing and recording, I seek to investigate the connections between the decisions made by the engineer and/or producer, and the possible effects (and affects) these decisions have on the listener. In other words, the aim of this approach is to "reveal" an engineer's or producer's communicative intentions, whether these are conscious or not, when making certain choices in a production.

This allows for a discussion of the producer or engineer as auteur, as the composition, lyrics and (studio) performance are no longer the only carriers of meaning in a recording. An extended production-analytic model involves describing several aspects of record production, such as the use of microphone techniques, equalizers, compressors, stereo panning, reverbs, delays, etc. to create virtual spaces that fit the song. Importantly, however, it means not only describing these effects as sound, but also their connections to the lyrics, the song's (and band's) socio-cultural environment, and their relation to music history.

3 Making sense of space

3.1 Acoustics and binaural hearing

I will start with a short introduction to acoustics – how different sounds behave in different spaces – and how listeners perceive horizontality through listening with two ears. This is meant as a background for understanding the three-dimensional recorded space, a virtual counterpart to real space. I will briefly describe what makes different sounds behave differently in different types of space, as well as how binaural hearing is necessary for the horizontal localization of sounds.

3.1.1 Acoustics

The greatest distinction between different types of space is that of indoors and outdoors, or enclosed and open, spaces. The most extreme case of open space is called a *free field* and it occurs when "a source of sound is small enough to be considered a point source and is located outdoors away from reflecting objects" (Rossing, Moore, & Wheeler 2002: 525). In a free field the sound pressure is halved when the distance doubles. Free fields are defined by their total lack of reflection, only occurring indoors if the room is reflection free and thus *anechoic* (ibid.: 526).

Outdoors, in natural surroundings, the only possible sound reflections are in the form of echoes. Reverberation only occurs in enclosed spaces: "Indoors, sound travels only short distances before encountering walls and other obstacles. These obstacles reflect and absorb sound in ways that largely determine the acoustic properties of the room" (ibid.). A room's acoustic properties are affected not only by its size and shape, but also of the material of which its surfaces are made, as well as its content.

When a sound travels from its source, what first reaches the listener is the direct sound. Simultaneously, the sound hits various reflective surfaces, from which the reflected sound travels to the listener. These reflections arrive shortly after the direct signal. "The first group of reflections, reaching the listener within about 50 to 80 ms of the direct sound, is often called the *early* sound" (ibid.: 527). Reflections arriving after the early sound are usually reflections of the early reflections arriving at increasingly shorter intervals, resulting in what is called reverberation. Reverb can thus be explained as a series of very short delays.

The level of a reverb decreases, after the original sound stops, at "a more or less constant rate until it reaches inaudibility" (Rossing et al. 2002: 527).

As mentioned, in a free field the level of a sound is halved with every doubling of the distance between the sound-source and the listener. Due to reverberation, this is not the case in enclosed spaces. "In a typical room, sound waves undergo dozens of reflections before they become inaudible. It is not surprising, therefore, that the acoustical properties of the room play an important role in determining the nature of the sound heard by a listener" (ibid.: 523). Indoors, the decay of sounds is determined by the room's acoustics, which also affect how different frequencies in the sounds resonate. For example will bass sounds often have some "frequency build-up" in small rooms (see, for example, ibid.: 568). In larger rooms, the bass resonance will be too low for us to hear it, and thus it will not cause any problems (ibid.: 523). Also, any room dimensions being of the same size increases the possibility of certain frequencies becoming more prominent than others. Rossing et al. show that the distribution of emphasized frequencies for a cube and a rectangle (with dimensions in the ratios 1:2:3) are quite different: "the cube has a very 'peaky' response with many coincident modes, whereas the 1 : 2 : 3 room has a more even spread of resonances" (ibid.: 569).

More relevant to this thesis, however, are the ways in which the size of a room affects reverberation. In a large room that has been treated to provide good acoustics, for example a concert hall, the room size mostly affects the length of a reverb. Larger rooms give longer reverb tails, and vice versa (Meyer 2009: 189). The pre-delay time – the time interval between direct sound and early reflections – might also be longer, depending on the distance between the sound source and the listener. A shorter distance increases the pre-delay. The timbral properties of the reverb are mostly affected by the materials of which the surfaces of the room are made, something that I will discuss below. In smaller rooms, however, there is a limitation to how long a reverb can be, as the shorter distances will restrict the length of reflections. There is thus more emphasis on early reflection, depending on the materials of the room.

The materials of which the different surfaces of a room are made determine how much sound those surfaces reflect. Different types of materials absorb sounds to differing degrees, something that also varies with sound frequency. Some materials absorb low frequency sounds better than high frequency sounds, while other materials absorb high frequency sounds best. The amount of absorption a material is capable of is defined by its *absorption coefficients*. For example, in Rossing et al.'s table of examples of various materials, a wood floor has an absorption coefficient of 0.15 for 125 Hz sounds and 0.07 for

4000 Hz sounds (Rossing et al. 2002: 531). Thus, the material's capacity for absorption is better for low frequencies than high. A carpet on concrete, on the other hand, has an absorption coefficient of 0.02 for 125 Hz sounds and 0.65 for 4000 Hz sounds (ibid.). The carpet dampens the high frequencies, but it is too thin to affect the low frequencies to any extent, while the concrete is too hard to have any noticeable absorbing effect.

The varying absorption coefficients of different materials thus affect the timbral content of a room's reverberation. This also applies to the contents of a room, including furniture, decorations and people. In turn, this explains why different timbral properties are commonly associated with different types of rooms, for example the different associations to the sound of a bathroom, a bedroom and a kitchen. In those cases, the room's lack of acoustical treatment is also determinative.

3.1.2 Binaural hearing and horizontal localization

"The most important benefit we derive from binaural hearing is the sense of localization of the sound source. Although some degree of localization is possible in monaural listening, binaural listening greatly enhances our ability to sense the direction of the sound source" (ibid.: 89). Lord Rayleigh, an early contributor to the understanding of acoustics, performed experiments regarding localization of sounds, finding that low frequency sounds were harder to locate than high frequency sounds (ibid.). He explained further, "a sound coming from one side of the head produces a more intense sound in one ear than in the opposite ear, because the head casts a 'sound shadow' for sounds of high frequency" (Rayleigh referred to in ibid.). Low frequency sounds are not as affected by this shadow effect, as their longer wavelength make them *diffract*, or "bend around", the head (ibid.). Another factor affecting horizontal localization is delay. A sound coming from the right side of the head will arrive at the right ear before the left ear: "the apparent direction of the sound source is the direction from which the first arriving sound comes" (ibid.: 90).

Although this short introduction to acoustics and binaural hearing only barely touches upon a vast field, I hope it can provide a basis for understanding how sound behaves in different spaces and how we perceive this. Furthermore, it will hopefully illuminate the relationship between real spaces and recorded spaces in the following discussions.

3.2 Dimensions

As explained in the previous chapter, the recorded space can be understood as a three-dimensional virtual equivalent to actual spaces. However, time should also be considered a dimension of the recorded space – not necessarily in the sense that time is spatial, but that sound (and thus the recorded space) evolves over time and cannot exist without it. In this subchapter I will discuss the roles of these four dimensions, as well as related parameters, in the context of record production; how they affect the resulting recorded space, and how they can be formed in the production process. Askerøi argues that the three-dimensionality in modern record productions is an audible feature, and that it thus is not based on illusions, like a painting or photography (Askerøi 2005: 1). But, as I will argue, similar to the way in which a painting cannot have "real" depth, a recording cannot have "real" height. This makes the two similarly based on illusions.

3.2.1 Time

Time is probably the most important dimension for music. Sound waves need time to evolve, meaning that without time there is no sound. In spatial models, however, this dimension is easy to forget. This is in particular because one often experiences the music in a bigger perspective than on the level of sound waves. In their spatial models, both Danielsen and Moore consider time an important parameter, but, as I discussed in the previous chapter, Danielsen's use seems to shed some extra light on this aspect. In her analyses, Danielsen uses the time dimension to explain radical changes to the overall sound made possible by digital mixing and sampling technology. Moore, on the other hand, deals with analog recorded music, in which radical changes, like the ones described by Danielsen, were difficult to achieve. A typical example of Moore's use of time is the varying panning of sounds across the dimension of width: "This panning effect moves sound sources as an entity, rather than moving specific sound sources. This results in the mix using the sounds to reveal the lateral extremes of the sound-box, as opposed to expanding the overall width" (Dockwray & Moore 2010: 191). Although Danielsen's and Moore's approaches differ somewhat, both rely on time for describing change within the recorded space.

Time is also vital to the perception of parameters normally associated with other dimensions. I already mentioned Dockwray and Moore's discussion of variable panning, an effect often associated to width. However, time can also be linked to dynamics, as this is related to how relative amplitude changes throughout a mix. In a production, variable

dynamics can be affected by style of playing, the arrangement of the song, or by volume automation, and it is often associated to depth since it affects how prominent a sound is. Another parameter often associated with depth, which is also related to height and width in addition to time, is reverb. As mentioned, a reverb consists of early reflections and a reverb tail. The time that passes between the direct signal and early reflections reaching the listener, as well as the time it takes for the reverb tail to decay, can tell us something about the position and spread of a sound across the depth plane.

Furthermore, time is connected to the tempo of the music, which should not be forgotten in the context of production. A slow tempo with slow rhythms can cause a sense of openness in the recorded space – it will increase the size of "holes" across the time dimension (more on holes in the section on density). In contrast, instruments playing new notes in fast succession can lead to a dense space. Again, the time dimension is closely connected to other dimensions like depth and height.

Another way of considering the perception of time in recorded music is provided by Mads Walther-Hansen (2015), who suggests that "diegesis equals the perceived temporally connected performance" (Walther-Hansen 2015: 37). *Diegesis* is a term derived from film sound theory describing sounds directly a part of the narrative (diegetic sounds) and unconnected, off-screen sounds (non-diegetic). Based on Gérard Genette's classification of narrative events and Claudia Gorbman's classification of film sounds, Walther-Hansen applies the categories of *meta-diegetic* and *extra-diegetic*, referring to sound events appearing to take place before and after the diegetic event (the perceived performance) respectively (ibid.: 35). The experience of meta-diegetic or extra-diegetic events often emerge, writes Walther-Hansen, "as a consequence of both a spatial and temporal hierarchy between sound events" (ibid.). This can include sounds that seem to have been recorded before or after the performance of a song. A meta-diegetic event can be a recording within the recording. Walther-Hansen exemplifies this with Queens of the Stone Age's song "You Think I Ain't Worth a Dollar, But I Feel Like a Millionaire", in which a rock song is played on a car radio (with its sonic characteristics) before the sound quality suddenly morphs into that of a high-resolution recording. There is thus a change from pre-recorded (meta-diegetic) to recorded (diegetic). As an example of an extra-diegetic event, Walther-Hansen discusses the song "Radio #1" by Air, in which a male voice appears quite far into the song. The voice is very loud, pushing the rest of the sound-sources to the back of the recorded space – a result of a ducking compression effect (ibid.: 43). It is obvious that this is an event happening "'outside' the performance universe", thus being extra-diegetic (ibid.). With his notion of diegetic, non-

diegetic, meta-diegetic and extra-diegetic events, Walther-Hansen thus adds to the idea of a narrative space, which I will discuss later.

3.2.2 Depth

Although probably not vital to music in the same sense as time, depth is an important dimension in recorded music. As Moore explains, the depth dimension has to do with the notion that sound-sources are located at different perceived distances from the listener (Moore 2001: 121). This is an effect of such parameters as volume (relative amplitude) and reverberation, and of perceived loudness and proximity, which is affected by the frequency content of sounds.

Moore classifies depth in terms of three categories: foreground, middleground, and background (ibid.: 121). Even though perception of depth in recordings is continuous, it may be difficult to point to differences within these categories. Moylan similarly uses three categories of distance, but with other names: proximity, near, and far (Moylan 2015: 219). He relates these categories to the level of detail in timbral components and changes in source timbre, as well as the difficulty of localizing far sounds (ibid.). Although Moore's and Moylan's categories are essentially representing the same features, they give slightly different associations, connected to their different perspectives (listener's and producer's, respectively). For a producer familiar with the proximity effect of microphones, the proximity category is easily associated with sounds at the extreme fore of the recorded space. However, Moylan allows for a "moderate level of definition of timbral components" in the far end of the proximity category (ibid.). Moore's categories thus seem more open to interpretation, also for listeners unfamiliar with production related terms. What is interesting in this context is that Moylan acknowledges the continuous nature of distance localization by describing how sounds differ within the three categories (ibid.).

But depth is not only defined by a sound's perceived distance from the listener. Doyle argues that the discussion of reverberation as depth-defining can be somewhat complicated due to the fact that highly reverberated sounds can also appear to be close, for example, when they are loud (Doyle 2005: 26). An example of this, addressed by Doyle, is the guitar sound in Link Wray's "Rumble". According to Doyle, in some cases, "the variable of depth seems to be at least partly if not wholly dependent on relative amplitude regardless of the presence or absence of reverb. At other times, […] reverb appears to be the major depth-creating effect" (ibid.).

Although the perceived distance of a sound from the listener is not always defined by the amount of reverb, I believe that the reverb, in such cases as "Rumble", defines the perceived depth of that single sound, i.e. its perceived size across the depth plane. In his discussion of saturation, Rumsey draws a distinction between distance and depth, explaining that "[d]istance is a term specifically related to how far away an individual source appears to be, whereas depth can describe the overall front-back distance of a scene and the sense of perspective created" (Rumsey 2001: 35). Moore, as mentioned above, gives a similar description (Moore 2012: 37). Based on this distinction, I will treat the distance and depth of single sounds as two different categories, both affecting the perception of depth in the recorded space.

Distance

As mentioned, the perceived depth of local sounds is mostly affected by reverberation, and more specifically, the length of the reverb tail. The time interval between the direct signal and the early reflections of the reverb, on the other hand, contributes to establishing the distance between the sound-source and the listener. As explained above, if a sound-source is close to the listener the sound reaches the listener shortly after it is made. At the same time this sound travels to reflective surfaces, gets reflected, and then these reflections are perceived by the listener. The ratio of distances between the listener, the sound source and the closest reflective surface thus defines the length of the pre-delay time the listener hears. In a recorded space, however, one must also take into account the perceived size of the space in which the sound is located. If the reverb tail is very short, the pre-delay time is not very effective, since small rooms don't afford much diversity in depth. If the reverb tail is long, on the other hand, the pre-delay time can be much more affective.

The proximity effect, as the term suggests, also plays a role in the perceived distance of sounds. Moylan describes proximity as "the space that immediately surrounds the listener, and is the listener's own personal space. Sounds in proximity are perceived as being close, as occurring within the area the listener occupies" (Moylan 2015: 218). According to Moylan, the perception of proximity is related to the level of detail in "timbral components" and definition of sound quality (ibid.: 218f). This means that highly saturated sounds with a lack of definition will appear as less proximate than more timbrally detailed sounds. Also an increase in bass response can add to the effect of proximity. The proximity effect is often mentioned in relation to microphones and microphone techniques. Directional microphones, like cardioid pattern microphones, have a tendency to increase bass frequency when they are

placed very near the sound-source (the sound-source's proximity). A typical example of the proximity effect is a vocalist singing softly, close to the microphone, and having high relative amplitude. The close placing of the microphone will pick up any details of the voice, as well as boosting the low end. The situation is thus not unlike when someone speaks softly very near you.

3.2.3 Width

Width is the only one of the dimensions discussed here that is not applicable to mono sound, as mono cannot be said to have any width. The width dimension is a result of the use of two speakers placed along a horizontal line (or speakers placed behind the listener in the case of quadrophonic and surround systems etc.). The use of width in both Danielsen's and Moore's models reflects that they have been developed with stereo recordings in mind. Moore comments on this that the sound-box "requires the development of stereo for its full manifestation" (Moore 2012: 31). This leaves the models more or less incapable of analyzing spatiality pre-stereo music (because they are based in the idea of three spatial dimensions).

In *The Audio Dictionary* (2005), White and Louie describe "stereophonic" as referring to "a sound system that provides the listener with an illusion of directional realism" (White & Louie 2005: 374). Although width in stereo playback in some ways can be regarded as "real", since it is partly based on physical properties, it is still only a *representation* of real events that took place in the recording studio. Stereo recordings are played back through two speakers, sending out different sounds that reach each of the listener's ears at different times. This makes the listener able to locate where the sounds come from on the horizontal plane, just like we can hear where "real world" sounds come from. However, as Théberge, Devine and Everett suggest, the "point of audition" in a recording is not necessarily realistic (Théberge et al. 2015: 7). This is mentioned in a discussion of early "auro-stereoscopic"[4] telephonic transmissions from the Paris Opera and Théâtre Français, where the location of the microphones at the lip of the stage suggests a "'point of audition' that would have been highly unusual for most audiences of the day" (ibid.). Talking of realism in modern popular music is even more difficult, as instruments are often recorded separately, track by track. I will return to a further discussion of the realism of recorded spaces later in this chapter.

[4] "Auro-stereoscopic" is a term used by Clément Adler to describe his arrangement of telephones for opera and theatre transmissions (Théberge et al. 2015: 7). The term refers to the *stereoscope*, a device for viewing "three-dimensional" pictures, based on the difference between what we see with each eye (ibid.: 6).

The opportunities of stereo recording and playback have been taken advantage of in several different ways. Dockwray and Moore (2010) study the development of a normative mix configuration, which has mostly to do with the placement of sound-sources across the horizontal plane. I will come back to this in a later discussion. I want, however, to mention, in relation to this short discussion of directional realism in stereo sound, that before a normative mix configuration was established, several different variations were common that probably would be regarded as peculiar today.

Width in a recording is mostly affected by panning and stereo reverbs, often in combination with stereo microphone techniques. An instrument can be recorded in mono or in stereo. This means that width is determined both by the perceived width of single sound-sources and by the horizontal placement of sounds. Rumsey writes that what he calls *apparent source width* "relates to the issue of how large a space a source appears to occupy from a sonic point of view" (Rumsey 2001: 36). In relation to his sound stage model Moylan (2015) similarly addresses the perceived width of sounds. Moylan, as mentioned, uses the term phantom images to describe sound-sources that seem to emerge from where there is no physical source. Phantom images are related to our perception of horizontal location through binaural hearing: a sound that reaches us with the same energy in both ears, at exactly the same time, will appear to come from the horizontal center, even if there is no physical sound-source (for example a speaker) there. Also, phantom images provide the illusion of the width of sound sources. Moylan suggests two types of phantom image: *spread image* and *point source* (Moylan 2015: 53). Spread image phantom images are sounds that have a certain perceived width, extending between two audible boundaries (ibid.: 54). The size of spread images can vary considerably (ibid.). Point source phantom images, on the other hand, are sounds than can be pinpointed to a precise point in the stereo field, thus appearing narrower than spread image sources (ibid.: 53.).

In Moylan's examples, spread images seem to be a result of stereo microphone techniques or use of stereo reverbs. I would argue, however, that a spread image could also be a mono source that sounds wide because of its low register content. Low frequency sounds are, as previously mentioned, more difficult to locate than high frequency sounds, and are thereby perceived as more spread than high frequency sounds. There is thus a distinction between the horizontal location and the perceived scope of sound-sources across the width dimension, in much the same manner as the difference between distance location and the depth scope of single sounds. The width of sounds also has high relevance to what Moore calls the *density* of the sound-box, which I will discuss below.

3.2.4 Height

In recorded music, height is defined in terms of frequency content. Moore uses the term register, whereas Danielsen uses frequencies and pitch. In the continuation of Danielsen's description I would argue that height should also be determined by timbral content, as the "darkness" or "brightness", which is determined by frequency content, seems to affect how high sounds appear to be. This might be more relevant for engineers than musicians, as different frequency areas are often defined based on their timbral character, like "boomy", "boxy", "muddy", "honky", "air", etc. (see, for example, Izhaki 2012: 211). The inclusion of timbral content to the notion of height means that it can be affected not only by the register in which the instrument is played, but also by microphone choice and placement, as well as equalizing.

Many don't regard height as a perceived dimension in sound. This is not surprising, as there is no actual height in music – nothing is actually moving up or down. Moylan, for example, keeps his sound stage model two-dimensional, explaining that the pairing of height and pitch/frequency/timbre "is not an element of the actual spatial locations and relationships of sounds" (Moylan 2012: 167). He also claims that "[t]he perceived elevation of a sound source is not consistently reproducible in widely used playback systems and has not yet become a resource for artistic expression (Moylan 2015: 51). This is difficult to disagree with. As Moore comments, "[t]his is true: musicologically, this dimension is not one of elevation of the sound, but is determined initially by its position in pitch space; how high or low its pitch is" (Moore 2012: 31).

In a model that is based on a reference to live stages, the only actual spatial locations are in the width and depth dimensions. In a virtual space, as conceptualized in the sound-box and the sound-room, the exclusion of height based on the above argument would also make the depth dimension irrelevant, since that is also not an element of actual spatial location. Sounds only *seem* to appear at different distances (in fact, they are coming from the same set of speakers, which have constant positions). But sense of height might be even more vague, especially when we define it as pitch and frequency. I believe thus that height, as a dimension in recorded music, must be regarded as culturally determined. We often describe melodies as going up or down, and tones as high or low, as well as the high and low sound frequencies these tones and melodies consist of. Also, in musical notation high and low notes represent high and low tones respectively. On mixing consoles the equalizers often (but not always) have high frequency bands at the top and low frequency bands at the bottom. Such things

might explain why height in music often feels natural to musicians and others working with music.

Returning to the claim I made above, that the three-dimensionality in images and recorded music are equally based on illusions, I would argue that audible perception of height in sound *is* possible to simulate in the stereo format. It is for example not uncommon in sound design for video games to do this, based on what is called *head-related transfer functions* (HRTF).[5] This is, however, dependent on the use of headphones for playback, as listening on loudspeakers will not have the same effect. It is also rather unusual to deal with this in record production. Moreover, the descriptions of height in music given by Danielsen and Moore connect it to register, i.e. frequency and pitch. A different explanation for why we experience height in recorded music, than the actual audible reproduction of height, is thus needed.

In his doctoral dissertation, Hans T. Zeiner-Henriksen (2010b) addresses the case of verticality in club music and its relation to body movements. He explains the relation referring to George Lakoff and Mark Johnson's term *primary metaphors*, which he describes as

> metaphors that have been incorporated into our worldview so thoroughly that we no longer see them as metaphors. They are based on correlations between expressions and embodied experiences and are, according to Lakoff and Johnson, fundamental to all thinking regarding subjective experience and judgement (Zeiner-Henriksen 2010b: 134).

The height dimension of recorded spaces is thus not a "real" dimension, like width, which is based on the physical directions of sounds. Instead, it is based on our many experiences with high-pitched sounds coming from above our ears (e.g. birds) and low-pitched sounds coming from below our ears (low-pitched sounds are more likely to resonate with floors or even our bodies than high-pitched sounds). This is, as suggested, reflected in different musical terms.

3.2.5 Density

Moore explains that "[t]he most important features of the use of [sound-box] are the types and degrees of density filling it (whether thin strands or 'blocks'), and the presence in this space of 'holes', that is, potential areas left unused" (Moore 2001: 121). Although not a

[5] An HRTF is, according to Rossing et al., "essentially a filter through which any sound can be processed", based on recordings made inside the ear canals of dummy heads (Rossing et al. 2002: 690). The filter is used to simulate pinna (outer ear) cues, the shadowing effect of the head, etc. (ibid.: 689).

spatial dimension per se, density, then, must be seen as an important parameter when considering spatiality in a recording.

"Holes" are perceived open spaces within the sound-box, or recorded space, such as gaps between low-frequency and high-frequency sound sources, or gaps in the stereo placement. Although holes can occur in any of the dimensions, they are probably most significant in the width dimension. On the vertical plane they can be hard to notice, except perhaps when bass frequencies are lacking. On the depth plane they are most noticeable if there is nothing in the foreground. Density is nevertheless something that applies to all the dimensions to the same degree, since clustering of sound along any of the axes can make the recording sound cluttered. Spreading the sounds, in height, depth, width, as well as time, may on the other hand result in a more "open"-sounding space. Although density applies to all dimensions separately, Moore's definition seems to focus on what happens in the sum of the dimensions. Thus, the density of each dimension affects the overall density of the recorded space.

The global density of a recorded space seems thus to be affected by the properties of each dimension in two ways: the location and the size of sounds. Similar distinctions can, as mentioned, be used in the dimensions of both width and depth, where there is a difference between horizontal location and perceived width, and between distance and depth. In the height dimension the size of a sound's register, or its frequency range, is relevant. All these factors, by affecting overall location and individual sound size, contribute to defining the density of the recorded space.

The recorded space consists thus of four basic dimensions: depth, width, height and time. All of these, and their related parameters, such as distance, dynamics and density, play an important role in the forming of stereo recordings. In much the same way as dimensions in a real room cannot exist separately, it is difficult to separate the dimensions of the recorded space from each other. Balance across the stereo plane is, for example, not achieved by having one instrument on each side, if one instrument is located high in the recorded space and the other one is low, or if one is in the foreground and the other in the background.

3.2 Development of a normative mix

Having discussed the outlines of a spatial approach to production and its different dimensions, I will now turn to a more historically oriented account of how the stereo mix evolved into a normative mix with clear parallels to live performances. Based on Moore's

sound-box, and its potential for mapping sounds within a virtual space, Dockwray and Moore (2010) investigate the development of horizontal distribution of sounds in early stereophonic popular music recordings to what is now a normative mix configuration. This development is discussed in light of technological development, as well as aesthetic aspects. The authors explain that their aim was "to enable the description of a historicised achievement of [the diagonal] mix from a range of other possibilities which existed during the early years of stereo, i.e. prior to the early 1970s" (Dockwray & Moore 2010: 183).

3.2.6 Functional staging of sound

In the late 1960s, when stereo playback equipment was still not something that everybody had,[6] engineers avoided placing sounds at the sides of a mix. This was done in order to prevent center build-up of sounds when stereo recordings were played back on mono equipment. The summing of the two channels made the centrally placed sounds louder in mono than in stereo, thus changing the mix balance (ibid.: 187). This effect of mono summing could also cause a centrally placed bass to make the needle on the turntable jump out of its groove. The solution to that was to place the bass on one of the sides, and balance with other instruments on the opposite side (ibid.). However, the later central placement of the bass was, according to Albin Zak, "a result of its sonic function 'as anchor of both groove and chord changes'" (Zak cited in ibid.: 188).

The technical and, maybe to a larger extent, the aesthetic reasons for the different mix configurations can be related to Zagorski-Thomas' term *functional staging*. The term refers to a functional adaptation, or staging, of productions for different purposes (Zagorski-Thomas 2010: 252). Zagorski-Thomas discusses how 1970s disco and rock was staged differently to function in different contexts. Disco records were played back in clubs – large, reverberant spaces that in themselves provided a lot of reverb (ibid.: 253). For the club-goers to hear the kick drum clearly, and thus be able to feel the rhythm, short, dry sounds were used. The kick drum, and often also the bass guitar, was muffled in order to give short sounds. Rock records, in contrast, were usually played back in living rooms, which are rather dry spaces (ibid.: 255). To simulate the live-experience familiar from large stages, rock productions had big sounds, with extensive use of reverberation.

[6] "During most of the mid to late 1960s, stereo mixes of albums were considered to be a minor adjunct to the dominant mono version" (Dockwray & Moore 2010: 187).

In Zagorski-Thomas' examples, the production aesthetics were formed with the playback environment in mind. In the case of early stereo mixes, on the other hand, aesthetics were adapted to the format and playback equipment the records would be played back with. It is thus not so much a case of environment as one of technical limitations. These adaptations have further affected how we recognize production styles from the period, much in the same way as 1970s disco is known for its "thumpy" drum sound, which is a result of the challenges of the music's common playback environment. We can thus differentiate between at least two different types of function for which the music can be staged – one deals with environment and social function, whereas the other deals with the purely technical issues of playback. The subsequent placement of bass guitar, and eventually also snare and kick drums, in the horizontal center can be related to both Zak's suggestion that these instruments are central to groove and chord changes, as well as the emerging norm of listening through headphones, where bass sounds (bass guitar and kick drum) placed on a side can make the recorded space appear as out of balance.

3.2.7 Mix types

In their analyzed material, which consists of sound-box transcriptions of 1000 songs, Dockwray and Moore identify four main mix types: *clustered*, *triangular*, *diagonal*, and *dynamic*. This taxonomy of mix types can tell us something about how the understanding of stereo recording and mixing evolved in the latter part of the 1960's. In what follows I will give brief descriptions of the four mix types.

Clustered and triangular mixes

As the term suggests, a clustered mix is a mix where many sounds are stacked in the horizontal center of the recorded space, resulting in a narrow stereo image. There are variations of this where the cluster seems to be placed towards one of the sides, and some in which a single instrument is detached from the main cluster (Dockwray & Moore 2010: 186). This means that the term can be used to describe what is essentially a mono mix, but also stereo mixes consisting a thick block of sound as well as big holes.

The triangular mix is, in contrast, a more spread configuration. The term refers to the triangular configuration between lead vocals, snare drum and bass guitar. Although a centrally placed vocal with snare and/or bass on the side(s) is more common, the term describes any situation where one or two of these elements are off-center (ibid.). This might

be related to early "duophonic"[7] records – 1-track mono records that were reprocessed for the stereo format. According to Barry (2015), RCA-Victor, after having success with duophonic classical records, released Elvis Presley's album *Elvis' Golden Records* in stereo. The reprocessing, which involved the splitting of frequency content between the two channels – low frequencies to one side and high frequencies to the other – as well as adding reverb, "put boom in one channel and scratch in the other", resulting in a "highly unstable" version (Barry 2015: 135). What is interesting here is the similarity in frequency distribution across the stereo field between such "fake stereo" records and true stereo triangular mixes with off-center bass.

Within the category of triangular mixes there are some variations that seem to have changed somewhat through time, perhaps as part of the "evolution" of the diagonal mix. A distinction worth mentioning between different triangular mixes is that of bass guitar location. Some mixes have the bass placed to one of the sides, together with snare, opposite of the snare or with the snare in the middle, whereas other mixes have the bass centralized while the snare is panned. The central placement of bass in triangular mixes occurs, according to Dockwray and Moore, mainly between 1967 and 1969 (Dockwray & Moore 2010: 187). Increased listening with headphones can clarify what importance this change had for the perceived balance of the recorded space. Furthermore, the mentioned point made by Zak, of the bass' central role in groove and harmony, becomes evident.

As Moore (2001) suggests, mixes like these are not necessarily originally intended for stereo listening. Mastering to two or three tracks was quite normal, even for mono recordings. But when mono recordings were re-mastered for the CD format in the late 1980s and throughout the 1990s, they were often released as stereo recordings, resulting in what is called "fake stereo" (Moore 2001: 122). Of Amen Corner's "Bend Me Shape Me", Moore writes that "[t]he original mono release was from a two-track master: the voice and rhythm instruments appear on one track, the voice, hand-claps, organ and horns on the other" (ibid.). With the voice on both tracks, it will sound central, whereas the other instruments are panned hard to each side (ibid.). The result of such fake stereo masters is that the recording can appear as heavy on one side, often more than in triangular mixes originally intended for stereo playback.

[7] A term coined by Capitol Records (Barry 2015: 132).

Diagonal mixes

What Dockwray and Moore call the diagonal mix configuration is what has over time been established as the normative distribution of instruments across the horizontal plane. This norm is still applicable in many genres, in particular rock related styles. The authors describe the mix type as identifying "the vocals, bass and snare as being on a slight diagonal line in a linear configuration (relative to the vertical axis), with other instruments placed to either side" (Dockwray & Moore 2010: 186). This also allows for the inclusion of mixes where the vocals and bass "appear to be in a perpendicular configuration" (ibid.).

This diagonality can for some be difficult to grasp. As with other aspects of listening, the noticing of any diagonality can be dependent on one's approach and background. With my background in production and engineering, where lead vocals and bass are usually not panned, I find it hard to hear the diagonality in Dockwray and Moore's examples, as well as in other normative mixes. The exceptions to this are cases where the vocals are recorded with stereo microphones, something that adds some dynamics to the stereo placement. Moore, on the other hand, points out that he, in most cases, hears the bass and vocals as panned slightly to each side (A. F. Moore, personal communication, February 5, 2016). In the cases of early stereo recordings, where the whole band was recorded in the same room, the stereo location of sound-sources might be affected by the placement of instruments in relation to stereo drum overhead microphones as much as of using a panning knob. My difficulty hearing Moore's diagonality between bass, snare and vocals also in more recent music emphasizes some of the complexities of the auditory analysis of sound related to auditory perception. Despite my lacking recognition of the diagonality, I will continue to use the term to refer to mixes where bass, snare, and vocals are centrally placed above each other (with bass at the bottom and vocals on top).

A diagonal mix will give a good balance in the recorded space, especially since bass sounds (kick drum and bass guitar) are more or less centralized, thus keeping the "heavy" sounds away from the sides. It also emphasizes the important roles of the three key sounds – vocal as mediator of meaning (lyrics), and snare and bass as important to the groove of the song. What's more, the central placement of important sounds and decentralizing of less important sounds can also be seen as very compatible with mono, as the hierarchy of importance is easily retained.

Dynamic mixes

"The dynamic mix refers to tracks where there is some level of movement within the sound-box" (Dockwray & Moore 2010: 186). A dynamic mix is thus a mix where certain parameters are varied throughout the song. In Dockwray and Moore's article, and in the music from the period they investigate, the parameters are mostly stereo panning and volume, i.e. a variation in width and depth. Variation of the amount of reverb and echo might also have been used.

This mix type seems thus to be one that appears in combination with other mix types, and it might accordingly be regarded as a mix characteristic rather than a separate mix type. In some examples discussed by Dockwray and Moore we see this combination. Jimi Hendrix's "Spanish Castle Magic" features a diagonal mix, but during solos the panning of the electric guitar is varied across the stereo channels. "Purple Haze", on the other hand, can be described as a clustered mix, where a cluster of drums, bass and electric guitar appears in the center, and vocal is on the far right. This placement of the vocals makes a triangular configuration between bass, snare, and vocals, thus making the mix fit the characteristics of both triangular and clustered mixes. On the left side of the sound-box there is a backing vocal appearing occasionally, as well as a "tinkling" sound appearing during the last guitar solo. This sound is eventually, at the end of the song, panned around in the stereo field.

Starting from this definition of the dynamic mix as a characteristic that can be applied to any other mix configuration, I would claim that in a contemporary context the term could be used to describe any automated mix, as well as mixes where dynamic changes are added "live" during mixdown. In contemporary music, this also includes automation of equalizers and filters. Dynamics, in this sense, can thus happen across any of the three spatial dimensions.

In the early stages of popular music stereo recording, there were evident changes from experimenting with different mix configurations to a normative mix type. This makes it, to some degree, possible to pinpoint the period in which early stereo popular music recordings were made, based on their mix configurations. There are necessarily some exceptions, and, as I will discuss later in this thesis, such exceptions can be contemporary records where a triangular configuration is used as a retro-element. A mix configuration can thus act as a marker of time, something I will return to in a later discussion of sonic markers.

The longevity of the diagonal mix as a norm for stereo recording since the early 1970s indicates its compatibility with the stereo format. I believe that as long as recordings with

traditional rock instrumentation – with bass guitar, acoustic drums and vocals at its base – are made, the diagonal mix will remain the norm. Although electronic pop music to a large extent relates to the same configuration, with at least bass sounds and lead vocals centralized, the mix configuration can be somewhat more blurred in such styles than in more traditional arrangements.

3.3 Transtextual relations in the recorded space

In "Intertextuality and Hypertextuality in Recorded Popular Music", Serge Lacasse (2007) applies literary critic Gérard Genette's terminology to propose a comprehensive typology of recorded popular music. Genette uses the term *transtextuality* for referring to "all that sets the text in a relationship, whether obvious or concealed, with other texts" (Genette 1997: 1). He divides the term into five subcategories: *intertextuality*, *paratext*, *metatextuality*, *architextuality*, and *hypertextuality*.

In Genette's terms intertextuality means "a relationship of copresence between two texts or among several texts: that is to say, eidetically and typically as the actual presence of one text within another" (ibid.: 1f). Intertextuality can, according to Genette, happen in the form of quoting, plagiarism or allusion (ibid.: 2). Metatextuality describes a commentary relationship uniting a given text to another, "of which it speaks without necessarily citing it (without summoning it), in fact sometimes even without naming it" (ibid.: 4). Hypertextuality is a "relationship that unites a text B […] with an earlier text A […], upon which it is grafted in a manner that is not that of commentary" (ibid.: 5). The categories of paratext and architextuality are more exclusively concerned with literature. A paratext can be a secondary signal, for example, "a title, a subtitle, intertitles; prefaces, postfaces, notices, forewords, etc.; marginal, infrapaginal, terminal notes; epigraphs; illustrations; blurbs, book covers, dust jackets" (ibid.: 3). Lastly, architextuality refers to the "architecture" of a text. It is usually only a paratextual mention, "which can be titular (as in *Poems*, *Essays*, The Romance of the Rose, etc.) or most often subtitular (as when the indication *A Novel*, or *A Story*, or *Poems* is appended to the title on the cover)" (ibid.: 4).

Transtextuality, as the collective term described by Genette, is thus not unlike what is often meant by intertextuality. Michael Riffaterre, for example, writes, "intertextuality is … the mechanism specific to literary reading. It alone, in fact, produces significance, while linear reading, common to literary and non-literary texts, produces only meaning" (Riffaterre

cited in Genette 1997: 2). In other words, Genette's use of intertextuality departs from the more common use, as he means by it one specific type of relation between two texts (ibid.).

Of particular interest to this thesis are Genette's subcategories of intertextuality, hypertextuality and metatextuality. These types of transtextuality describe a direct relationship between the current text and a specific previous text. Such connections occur on regular basis in music. Although the differences between intertextuality, metatextuality and hypertextuality would be interesting to discuss in this context, as they are relevant to the study of spatiality in recorded sound, I will use the term transtextuality as a universal term to describe any of these subcategories as well as more vague connections (for example appropriations of musical styles or extra-musical meanings).

What is a text?

In this context it is appropriate to ask what is meant by a text. In the introduction to *Critical Readings in Popular Musicology* (2007), Moore describes "text" as anything that can be read or interpreted (Moore 2007: xi). This expands the definition beyond the realm of literature. Moore further suggests that,

> [t]his also implies a degree of involvement with the experiencing of a musical text which, also, may not be part of the explicit experience of many listeners, but is normally part of the implicit experience […] The interpretation does not need to be conscious, nor does it need to be involved, but is something that we are inevitably engaged in (ibid.).

Thus "text" can be regarded as a collective term for anything that is subject to interpretation. In music this can include the lyrics of a song, but the term also extends to musical parameters, for example, sound. When discussing music as a transtextual text it can mean that certain musical parameters, for example melody, sound or rhythm, refer to other texts. The texts to which musical parameters refer can, as mentioned, be anything, from specific sonic or structural aspects of other songs to an emotion. In what follows I will discuss the different ways in which the recorded space can act as a reference to the different notions of real and unreal spaces.

Actual space as text

Based on the notion of text as discussed above, spatiality, and more specifically, the sound of spatiality, can be seen as a subject of interpretation. The sound of spatiality signifies something about the spatial location of any sound-source as well as the size, shape and

material of the space. Moreover, when being in and hearing the acoustics of an unfamiliar space, one compares the sound of this space to the sound of previously experienced spaces. Gibson's theory of ecological perception suggests that we approach and understand new environments according to our previous experiences with similar environments (see, for example, Clarke 2010: 106). This results in an interpretation of new spaces with reference to one's experience with other spaces. Also, experiences of certain spaces can be connected to different meanings, for example in the form of emotions, and such meanings can be important in the interpreting of new spaces.

This focus on our sonic surroundings is related to R. Murray Schafer's notion of the *soundscape* (1977, revised 1994). With this term Schafer treats the world as "a microcosmic musical composition" (Schafer 1994: 5). He describes the soundscape as "any acoustic field of study. […] We can isolate an acoustic environment as a field of study just as we can study the characteristics of a given landscape" (ibid.: 7). Although Schafer's definition of soundscape involves the sounds that comprise our spatial surroundings, I would argue that it also allows for a study of the environment of the sounds.

Spatial environments could then be considered on par with sound-sources as a case of study. However, in contrast to sound-sources, spatial environments are not able to produce sound, but have rather a mediating effect. The mediation of sounds is also an important factor in how sounds are perceived. We can thus distinguish two main categories of contribution to the soundscape: sound-sources and mediators. I consider the sound-source to be what actually produces the sound, while a mediator is that within which the sound propagates, and thus is formed by.[8] The acoustic environment of a sound takes partly the same role as a recording or playback medium, in that it forms the timbral content of sounds. It can thus be compared to the different signal modifying processes involved in the electronic (re)production of sound. It also takes part in the forming of the sound's context and thus the sound's meaning(s).

The notion of soundscape directs our focus towards the sounding environment and invites interpretation. Furthermore, this reminds us that sounds do not exist in a vacuum without any external influences. Even in what I would call a sonic vacuum, or completely "dry" spaces (for example in an anechoic chamber), sounds will have certain associations to them, and are thus not "free" from external meanings.

[8] Sound-sources can further be divided into a *driving system* (for example a guitar string) and a *driven system* (the resonating guitar body; Hartmann 2013: 23). This suggests that there is a mediation happening also within the sound-source.

Recorded space as text

The idea that sounds do not exist in a cultural vacuum is especially relevant in the context of the recorded space. Even when sounds are recorded separately, they are eventually combined in the common global space of the recording. Furthermore, the features of each local space, as well as the global space, of a recording are related to the realm of actual space, through their similarities to or differences from it. However, the recorded space cannot be said to be the same as the real spaces of the recording. Even when recording all sounds in the same room with just a stereo microphone pair (for example with a Sennheiser dummy head, which provides a high degree of realism) the resulting sound will be formed by multiple factors. Such factors can be the producer's perspective (for example in the form of microphone placement), signal path (preamplifiers, cables, microphones, etc.), or the playback equipment and environment (type of loudspeakers and how they are placed in the room). All this makes the recorded space a result of the producer and engineer's personal approach and aesthetics, as well as the producer and engineer's associations to specific sounds and sonic environments. Also, the listener's associations, whether they are the same as the producer's and engineer's or not, influence the perceived meanings of the recorded space. The recorded space is a result of microphone techniques, the processing of sounds and use of effects, for example reverb and echo. The factors discussed in the previous subchapter – panning and volume – are especially important. All this constitutes a global spatiality, in addition to individual internal spatialities within this totality.

Record listening is a type of what composer Pierre Schaeffer calls *acousmatic listening*, or reduced listening, in that it is listening without a visible source (Schaeffer 2004: 77). Referring to the *Larousse* definition of acousmatic,[9] Schaeffer describes radio and methods of reproduction as a curtain visually separating the listener from the sound-source (ibid.: 76f). Modern listeners thus find themselves in conditions similar to those of the disciples of Pythagoras. When listening to a recording the playback source (or medium) might be visible, but the initial source of sound (instrument, voice, etc.) is hidden. Schaeffer makes this distinction to be able to discuss the sound itself, without having to relate to the origin of the sound. This is related to the technique that R. M. Schafer calls *schizophonia*, describing "the splitting of sounds from their original contexts" (Schafer 1994: 88). Schizophonia, according to Schafer, was introduced as part of "the electric revolution", and is central to the new possibilities of the electroacoustic transmission and reproduction of sound

[9] "Name given to the disciples of Pythagoras who, for five years, listened to his teachings while he was hidden behind a curtain, without seeing him, while observing a strict silence" (*Larousse* cited in Schaeffer 2004: 76).

(ibid.). Schizophonia can thus be seen as a basis for the modern understanding of acousmatic listening.

Although Schaeffer thought that acousmatic listening could encourage reduced listening, decoupling the sound from its source, Michel Chion suggests that the visual separation of sound from its source actually makes the listener ask what causes the sound (Chion 1994: 32). Smalley claims that, "we cannot separate space itself from what produces it, nor from our experience of space in nature and culture" (Smalley 2007: 54). With the term *source bonding* he describes this "intrinsic-to-extrinsic" link between the inside of the work and the sounding world outside (Smalley 1997: 110). He further defines source bonding as "the *natural* tendency to relate sounds to supposed sources and causes, and to relate sounds to each other because they appear to have shared or associated origins" (ibid., emphasis in original). As previously mentioned, Brøvig-Hanssen and Danielsen suggest that we make sense of virtual spaces by comparing them to previous experiences with actual spatial environments (Brøvig-Hanssen & Danielsen 2013: 71).

In the case of recorded space the sound of spatiality is separated from any visual clues (as well as those of smell and tactility). Some details of the room(s) are lost, but it is still, to a large extent, possible to perceive some sense of spatiality. The recorded space still functions as a kind of reference to real spaces, due to source bonding. The listener tries to figure out what the origin of the sound is, thus comparing the recorded space to previously experienced spaces (both actual spaces and recorded spaces). The recorded space might thus, due to its references to other spaces, be a transtextual space.

3.3.1 The realism of virtual spaces

Based on this positing of recorded space as a transtextual space, some recorded spaces can be interpreted as more realistic than others. Brøvig-Hanssen and Danielsen (2013) study the realism and surrealism of different recorded spaces and how surreal spaces become "accepted" through a process of naturalization. They mention the early use of electrical microphones to amplify the voice and how this was received by the audience as something unnatural (Brøvig-Hanssen & Danielsen 2013: 75). But as people grew accustomed to the microphone it gradually became part of the natural. The culture of listening is thus in constant change, as new ways of "staging" sound become naturalized. A basis for this is their understanding of recorded sound as virtual space-forms for "denoting a spatial environment that, rather then [sic] actually existing as physical space, is solely implied by the sound"

(ibid.: 72). This suggests, as I also discussed above, that the recorded space is never an actual space, but rather some kind of illusion. However, as we surround ourselves with such illusions, we eventually stop recognizing them as such.

The norms of the individual sounds within a recorded space seem to have different frameworks from genre to genre. In country music it would, for example, be absurd to hear a drum machine, whereas a pedal steel guitar, which is quite normal in country, is highly unusual in hiphop music. When it comes to the more general traits of spatiality, and location in particular, as previously discussed, the norms are more global, particularly regarding more traditional band line-ups. This might be related to the norms of live staging. In rock (and similar genres) the instruments are arranged across the stage in much the same way as in a normative/diagonal mix. Genres within electronic music are, in contrast, more explorative, and are less bound by the norms adhering to a live band playing on a stage. It can, for example, be played live on stage by only one person. However, it might also be partly based on the norm of centralized bass in rock music.

Brøvig-Hanssen and Danielsen relate the notion of natural and surreal in virtual spaces to source bonding. As mentioned, this has to do with the human inclination to recognize the origin of sounds. Although it could be physically impossible to create a real space equivalent to a particular virtual space, for example due to a combination of different local spaces within the global space, the "impossible" virtual space can still make some sense to us, because of our familiarity with the different local spaces. I discussed Brøvig-Hanssen and Danielsen's three examples of such "impossible" spaces in chapter 2. Although those spaces are also far from what could be regarded as realistic or natural, they, too, are no longer perceived as especially inaccessible to listeners (of course, depending on what the listener is used to). Such spaces can be the results of the use of several different reverbs, distortion and extreme use of equalizers and filters.

More discreet examples of surreal spaces deal with relative amplitudes, and are applicable to nearly all rock music (and much more). The volume balance between lead vocals and the rest of the band in a recording is usually the opposite of what it would be in the rehearsal room (at least without the help of a PA system). The change is due to both the amplifying of the voice through a microphone *and* the "attenuating" of drums and guitars, also through microphones. This indicates that especially our understanding of the voice as instrument, but also our understanding of the sound of different instruments, has been formed by the extended use of microphones to amplify sound, both in recording and in live performance.

Within this category of virtual, but natural, spaces, however, there are different degrees of realism. There are several different approaches to achieving realism in a production. Glyn Johns swears by recording the full band simultaneously, in the same room with few microphones (see Crane 2015: 48ff). Steve Albini also records bands simultaneously, but with a number of microphones. He also avoids using dynamic compression, so as to preserve the played dynamics of the musicians (Tingen 2005). Both Johns and Albini claim that "live" recording, even though some flaws might be introduced, is more adequate for capturing the interaction between musicians (Crane 2015: 48; Tingen 2005).

A contrasting direction that grew common in pop and rock in the 1980s, after Solid State Logic (SSL) had introduced their 4000 series mixing consoles, with inline dynamic sections (compressors and gate/expanders) on all tracks, in the late 1970s,[10] was to take great control of the recording by using a large number of tracks, and adding compression to several of them, as well as to the mix buss. The track-by-track approach allows, to a much larger degree than the live approach, the re-recording and perfection of all parts of the recording. Some, such as Johns and Albini, argue that some of the "magic" is lost with this approach.

Although many recorded spaces that would be unthinkable 20 years ago have become naturalized to most listeners, there are still examples of spaces that sound peculiar. As an example of extra-diegetic events in music, Walther-Hansen discusses Air's "Radio #1" from their 2001-album *10.000 Hz Legend*:

> At 3'14 we suddenly hear a male voice humming along to the chorus. The male voice sings along to the track and, thus, takes part in the performance. He appears, however, temporally separated from it. When the voice enters, the listener suddenly hears the initial performance through another diegetic layer. In this way the listening perspective changes (Walther-Hansen 2015: 43).

The effect is achieved by ducking other sounds with side chain compression triggered by the male voice (ibid.). The recorded space is already rather fragmented, consisting of several very different and separate spaces. However, this is not something especially radical, as similar spaces can be heard in 1970s and 1980s art rock.[11] What is interesting in "Radio #1" is that the humming male voice, when it enters, breaks the previously settled boundaries of

[10] This feature, along with the SSL mix buss compressor, has certainly been a great contributor to what has been dubbed "the loudness war".
[11] A more peculiar example discussed by Doyle (2005) is Johnny "Guitar" Watson's 1954 single "Space Guitar", in which both saxophone and electric guitar alternate between completely dry and deeply reverberant spaces (Doyle 2005: 159).

the recorded space of the song. Suddenly the front of the space comes even closer to the listener.

A similar example, which could also be regarded as an extra-diegetic event, appears in Daughter's song "Winter" from *If You Leave* (2013). The song opens with a dense space, filled with reverse reverbs and wide delays. The voice sings softly and is doubled in unison. Everything seems to be floating. That is, until 1'14, when a rhythm played with rim shots is introduced. While all other sounds are characterized by the use of digital reverbs, the rim shots sound like they are recorded at some distance in a rather small room, having a distinct pre-delay but a short reverb tail. The rim shots appear thus to be in a space separate from that of the rest of the band. This space is much more "live", or "homey", than the band's space, which sounds very "produced". It is a space that can easily be associated with do-it-yourself home recordings, and thus acts as an intrusion to the well-produced studio space of the rest of the band. Returning to the discussion of diegesis, this homely space suggests that the rim shots are played along with the previous events (the band), in much the same way as the male voice in "Radio #1" hums along with the diegetic events.

In light of the discussions above, about "live" and "produced" records, Schafer's distinction between hi-fi and lo-fi soundscapes can be interesting. According to Schafer, hi-fi soundscapes are sonic environments in which each sound source is easily separable and has a great sense depth (Schafer 1994: 43). The opposite, lo-fi soundscapes, are defined by their lack of clarity. Here, the sounds are more blended and immediate (ibid.).

The case of the recorded space is somewhat different, but there is still some sense of Schafer's notion of hi-fi and lo-fi. For example, in the Ryan Adams song "Dirty Rain" from the album *Ashes & Fire* (2011), produced and engineered by Glyn Johns, all sounds are clearly identifiable. Johns' approach to dynamics allows for the space to have a lot of depth. In some regards, the sense of hi-fi can be connected to the sense of realism in Johns' productions. In the common sense of fidelity, as the "degree or exactness with which something is copied or reproduced" (*Oxford Dictionaries* 2016), hi-fi in recordings can then seem to be reflecting realism. Schafer's notion of fidelity, too, is applicable to Johns' productions, as the natural dynamics give room for depth, and sounds are clearly separable. However, what is less realistic and "true to nature" in "Dirty Rain" is the use of artificial reverb, most prominent on the lead vocals. This effect can still be regarded as contributing to Schafer's notion of hi-fi, however, as it gives even more depth to the recorded space.

A less hi-fi space can be heard in "A Shot in the Arm" by Wilco from their album *Summerteeth* (1999). The space has some depth, but the boundaries between the sounds are

blurred by the density of the space – a feature resulting from the many chaotic sounds as well as extensive use of reverb. This also makes the depth of the space less noticeable. It is still problematic to call it a lo-fi space, though, since there is some sense of separation and some depth. Most well produced spaces have some degree of separation, and it is difficult to find examples that lack both depth and separation. In the case of "A Shot in the Arm" the lack of separation is used deliberately to create a sense of chaos.

There seems thus to be a relation between the notions of naturalized and surreal, realism, and hi- and lo-fi spaces. Although no recorded space can be regarded as the same as a real space, many spaces can, as a result of naturalization, be perceived as familiar as natural spaces. Naturalized and realistic spaces are not necessarily the same, but the processes of naturalization seem to have made listeners less aware of the difference between unrealistic and realistic spaces. Furthermore, realistic recorded spaces seem to represent the fidelity of their actual space equivalent to a larger extent than unrealistic spaces, meaning that if an actual spatial situation can be described as a hi-fi soundscape, the realistic recorded space will also resemble a hi-fi soundscape. The recorded space, as a virtual space, can thus be regarded as realistic or unrealistic. However, even the unrealistic spaces can be naturalized and thus be perceived as less unnatural than they actually are.

3.3.2 Sonic markers and transtextuality

Elaborating further on the possible transtextual functions of recorded spaces, I will employ Askerøi's term sonic markers. Sonic markers are, according to Askerøi, "musical codes that have been historically grounded through a specific context, and that, through their appropriation, serve a range of narrative purposes in recorded music" (Askerøi 2013: 16). He suggests that sonic markers "can be identified through the explicit use and construction of expressive devices in music that range from vocal peculiarities to instrumental stylings and the technological aspects of production" (ibid.: 2). Sonic markers can supplement transtextuality as a way of understanding origin and meaning in recorded sound.

Sonic markers can relate to several types of narratives, including those of place, politics, musical style, and time. The application of sonic markers as a part of the recorded space, and the re-contextualization of these, can link the recorded space to an array of contextual meanings. Relevant to the discussion of transtextuality is the appropriation of sonic markers into new texts, as when a sound or feature (for example a type of space) previously established as a sonic marker is simulated. Askerøi argues that "sonic markers are

constructed thanks to different forms of appropriation or recontextualization of musical codes" (Askerøi 2013: 139). Thus, musical codes gradually achieve their meaning as sonic markers when they are repeatedly transtextually referred to with the same purpose. In the discussion of Joy Division, Askerøi suggests that the band appropriated already established sonic markers of punk: "one could argue that the band appropriated sonic markers of punk in order to communicate something more complex, and at the same time keep its relation to punk intact" (ibid.). The sonic markers of punk, then, communicated the narrative in which they were originally constructed, as well as a retrospective relationship to punk.

In the case of spatiality, sonic markers can for example be the spatial arrangement, or the internal spatial features of an instrument. A sonic marker can also refer to a specific production style, for example what is called the "Motown bass", which mostly relates to the playing style of James Jamerson, one of the regular bassists for Motown Record's releases in the 1960s. The Motown bass was constructed as a sonic marker through the repeated use of Jamerson (and his playing style) for a majority of the recording sessions at Motown. When this sonic marker is then reused in new contexts, for example in contemporary recordings where Jamerson's playing style is imitated, it still functions as a sonic marker of the Motown sound. However, its narrative purpose is altered, as it is not actually a bass recorded in Motown's studio in the 1960s with Motown's producers, engineers and musicians. It is used at another time and another place, and possibly also in another genre of music, but as a reference to the aesthetic desires of the Motown producers, engineers and musicians.

The Motown bass can then also be interpreted as a sonic marker of time, as the Motown sound is mostly related to the label's music in the 1960s. Askerøi applies the term *retronormativity* to describe "the mechanism of placing the 'past' in the 'present' […]. Retronormativity further implies a nostalgia for technological artefacts in the return to the 'sound' of the 1950s, 1960s, and so on" (ibid.: 42). Further, he explains that "[r]etronormativity builds strongly on appropriations, meaning that sonic markers are never really copied but rather constructed by being incorporated into an entirely new musical language" (Askerøi 2013: 67). This means that a Motown bass sound, when used in contemporary music, brings previous narratives of for example technological development into a new narrative. The Motown bass as a sonic marker of time, when appropriated in new music, then communicates the role of the bass, and its production, as it was considered in the 1960s, while also maintaining the music's nostalgic relationship to the Motown sound. It can thus be seen as not only a sonic marker of 1960s soul music, but also as a sonic marker of nostalgia.

An interpretation of Motown bass in new music as a transtextual reference must be seen slightly differently. Rather than placing a narrative of the past in the present, it points backwards to a specific text or a common feature of a set of texts. The result is an interpretation of the current text where the original text is in focus rather than the context in which it occurred. However, the context will inevitably have an impact on the text, meaning that a transtextual reference also brings the context of the original text into the interpretation of the current text.

A sonic marker, then, refers to a convention specific to a time, place, style, etc. It is, as previously mentioned, a code that has been repeatedly used in a specific context, which then becomes a kind of hook to which the listener refers. Transtextuality, on the other hand, denotes a reference to a specific text. The reference can be to a single feature that happens only in one text, for example a melodic line, or more relevant in this context, a type of delay that was used in a song but was atypical for the genre and/or time.[12] However, the original text to which the transtextual element refers also necessarily has its place in some kind of larger discourse which will affect the current text. The two approaches can thus seem very similar, while taking different points of view. A sonic marker can be recontextualized in a new text, thus being the subject of transtextuality. Whether it becomes a new type of sonic marker in the new (con)text or remains a recontextualized sonic marker depends on the new narrative it becomes part of. Moreover, the term sonic marker refers to the larger narrative in which the musical code appears. Transtextuality refers to the specific text that the element is taken from. Thus, the focus of transtextuality is the meaning the reference brings to the current text, whereas sonic markers describe the roles of musical codes when they are repeatedly recontextualized with the same purpose.

Sonic markers, then, denote the identity and narrative functions of sounds that are produced in a specific way. Thus, they affect how the song, or text, is interpreted. This is also the case of transtextual references, but a juxtaposition of transtextuality and sonic markers is problematic. A sonic marker is initially a function of an original text, one that *can* be re-contextualized in a new text. It can refer to external meanings, but it can also, for example, be a result of technological development and not act as a direct textual reference. On the other hand, transtextuality is a feature of the new text when a previous text is re-contextualized in it.

[12] An example that comes to mind is a delay effect that is applied to Roger Daltrey's voice in the song "Bargain" from the album *Who's Next* (1971) by the Who. At the end of each chorus, Daltrey repeats the last line, "the best I ever had", and his voice is "pulled" towards the back of the sound-space.

Spatiality as sonic marker and transtextual reference

Another example of a sonic marker of time is the triangular mix. As previously discussed, the triangular mix is related to the early stages of stereo in recorded popular music, as a result of technical limitations. The label and studio Daptone has a distinct retro orientation to their releases, and in the music of one of their most popular acts, Sharon Jones & The Dap Kings, the triangular mix is used actively to revoke the sound of 1960s soul and RnB. For example, on the album *Give the People What They Want* (2014), most of the songs are mixed with bass and drums on opposite sides, and with lead vocals (sometimes also backing vocals) centralized. This creates a triangular configuration, and the centralizing of only the vocals hints at the technical limitations in early stereo mixing and the need for mono compatibility. Sharon Jones & The Dap Kings, then, make use of the recontextualization of sonic markers of time to connect their music to a specific era. As a result, the narrative of the technological development of stereophonic sound is brought into the music. At the same time, as in the case of Joy Division and punk, the use of sonic markers of time communicates the band's relationship to the recordings of the late 1960s.

I have mentioned some examples of how the recorded space can act as sonic marker or transtextual reference, for example that the triangular mix can be seen as a sonic marker of time. Also, the slap-back echo in 1950's rock and roll is a distinct sonic marker of time. To conclude this subchapter I will discuss another two examples of how aspects of the recorded space can work as sonic markers and/or transtextual references.

Intimate spaces and authenticity

As I suggested in chapter 2, the exposition of Lucinda Williams' voice on the album *Car Wheels on a Gravel Road* (1998) can be read as honest and authentic. I would further argue that the proximity in this case also functions as a sonic marker of intimacy and authenticity. Proximity as a sonic marker of intimacy is a historically grounded narrative function, starting with the advent of crooning in the 1930's. As Simon Frith (1986) points out, crooning was made possible by the invention of the electrical microphone (Frith 1986: 263). The crooners sang softly, but the amplification of the voice made for an "unnatural" illusion of intimacy. This was not well received by all. Frith writes: "Microphones enabled intimate sounds to take on a pseudo-public presence, and, for crooners' critics, technical dishonesty meant emotional dishonesty" (ibid.: 264). The sound of crooning was described as sentimental and effeminate (ibid.: 263). Brøvig-Hanssen and Danielsen refer to Michael Chanan, suggesting that "this new voice was no longer intended to reach every listener in a public space; instead, it was

designed to sound like each listener's best friend" (Brøvig-Hanssen & Danielsen 2013: 75). The origins of the proximate voice hint thus to an interpretation as sonic marker of not only intimacy, but also of sentimentality. In the opening track of *Car Wheels on a Gravel Road*, "Right in Time", the sound of the voice then underlines the lyrics: "Not a day goes by, I don't think about you. You left your mark on me, it's permanent, a tattoo" (Williams 1998).

The sense of authenticity in a proximate voice follows from its intimacy and a transtextual reference to actual spaces. As suggested above, the sounds of actual spaces can act as texts, in that they can be interpreted to reveal clues of size, materials and position. The proximate voice functions then as a reference to an actual spatial situation where the sound-source (the person whose voice is heard) is very close to the listener. In an actual situation, our personal space is reserved for a very limited number of people that we trust. When Williams, with some help from the proximity effect, enters the listener's personal space, it does not feel like an intrusion. She can be trusted. We get a notion that Williams' true emotions are conveyed through the song. Is the proximate voice in Williams' music, then, a recontextualized sonic marker of authenticity? While the proximate voice as a sonic marker of authenticity can be seen in many songs throughout the history of recorded popular music, it is difficult to point to the beginning of such use. The reference to intimacy, although not necessarily used as a compositional tool, is easier to pinpoint to a specific time. The reading of the proximate voice (a sound that was first received as unnatural) as authentic has been made possible by its naturalization, which has opened up for more positive interpretations with reference to actual space. The proximate voice as a sonic marker of authenticity is then a result of a long process – it has been culturally grounded through repeated application.

"Live" and realistic spaces

I discussed above how some recorded spaces could be interpreted as more realistic than others, related to their sense of "liveness". Building on this, I would suggest that a recorded space of specific features could act as a sonic marker of liveness. This has become more relevant after the ascendancy of track-by-track recording, and even more so in the contemporary realm of home recording and digital manipulation of sounds. In this context, liveness as a sonic marker is based on our interpretations of live/actual spaces. It is a code that has been established through our knowledge of, and continued listening to, live music as well as live recordings. It is also often used as a compositional tool, for example as a sort of statement against the increasing tendency to record musicians separately. Johns comments,

> there have been loads of cases where I've been forced into a situation where I had to build [songs]. But I don't think it makes for great records. […] The recording process was invented originally to capture the performance of a piece of music. That's what it was there for. Unfortunately now the tail's wagging the dog a little bit (Johns cited in Crane 2015: 50).

Even when not actually recording live, it is still possible to create sonic markers of liveness in a recorded space. Those cases might, however, be seen as simulations of live recordings, as they strive to sound like all instruments were recorded simultaneously in the same room (even though the song can actually have been recorded track-by-track over the course of several years).

The different cases of live recording and simulated liveness can be difficult to distinguish from each other when listening to them. They can thus appear as the same type of sonic marker (of liveness or realness). However, the simulated liveness also takes advantage of new technology to preserve the option of "fixing" individual tracks at any time. Simulated liveness acts thus less as a statement (a sonic marker of opposition) against a recording technique, and more as a statement against a modern "overproduced" sound. Although liveness in recordings (or the lack of it) has clear transtextual references to actual spatial situations, it also acts as a sonic marker, since it is based on the active choice of the producer/engineer. All in all, it is problematic to discuss recorded spaces as not being the results of compositional choices and tools. A producer/engineer has, in most cases, ideas on how the recorded space should be formed, thus contributing to the "composition" of the sound.

3.4 The corporeality of recorded space

Going even further into the discussion of the transtextual connections between virtual recorded spaces and our experiences with real spaces, I will revisit Middleton's search for a theory of gesture, addressed in the article "Popular Music Analysis and Musicology: Bridging the Gap" (1993). The title refers to his aim to bridge the gap between formalist musicology and the very context-oriented popular music analysis (Middleton 1993: 177). He suggests that, "how we understand musical sounds is organised through processual shapes which seem to be analogous to physical gestures" (ibid.), and he further points out that popular music, more clearly than certain other musics (Middleton mentions classical symphony and chamber music) is "rooted in the structures, inner processes and operational patterns of the secular human body. Even with pieces not intended for dancing, listeners usually find themselves *moving*, really or in imagination" (ibid.: 178). In other words, he suggests that sounds can in

themselves be understood *as* gesture, and thereby also *through* gesture. In this context, he finds Moore's sound-box to function as a bridge between the popular music recording and a gestural totality: "it is within the four-dimensional space-time of this imaginary cuboid that the gestural intersections take place" (ibid.: 179).

Moore refers to what he calls the "tactility" of recorded sound to describe the sound-box, suggesting that a recording carries a "feel" (Moore 2012: 29). This feel is often the first thing to "hit" the listener, in much the same way as Brolinson and Larsen's description of sound as something that can be heard in a very short section of music (Brolinson & Larsen 1981: 181). It is also, comments Moore, "often the hardest to discuss" (Moore 2012: 29), being somewhat vague and comprising a complex set of parameters that have no established terminology to describe them.[13] The feel carried by a recording is the same as its textural space, which is what the sound-box attempts to conceptualize. I will discuss, in the remainder of this chapter, the similarities between the perceived tactility of recorded space with the "feel" of real spaces based on the understanding of transtextuality, as discussed in the foregoing chapter. What I seek to do is to elaborate on why and how the recorded space is relevant to the listener in terms of corporeality. In other words, how are listeners' experiences with body movements in different spaces relevant to their experience of recorded music? To begin, I will address the subject of corporeality, and try to define a notion of corporeal spaces as an alternative to what Moore calls "feel".

Corporeality

Oxford Dictionaries define "corporeal" as "[r]elating to a person's body, especially as opposed to their spirit" (*Oxford Dictionaries* 2015). Corporeality is thus descriptive of physical experiences, for example of being and moving in physical spaces. It includes, I would argue, the processes of listening, which are results of physical reactions to sound in the auditory system. In turn, this includes experiences of orienting oneself in a space (for example a room) by the help of its acoustical properties with respect to the corporeal. As explained in the beginning of this chapter, the acoustics of a space are determined by the space's physical dimensions and surface materials, which again result in spatial features like reverberation and echo (or the lack of such). The acoustics of a space also affect the

[13] There have, however, been attempts to find a terminology, for example Moore's sound-box (2001), and Danielsen's sound-room (1993), but these cannot be said to have settled to any extent outside of popular music studies.

perceived location of sounds in relation to the listener, thus communicating something about the sound's whereabouts across the width, height and depth dimensions.

As a way of getting to grips with a relation between acoustics (the sound of a space) and gesture, the understanding of sounds as gesture can be fruitful. The connection between sonic spatiality and body movement is somewhat blurry, but it is of interest to the understanding of the ways in which we perceive spatiality in recorded sound. In his investigation of gesture in music, Middleton focuses on the similarities between gesture and rhythm, suggesting that, "in some sense a theory of gesture would be also a theory of rhythm" (Middleton 1993: 177). This theory of rhythm includes "the micro-rhythms responsible for the inner life of sounds themselves, and the quasi-'spatial' rhythms organising the hierarchies of relative pitch strength and tonal tension" (ibid.: 178f). In other words, through the frequency of sound waves, all sounds are comprised of some sort of rhythm, and thus also of gesture.

Kinesphere

An approach based on the micro-rhythms of sounds alone might, however, be difficult to relate to corporeal experiences of spatiality. There are certainly micro-rhythms, in the sense that Middleton uses the term, in for example reverberation (pre-delay time and the intervals between each reflection), but to grasp the larger aspects of corporeal spatiality I will apply the term *kinesphere*. Kinesphere is used in the Laban Movement Analysis to describe an "imaginary box surrounding a person that defines the maximum movement possibilities from a certain point in space. Laban argues that the kinesphere is a mental construct that one is always aware of in the course of interacting with the environment and with others" (Jensenius, Wanderley, Godøy, & Leman 2010: 20). Based on this description, the kinesphere could thus be defined as the amount of space in which one feels comfortable moving. Rudolf Laban's own definition suggests that the kinesphere is "the sphere around the body whose periphery can be reached by easily extended limbs without stepping away from that place which is the point of support when standing on one foot" (Laban 1966: 10). He clearly limits the kinesphere to actual bodily movement. However, the notion of the kinesphere as a mental construct suggests that its size is determined by how populated the subject's environment is (how much movement is physically possible), and, I would argue, by the size of the surrounding space (depth, width and height). Furthermore, spatial awareness is partially determined by the acoustics of a space, suggesting that the materials of the space's surfaces (which affect sound reflections) affect the size of the kinesphere.

Gibson's affordance theory is relevant here. The theory describes the relationship between the needs and capacities of a perceiving organism and the properties of its environment, as well as the relationship between perception and action (Clarke 2010: 106). Similarly, the kinesphere is determined by the movement possibilities afforded by the environment to the perceiver. A large empty room can afford large movements, resulting in a large kinesphere, whereas a small crowded room is likely to afford only small movements, resulting in a small kinesphere. As in the affordance theory, this relationship is reciprocal,[14] meaning that the kinesphere is also determined by the perceiver's ability to move, comfort with moving, and the perceiver's size in relation to the size of the room.

The idea of a kinesphere can be transferrable to music through the theory of motor-mimetic processes. Zeiner-Henriksen discusses this theory in relation to his study of body movements to electronic dance music: "The theory of motor-mimetic processes of perception concerns the way motor networks used to perform a certain action are partly activated already in the act of observing someone else performing it" (Zeiner-Henriksen 2010a: 127).[15] Could it be, then, that our perceptual systems respond in similar ways to the sound of a space as to the sight of it? If so, the density of a recorded space, and any holes in this density, can be connected to what I will call a *sonic kinesphere*.

Sonic kinespheres

There are two possible ways to describe sonic kinespheres: one in which the listener is regarded as the subject, and one which, somewhat more abstractly, considers the sounds within a recorded space as subjects. The latter builds upon the understanding of sound as gesture, suggesting that the kinesphere can be used as a corporeal metaphor for the perceived space any sound in a recorded space has at its disposal. Every sound can thus be regarded as having its own kinesphere, not unlike the notion of internal spaces, based on the density of the space and the location of every sound in relation to the spatial dimensions. In a dense space, for example, the sonic kinespheres of sounds will be smaller than in an open space. The presence of holes, to use Moore's terminology, will give a greater sense of separation, giving the sounds a greater perceived potential for "elbowroom" (although the location of the

[14] The same object can afford different uses to different perceivers. Clarke's example is that "[t]o a human being, a wooden chair affords sitting on, while to a termite it affords eating" (Clarke 2010: 107).
[15] As Zeiner-Henriksen comments, this link is supported by the more recent discovery of "mirror neurons", which also apply to sound. Zeiner-Henriksen refers to the neuroscientist Christian Keyser and his colleagues, who, in 2003, discovered that audiovisual mirror neurons discharged "regardless of whether the subjects (monkeys) performed, heard or saw a specific sound-related action (Zeiner-Henriksen 2010a: 127).

sounds might be unchanged). This also means that the perceived size of sounds has an impact on their kinespheres.

The other way of describing sonic kinespheres regards the listener as the subject, placing her "inside" the space of the recording. This definition is more directly related to our role as perceivers of recorded space. The perceived distance between each element affects the size of the perceiver's sonic kinesphere. A crowded recorded space can be transtextually compared to a crowded actual space, as the perceiver experiences a sense of possibility for imaginary movement within the recorded space. Returning to the case of body movement, one could perhaps say that holes in the recorded space also afford actual movement, since the differences between up and down, and right and left are more prominent. When the extremities of the recorded space are emphasized we are more easily drawn towards them.

Zeiner-Henriksen (2010a) suggests a correlation between verticality in music (high and low frequencies) and vertical body movements. He cites Björn Vickhoff, who comments that "[a]lthough there are no obvious directions of melody movement, most listeners feel directions in music. When the melody is moving 'upwards' or 'downwards' you get a feeling of spatial direction" (Vickhoff cited in Zeiner-Henriksen 2010a: 128). A similar sense of horizontal direction can be detected, I think, suggesting that the opening up of space on one side of the recorded space will give a larger kinesphere in that direction, a feature that might "pull" the listener towards that side. As explained above, these kineto-dynamic aspects of music don't necessarily result in actual body movement, but they might have a parallel effect on the perceptual system. This idea of direction in music might thus be a foundational element of a sonic kinesphere as a concept for describing a sense of possibility for (imaginary) movement.

Summing up, recorded space contains sense of corporeality. I have argued that the perceived corporeality of a recorded space is a result of transtextual references to the corporeality of real spaces, and that this corporeality can be understood through the idea of a sonic kinesphere. This suggests that when listening to music we immerse ourselves in the recorded space, and that the sounds in the recorded space have a "tactile" distance both between themselves and from us. Much like being in a room with various objects of differing size and distance from one another, we get a sense of the potential for movement in some directions rather than others. Thus, the notion of a sonic kinesphere substantiates the idea that recorded spaces can act as transtextual references to actual spaces, in a way that involves the corporeal experiences of the listener.

The idea of a sonic kinesphere offers an additional way of exploring the spatiality of recorded sound. This builds partly on Moore's notion of tactility in music, and it can be seen as an alternative to his understanding of holes in the density of a sound-box. However, the concept of a sonic kinesphere further emphasizes the inextricable link between music and the body, connecting our experiences of music to our everyday experiences of being and moving within different spaces. What this means is that, for example, the feeling of being in a room where the possibilities of movement are few can be evoked by a recorded space in which the sonic kinesphere is limited by the density of the space.

3.5 Definition of sound-space

In the previous chapter I discussed different approaches to spatiality in recorded music. Most relevant to this thesis are Danielsen's, Moore's and Moylan's different models. All these consider the various ways in which recorded spaces have been or can be formed, but with differing focus. What they do not explore to any considerable extent is how the forming of a recorded space affects the meaning of a song. Although spatiality in itself can be said to have some sort of implicit meaning, its potential for affecting the interpretation of a song needs to be further emphasized.

A focus on how spatiality can be used to affect meaning is particularly useful in production contexts, to comprehend the role of producers and engineers in the forming of a recording. An approach also including the theories discussed in this chapter – transtextuality, sonic markers and corporeality – has the potential grasp the impact of spatiality on the total impression of a recorded song. The discussed theories are in themselves powerful in the interpretation of music, placing the music in certain contexts. However, I believe that the incorporation of these theories in a spatial approach is easier to adhere to for a producer or engineer. In addition it provides a comprehensive understanding of the ways in which spatiality affect musical experiences. As I will demonstrate in the following chapter, an approach like this can also provide insights into how the spatial location and forming of different elements can affect the narrative of a song.

As there are no existing spatial models that incorporate such aspects of the meaning of recorded music, I will introduce a new concept, which I call the *sound-space*. The sound-space is *a production-analytic model that comprises the four-dimensional spatiotemporal organization of a recording as well as the ways in which that organization evokes or emphasizes meanings in the recorded music*. In the following I will discuss the theories of

transtextuality, corporeality and sonic markers in relation to the sound-space as a production-analytic tool, in order to show how producers and engineers can use this model to affect the resulting narratives of their productions. Furthermore, I will point to the ways in which it raises awareness of the corporeal involvement of the listener in recorded sound.

4 Working with space in recording practice

In this chapter I will put the concepts discussed in the theoretical part to a test by describing two closely related sound-spaces in a specific production context. This test of theory is based on two productions of the same song, which are produced, recorded and mixed by me. The first of the productions was produced, recorded and mixed before the work on this thesis had started. Thus, this first production process was not documented, and the discussion of it will be based on my recollection and what can be deduced from the audible result. In the second production, I have tried to apply the notion of sound-space as a production tool, and my aim is to see if this has had any effect on the mix, when compared to the first production. Because this approach resembles practice-led research, I will start with some methodological consideration regarding the use of own practice as part of a research process.

4.1 Methodology

The use of own practice as part of research poses certain challenges. Most important is the question as to how practice and the artifact resulting from that practice can be used as research, and to what extent it can be a substitute for traditional research. Mäkelä and colleagues (2011) discuss such issues by investigating the ways in which the terms *practice-based research* and *practice-led research* have been applied differently in Australia, UK and Finland. In the case of Australia, the authors refer to Linda Candy, a researcher interested in using practice-led methods. Candy differs between practice-based and practice-led research. "According to [Candy], practice-led research is concerned with the nature of practice and leads to results which have operational significance for that practice. […] On the other hand, she defines practice-based research as an original investigation undertaken partly by means of practice" (Mäkelä, Nimkulrat, Dash, & Nsenga 2011: para. 2). In academic work, she suggests, creative outcomes can help to demonstrate claims of originality as well as "contribution to knowledge": "She suggests that although the context and significance of claims are described in words, a richer experience can be obtained through the practical outcomes (i.e., artefacts) created" (ibid.).

This could mean that the use of a practice-based approach as part of this thesis, through exemplifying the outcome of the theory discussion, could contribute to a "richer experience" or a better knowledge of the subject, possibly demonstrating the benefit of

applying the sound-space in production contexts. Still, a practice-led research is more relevant to the practice related to this thesis, as it is meant to affect certain operational processes of record production. In Finland, "practice-led research has been adopted to highlight the active role of professional practice in the research process" (Mäkelä et al. 2011: para. 2). Also, since 2007, some authors have preferred the term practice-led over practice-based, "in order to acknowledge the change in emphasis from the production of original artefacts to the integration of artistic practice into the research process" (Lycouris referred to in ibid.).

In central Europe, "artistic research" is yet another term used to describe the practice-based research discussed above. Mäkelä and colleagues recognize three trends in artistic research and practice-led research. One trend suggests that, "art can stand on its own in a university context" (ibid.: para. 4), meaning that the artistic process can be regarded as a line of study equal to scientific research, and that "further explanation and validation of the practice are not as important as the practice and the artefacts created" (ibid.). A second trend suggests that a textual explication, following the basic structure of traditional research, should be presented alongside the artifact (ibid.). This requires "equal partnership between artistic practice and research practice", "the role of an artist as a researcher investigating a research question", and "thorough documentation of the researcher's artistic practice which gets involved in tackling a research problem" (ibid.). Mäkelä et al. call this trend an "academically-attuned practice-led research" (ibid.). The third trend suggests that practitioners or academics "need to prevail upon the university to change the understanding of research and to acknowledge the unique types of results that may be possible through artistic research" (ibid.). For this trend the authors suggest the following core question: "Can the academic notion of research be extended to include the unique results possible through artistic research?" (ibid.).

All three trends discussed by Mäkelä et al. tend to rely heavily on the artifact resulting from practical work. In this thesis it is not the resulting artifact, but rather the process of making that artifact that is under consideration. Therefore I will not take a position regarding the three trends. I believe, however, that although the process of making art cannot necessarily be used as a method of research in itself, the involvement of the researcher in the investigated process can be an exceptional resource to the research. Accordingly, I think practice-led research is a suitable description of the practice-related work in this thesis, as the resulting artifact of my practical work is not supposed to be counted as part of the research but will complement the discussion of theory rather than stand out as a separate contribution.

(That said, I hope that the artifact resulting from the practical work of this thesis will also stand out as an artifact as such, but that's another matter.)

My aim for the following part of this thesis is to exemplify the theories discussed in the first part. This process has given me a unique ability to test and revise the theories in quick succession. There are, however, certain potential downsides to an approach like this. The most obvious is that it involves a high level of subjectivity. However, the same can be said, at least to some extent, of all research, especially within the humanities. Nevertheless, since what is to be investigated is my own practice as a producer and engineer, this means that my conclusions will necessarily be somewhat saturated by my preconceptions about the production process. So although some degree of subjectivity is inevitable in any type of research, it is all the more important to take a critical approach to keep the research relevant outside of a personal realm.

A second disadvantage of a practice-led approach is that it does not account for the fact that different producers and engineers in many, if not most, cases will have different ideas about the record production. Thus, focusing on only one producer/engineer will necessarily ignore many nuances of the art of record production as a whole. The practice-led approach I am taking in this part of the thesis must therefore be critically and theoretically substantiated, and can thus be seen as a variant of interpretative humanities-based research. Thus, although the single point of view to record production investigated in this thesis is my own, it is not private. It is influenced by the different approaches of other producers and engineers, and related to these other approaches. It might therefore also be applicable to other productions and relevant to other producers and engineers than myself. Furthermore, my thoughts on record production also draw on established theories on how perception of recorded popular music can be understood in terms of spatiality, for example Moore's sound-box.

The upside of this method is, as I have already mentioned, that being a part of the process that is investigated offers unique insight to the subject. I have first-hand access to the research material, and receive instant "feedback" on how the discussed theories work in practice. I have followed the investigated process systematically through reflection and analysis, making me what Donald A. Schön calls a "reflective practitioner" (see Mäkelä et al. 2011: para. 3). This means that I have a close access not only to how spatiality *can be* an important factor in record production, but also to how spatiality *is used* in record production.

4.2 Review of practical work

Before analyzing the two versions of "Die Young", I will describe the practical work of producing, recording and mixing a new version of Dingus' "Die Young". As mentioned above, the song has already been released as a single, which I also produced. What is different in the new version is that I have based most of the production-related work in the concepts discussed in this thesis. Although the first version's production process is largely undocumented, I will include some notes on my ideas for the production. The song is an upbeat song, but with a tragic story about a young man recalling his love, a woman who died at young age. The story is set some years after her death, when the young man realizes that she will never come back, thus concluding that he wants to die himself so that he can be with her.

4.2.1 Production process

As mentioned, the practice-led research in this thesis consists of two different productions of the same song. The song is by the alternative country band Dingus and is called "Die Young". The two versions of "Die Young" were recorded at different points of time, one in the spring of 2014 and one in the summer of 2015. The 2014-version was recorded in the recording studio at the Department of Musicology, University of Oslo (IMV), and was released as a single in January 2015. It was not initially intended to be part of this thesis, and, as mentioned above, the production process was therefore not documented. The 2015-version was recorded at Parachute Studio in Oslo, as well as at IMV, as part of Dingus' self-titled debut album (forthcoming August 2016). The structure of the song is the same in both versions, with any differences between the two amounting to mostly a matter of arrangement, performance, recording techniques and mixing.

In light of the concepts discussed throughout this thesis, I will analyze and compare the audible result of the two productions of "Die Young". As in any comparative analysis there are, however, certain uncontrollable factors to consider. The most prominent one is my development as a recordist, producer and mixing engineer in the time between the two productions, as well as the use of different studios in the two versions. These factors will be impossible to distinguish from the effects of the use of sound-space as record production tool in the analysis. On the other hand, my growing involvement with the sound-space has also been a central part of this development.

Recording and mixing the 2014-version (version 1)

The 2014-version of "Die Young" was recorded as part of an EP that was eventually released as a series of singles in 2014 and 2015. The recording was done on several occasions, with all instruments recorded separately. The basic tracks, including drums, bass and acoustic guitar, were recorded over the course of a weekend. Drums and bass were recorded simultaneously, but in different rooms, together with an acoustic guitar as "guide track". The bass was tracked through a microphone on the amplifier, as well as through a DI.[16] On the drum kit I applied what is called the Glyn Johns overhead technique (one microphone straight above the kit, and one by the floor tom), together with close microphones on all the drums. I also used a stereo ambience pair at some distance from the drums. However, since the drums were recorded in a rather "dead" room, with very little reverberation, there is little sense of ambience in the recording.

The "guide guitar" was tracked using a single ribbon microphone. Another track of acoustic guitar was recorded with a Blumlein pair (two ribbon microphones placed at a 90° angle). This one was supposed to be used in the final mix, without the guide guitar. However, both tracks were used. On both lead and backing vocals, as well as the banjo, I used large diaphragm condenser microphones. Electric guitar and lap steel guitar were captured with ribbon microphones at close distances. Thus, ribbon microphones were used on many of the sources, also including drums. Ribbon microphones are generally less bright than condenser microphones and many dynamic microphones, something that eventually contributed to a lower overall sound-space on this initial recording. All in all, the tracks do not reveal much of the spatial environment in which they were recorded, thus giving the opportunity to form the spatiality to a greater extent in the mixing process.

I had some ideas about using a spatial model as a guide when mixing, but mostly for working with panning and volume, almost like Moylan's sound stage. My initial goal was simply to complete a good mix. Within this I had some ideas on use of references. For example, I attempted to imitate the reverb sound used on Daniel Romano's *Sleep Beneath the Willow* (2011), a sound clearly inspired by 1960s records by country artists like George Jones and Porter Wagoner. This retro-orientation also inspired me to an extended use of ribbon microphones and a somewhat limited quantity of microphones when recording. All of this necessarily had an effect on the resulting sound-space (which will be discussed in the analysis), although it was not consciously planned.

[16] Direct input unit.

Recording the 2015-version (version 2)

As I will discuss in the analyses, I overlooked an opportunity in the 2014-version of "Die Young" to emphasize the contrast between the singer-songwriter, reminiscent of an early Bob Dylan, and the electric band. In the new version I have applied sound-space as a tool to work with this, for example by making contrasts in the recorded space through the mix. This has affected my choices in both recording and mixing, including microphone techniques, microphone choice and use of reverb and echo effects. I also wanted to create a sense that the band was playing together in the same room, and that this room would sound "homey", so that the listener could relate to it in a positive way, thereby emphasizing the joviality of the band's image.

Drums and bass

The drum kit was recorded using a variation of different microphones, for both getting a unifying overall sound of the kit, while also having control over each drum. For picking up the overall sound of the drums, I used a pair of AEA ribbon microphones in mid-side configuration. This technique is normally used with one cardioid microphone as a mid-microphone directed at the source and one bi-directional microphone, the side-microphone, with its null-point directed at the source. The signal from the side-microphone is split in two, and one of the two signals is phase-reversed. This technique is fully mono-compatible, as the side-microphone signals are cancelled out when in mono. It also gives a nice stereo image without being too wide, as the sound in each stereo signal is essentially the same, meaning that there is no imbalance between the two channels. The use of a bi-directional microphone, like the AEA, as mid-microphone gives more room reflections in the mix. The AEA microphones also gave a bright but warm sound that corresponded well to the other microphones used for the drums.

The kick drum was picked up through a Neumann U47, modified as a U47 FET, known for picking up very low frequencies, as well as an Audix dynamic mic. I used both a dynamic and a condenser mic for the top of the snare drum, and a condenser mic for its bottom. Royer ribbon mics were used for toms, giving a controlled and warm punch. In addition to the AEA mics, I used a Coles ribbon mic as overhead mic. This made the overall sound of the drums somewhat more focused. Lastly, I used a pair of small diaphragm condenser microphones for stereo ambience, as well as a single ribbon mic from which the signal was compressed hard for providing even more "room sound". The combination of close microphones with the mid-side pair, the overhead, and the ambience microphones gave

me the best of both worlds regarding realism (capturing the drum kit as one instrument) and controllability (the possibility of forming and placing the kit in specific ways in the sound-space).

The bass guitar, a Fender Precision Bass, was played through a Musicman HD amplifier and an Ampeg 810 speaker cabinet, picked up through a tube condenser microphone. The direct signal from the bass was also recorded through a DI. These two signals (DI and mic) were combined and the volume relationship later adjusted to provide the sound I wanted. The forming of the bass sound started in the recording process with the controlling of the amplifier's equalizer, to make the bass the lowest element in the sound-space, while still having definition without interfering with the kick drum. Muting the bass with a sponge also contributed to this.

Although the drums and the bass amp were placed in different rooms to give total separation, the musicians were in the same room. This gave them the most spontaneous and unhindered communication possible. I played the guide guitar in the control room with the guitar amp in yet another room, and the signal was sent to the drummer and bassist's headphones, together with the bass guitar signal and some of the drums. The guide guitar track was later replaced.

Acoustic guitar

For the acoustic guitar, I wanted a bright sound and a lot of natural ambience. I accomplished this by using a close microphone, as well as a stereo mic pair for ambience. The close mic was a large diaphragm tube condenser mic, with a somewhat boosted top end. Two small diaphragm condenser mics were used for stereo ambience. The signals from these were also delayed 20 milliseconds, which separated the ambience sound from the close mic sound and thus made it stand out more clearly.

Electric guitar

The electric guitar was played through a Peavey Classic combo amplifier. I wanted a roomy sound for the electric guitar as well. In addition to two close mics, a large diaphragm tube condenser mic and a ribbon mic, I used two ambience mics. An AEA ribbon mic was placed in front of the amplifier, about three meters away and pointed to the floor. The second mic, a Coles ribbon mic, was placed behind the amplifier, something that gave a rather different sound. The combination of the two close mics and the two ambience mics gave a nice, "roomy" sound. To create even more sense of space, I sent the signal of the close ribbon mic

to a tape delay machine, making a slap-back echo. Placed relatively far back in the final sound-space, the slap-back echo provided an illusion that the guitar was played in a room with hard surfaces giving a distinct pre-delay to the room reverberation.

Lap steel guitar

The lap steel guitar was played through the same amplifier and recorded with the same close microphones as the electric guitar. There were no ambience mics used. The amplifier has a built-in spring reverb that was turned up, giving a sense of a narrow but deep space. The tape delay was also used, but with a slower repeat rate than on the electric guitar.

Banjo

To capture both the mid-range energy of the banjo's body, as well as the high frequency sound of the strings, I used both a Coles ribbon mic and a small diaphragm condenser mic. The high frequency energy from the condenser mic makes the banjo take place in the top of the sound-space. In the mixing process, I also added a phaser effect that makes the banjo blend with the mix without disappearing. I reduced the effect level at some points to make the banjo "pop" out of the mix.

Vocals and harmonica

Just as with the acoustic guitar, I wanted to enhance the room ambience for the lead vocal and harmonica in the first part of the song. In addition to the main vocal mic, I therefore used the same stereo ambience mics as for the acoustic guitar. The main vocal mic was a large diaphragm tube condenser. This was the mic that best matched the singer's voice, from a selection of four different mics. These mics were also used for the harmonica, which was played by the singer. For the backing vocals, both the female and male, I used large diaphragm condenser mics. These were recorded separately, one track per voice.

Mixing the 2015-version

Regarding sonic quality, the sound I achieved in the recording studio was very close to the desired result. During mixing I have used equalizers mostly to achieve a tidy sound without clustering in any frequency spectrums. An example of this was getting the kick drum and the bass guitar to occupy different vertical points of the sound-space so as to avoid masking. I kept dynamic range compression at a minimum to retain natural dynamics.

Although much of the spatiality of the track also was formed during recording, the mixing process was mostly concerned with organizing this spatiality, as well as further forming the recorded space. I concerned myself especially with the song's width and depth dimensions, as its height dimension was mostly formed through arrangement and the choice of microphone techniques. I will further discuss the distribution of sounds in each spatial dimension during mixing.

Width

It felt natural to place lead vocal, harmonica, and acoustic guitar together in the center of the stereo image to emphasize the role of the singer-songwriter as a main character. Backing vocals are panned a bit to each side, surrounding the lead vocal. The drum kit as a whole has some width, but its focus is in the center. Bass guitar and kick drum are placed in the center to achieve frequency balance across the stereo picture –placing one or both of these on a side could easily have made the sound-space appear side heavy, and thus unbalanced. The stereo placement of electric guitar, lap steel guitar and banjo was based on a hypothetical rehearsing room-situation whereby these instruments are placed on each side – electric guitar on the left and lap-steel and banjo on the right. The guitar solo at the end of the song is in the center, taking a larger focus than what would have been the case if it had been on either of the sides. This also contributes to clarifying that the guitar takes over the voice's role as lead instrument.

Distance and depth

Most of the drum kit is placed in the background of the sound-space. However, the kick and snare drums are somewhat more up front, emphasizing the rhythm. The bass guitar is quite loud, but because of its low vertical positioning, it does not appear as proximate. Guitars, both acoustic and electric, and lap-steel are in the middleground, whereas the banjo is just a bit behind the lead vocal, which is in the foreground. Backing vocals are at approximately the same distance as the banjo. Lead guitar is in the foreground. The distances of the sound-space are affected mostly by adjusting the volume of each element. However, some elements have been compressed, resulting in less dynamic variation. This applies to the kick drum, the bass guitar and the vocals. Regarding the lead vocal, the compression acts to bring it further to the front of the sound-space.

To form the depth of the recorded space, which also affects distance to some extent, I set up two reverbs: one studio ambience simulation and one plate reverb simulation. The

studio ambience simulation was meant to give even more "room sound" than what was achieved with the ambience mics, as well as a notion that all instruments were played in the same room. The plate reverb was used with a long reverb setting to make more abstract spaces, especially for vocals and lap steel guitar. Generous use of reverb on the lap steel makes its local space very deep. The lead vocal space is deeper than the space in which the drums and guitars are located in, but it does not sound as deep as the lap steel's space. For the guitar solo at the end of the song, I used a large amount of the studio ambience reverb. This gives the guitar its own space in the middle of the global sound-space. The use of two different reverbs in mix, as well as the recorded ambience and the amp spring reverb on the lap steel makes the global space a rather complicated one, consisting of several different spaces. I will discuss this further in the following analysis.

Time

The time dimension, which also concerns contrast, is mostly affected by the arrangement of the song, as well as the performance of each musician. During mixing, the greatest contribution to creating contrast was that I muted the guitar and lead vocal's ambience microphones when the band starts playing. Apart from this, most of the mix-related contrasts have to do with the spatialities of elements that are only present at particular times, such as the solo guitar.

Overall space-form

As will be further discussed in the analysis, I have formed the overall space-form of the song to consist of several different local spaces. These spaces include a "band space", which has a short reverb time and little difference between front and back; a "lap steel guitar space", which is very deep and narrow, but quite low due to the small frequency spectrum of the reverb; and a "vocal space", which is wide, and deeper than the band space, but not as deep as the lap steel guitar space. There is also a "solo guitar space" at the end of the song, which is of mostly the same proportions as the band space, but still it seems to be well within the boundaries of the band space. All these local spaces comprise what I call the global sound-space of the song.

4.3 Comparative analysis

With reference to the different theories of using sound-space as a production-analytic tool in chapters 2 and 3, I will now embark on a comparative analysis of the two different versions of "Die Young". The aim is to investigate the effects of using the notion of sound-space in production. As discussed above, this includes an understanding of the ways in which spatial organization of sound is able to affect the possible interpreted meanings of a song, through reference to other spatial situations or different sociocultural conventions. I believe that such an awareness has helped me achieve the result that I wanted. The focus of the analysis will be on the configuration of the recorded space, thus leaving out discussions of other musical parameters. The analysis will start with the latest recording of the song. As there are naturally many similarities between the two versions, I will, rather than comparing every point of both, only discuss the 2014-version where it clearly differs from the 2015-version. I will divide the analysis into four main parts – namely "Intro and first verses", "The band enters", "Last verse and chorus" and "Guitar solo" – based on changes of instrumentation and spatiality. Within these parts I will try to elicit the different aspects of how the sound-space is formed, including location, transtextual references, sonic markers, realism and corporeality. The differences between the spatial forming of the two sound-spaces will also be presented in a table at the end of this chapter (table 4.3.1).

Intro and first verses

The song starts with acoustic guitar and harmonica playing an intro, before the lead vocal takes over for the harmonica and starts singing the first verse. The 2015-version starts with acoustic guitar in the near middleground. Its sound is focused in the horizontal center, but is somewhat spread out in the room, a result of the use of stereo ambience mics. The guitar's frequency content is very rich, reaching towards the top of the sound-space while also having a satisfying amount of low end. The harmonica is placed slightly in front of the acoustic guitar. It has the same spatial environment, which is especially evident in the parts where its amplitude is higher. The lead vocal enters in the foreground. The voice, too, shares the spatial environment of the acoustic guitar, but there is also a longer reverb, as well as a slap-back echo, applied to it. The result of this is that the lead vocal exists in two spaces simultaneously. There are thus two main local spaces within the global sound-space, one in which both guitar and vocal are placed, and another, which is the deep vocal space. I will call the former the "performance space", as this is later inhabited by several of the other

instruments. It should not be confused with the global space, which denotes the totality of all the local spaces. The contrast of the vocal space against the performance space tends to give the long reverb a rather obvious role as an effect.

In the 2014-version of the song, this space is quite different. The acoustic guitar sound consists of two guitars. One is placed to the far left of the sound-space, while the other is spread out across the horizontal plane. The latter seems to lean towards the right, a result of the occupation of space on the left side by the other guitar. In other words, its sonic kinesphere is larger towards the right than to the left. Both guitars are in the near middleground. A considerable amount of low mid and low frequency content contribute to this placement. On the contrary, there is not much high frequency content in the acoustic guitars, a feature that gives them a relatively low position in the sound-space, compared to other contemporary acoustic guitar sounds. Their width and proximity combined makes it difficult to get a grasp on the depth of the virtual space in which they are located. The space sounds neither very dry nor very reverberant, and the guitars mask the sound of any reverberation. (In contrast, the narrower guitar sound of the 2015-version tends to emphasize the local space in which it is placed to a greater extent.) In the 2014-version it is the harmonica and lead vocal that draw attention to the space. The harmonica is placed centrally, a bit behind the guitars. Its higher position makes it audible, nonetheless. Also the lead vocal is placed farther away in the 2014-version than in the 2015-version. A long, but not very distinctive, reverb is applied to them. The reverb blends well with the local space of the guitars, and it gives the impression that the singer is placed at some distance from the microphone in a reverberant room.

The overall spaces in the intro of the two versions are thus quite differing. Whereas in the 2014-version the full stereo width is employed, giving a great sense of proximity to the guitars, but somewhat distant vocal and harmonica sounds, the 2015-version emphasizes the spatial environment in which the singer-songwriter is staged. These different types of staging interact with the thematic content of the song in quite distinct ways. The distant vocal sound of the 2014-version might be seen as an imitation of the girl in the story, being distant or faded. In other words, the story of the lyrics is emphasized. On the contrary, the 2015-version emphasizes a possible narrative related to the performance space, as it bears a clear resemblance to the acoustic singer-songwriter tradition of American folk music established by the likes of Woody Guthrie and Bob Dylan (for example, on Bob Dylan's early records, he plays alone, without additional musicians, sounding like he is in a small, non-reverberant, room). In that respect, the sound of a single, seemingly unprocessed, acoustic guitar functions

as a sonic marker, related to a notion of protest song, storytelling, the travelling folk singer, etc.[17] This also gives an opportunity for emphasizing the contrast between the "Bob Dylan"-figure and the electric band, which I overlooked in the 2014-version.

Furthermore, the combination of a full sounding (frequency-wise) acoustic guitar with a very present vocal gives a credible representation of the performance. In addition to the realistic performance space, the reverb that is applied to the voice gives another dimension to it, suggesting not only a spatial simultaneity, but also a narrative simultaneity: the short room-reverb suggests a sense of intimacy, while the longer reverb suggests loneliness or longing (for the lost girl).

After the first verse the lap steel guitar enters. In the case of the 2015-version, this is the first instrument that is placed to the side of the sound-space. It is at the far right, in the foreground. It has a lot of depth, though, having a long reverb applied to it that is different from those used on lead vocal and acoustic guitar. This results in a separate local space for the lap steel that reaches well beyond the boundaries of the already settled global space. The internal lap steel space is big and open, and thus it evokes Schafer's notion of hi-fi soundscapes in its clarity and sense of depth. The long reverb also places the lap steel in a long tradition of steel guitar sound and its related narratives: In the discussion of the use of reverberation on "Hawaiian" guitars, as steel guitars were called, in 1940s *hapa haole* music, Doyle draws on the Echo and Narcissus myth to suggest that "'[s]obbing' reverberant Hawaiian guitars lament the departure of the visitor, call vainly for his or her return" (Doyle 2005: 132). The sobbing refers here to the steel guitar's ability to slide between tones. An example of how the sound of steel guitar is still associated with sadness is Daniel Romano's introduction of his pedal steel guitar player in a live performance of the song "Hard On You": "That's mister Aaron Goldstein on the sad machine" (Music City Roots 2013). The reverberant steel guitar can thus be said to have been established as a sonic marker of sadness. This goes well with the story of "Die Young", in which the singer longs for his reunion with the girl that he lost. This sense of a local "lap steel space" also applies to the 2014-version to some extent. Here too, the lap steel is placed on the far right in a deep internal space, but one that is much more distant than in the 2015-version. It is thus not as clear sounding as in the 2015-version, and cannot be identified with a hi-fi soundscape.

[17] It must be mentioned that the acoustic guitar has been used in more commercial pop music, including what is called contemporary country music, but then often compressed hard. The typical "folk guitar" sound is characterized by natural dynamics.

The overall impression of the sound-space in the intro to the 2015-version is that each sound has a rather large sonic kinesphere, acting in an open hi-fi space with clear separation. When the lap steel enters, a greater sense of listener kinesphere is introduced, pulling towards the right. This emphasizes the coherence between extreme sound-space positions and corporeality. With the addition of a tambourine entering just after the lap steel, slightly to the left, the sound-space starts to encircle the listener. The characteristics of the performance space, suggestive of a small room, in combination with the instrumentation and musical style, lead my thoughts to the end of the documentary film *Heartworn Highways* (Leader & Szalapski 2005), where several musicians associated with the outlaw country movement are gathered at Guy Clark's house to drink and play music. There is a good atmosphere, much from it stemming from a group of friends hanging out together in a home. The small space of the 2015-version of "Die Young" has a similar homey feel to it. To this realistic space come the internal spaces of the lead vocal and the lap steel guitar as a contrast. There is no obvious dissonance between the different types of space, but rather a tension between the narratives related to the different spaces, in which one relates to the jovial setting of the performance, and the other to the tragic story of the lyrics. As suggested above, the 2014-version does not have this distinctive simultaneity of different spaces. There is simultaneity, but it is less defined when compared to the 2015-version.

The band enters

The 2015-version of "Die Young" has an obvious change of scene at the end of the second verse. As the singer ends the verse singing "oh my, oh my", the electric guitar strums a chord, followed by the drums playing a rhythm that is emphasized by the electric bass guitar. The introduction of these traditionally loud instruments "pushes" the acoustic guitar backwards to the far middleground. This is not due to a change of the acoustic guitar's volume, but more because of its lack of mid-tone energy compared to the other instruments. Its role is thus changed from the main harmonic instrument, to one of rhythmic subdivision and a "texture-filler". At the same time, it takes a clear role as a part of the band, in contrast to the 2014-version, where it is placed in near middleground throughout the song.

The drum kit in the 2015 version is at the far end of the sound-space, having the longest distance to the listener of the instruments. The cymbals are placed clearly in the background, taking up a lot of width, but not the entire stereo spectrum. Kick and snare drums, which are centralized on the horizontal plane, push forward into the middleground. This is also the case of the tom toms, although not to the same extent as the snare drum. The

tom toms are panned somewhat to the sides, giving a realistic representation of width of a drum kit in a small room. This spatial environment has the same characteristics as that of the acoustic guitar – a short reverb with distinct early reflections. However, the reverb is more prominent in the drum sound, something that is realistic, since the loud sound of the drums would, in an actual space, generate more reverb than an acoustic guitar. The drum kit seems thus to be placed in the same space, which I previously called the performance space. The lack of depth in this space makes each instrument, maybe except for cymbals, sound "reachable" from the listener's stance.

Located in the middleground of the sound-space of the 2015-version is the bass guitar. Since it resonates more with some frequencies than others, it tends to move slightly to the back and to the fore when different tones are played. The bass is the lowest placed element in the sound-space, with the kick drum just above it. Even though the bass is centralized on the horizontal plane, it occupies some width due to its low placement (as previously explained, low frequencies are more difficult to locate than high frequencies). In the far middleground, on the far left side, is the electric guitar. It shares the drum kit and acoustic guitar's reverb, giving the impression that the three instruments are played in the same room. The electric guitar balances the lap steel guitar on the far right, also being located at approximately the same middle height. While the wide-reaching bass guitar sound tends to reduce the listener kinesphere from the bottom of the sound-space, the distance in width to the electric guitar enlarges it towards the left.

On the last line of the second verse of the 2015-version, the first backing vocal comes in. This voice is in the horizontal center, somewhat behind the lead vocal, in near middleground. When the first chorus begins, a female backing vocal enters somewhat to the right, also in the near middleground, while the male backing vocal is spread out, reaching across the entire stereo spectrum. This is actually the same voice recorded three times. Two of the tracks are panned out to the sides and one is left in the middle. The slightly different dynamics in the three tracks makes the voice sound like it is moving randomly between the sides. This contributes to a difficulty in pinpointing the location of the backing vocals, while also providing an unrealistic width to the male voice. The result is that the listener's attention to the backing vocals alternates between the sides. We are at all times "pulled" towards the perceived location of the vocal. Moreover, there is a contrast between the realistic spatial environment of the voices and the unrealistic sideways movement. Also the distance of the voices is somewhat blurry. They seem to be pushed back by the band, but the present sibilants suggest some degree of proximity. Again, the spatial simultaneity of realistic and

unrealistic spaces tends to underline the contrast in the song between the upbeat music and the dreary story. Furthermore, the simultaneous presence and non-presence of the backing voices emphasizes the paralysis the protagonist in the lyrics finds himself in.

After the chorus there is a theme played on the electric guitar, which comes somewhat forward in the sound-space, just as musicians that are playing "un-plugged" have to come closer to the audience for their soloist parts to be heard. The theme ends in an alternate bass pattern, played by bass guitar and electric guitar, and joined by the drums, with bass drum and the low tom on downbeats and snare on upbeats. There is thus a clear upward and downward movement, amplifying a sense of corporeal verticality. After the theme, the electric guitar returns to its previous location.

With a diagonal mix established in the chorus, the bass guitar and snare drum being centralized, and electric guitar and lap steel panned, the sound-space of "Die Young" has several similarities with many other sound-spaces. This is the case of both versions. However, the 2015-version's space seems to have more separation, for example in the vertical plane, where bass guitar and kick drum are clearly separated, something that is not the case in the 2014-version. In the 2015-version, the band sounds like it is playing together in the same room. A possible exception would be the lap steel with its big reverb. However, this deep local space seems, in some peculiar way, to still be located within the smaller global space. This is, of course, an extremely unlikely case in "real life". But the naturalization of reverb effects in popular music recording (and live performance) makes it nevertheless somewhat realistic, in the sense that it does not prevent a sense of liveness in the recording. In the 2014 recording, in contrast, the distant location of the lap steel does exactly this. Its distance, in combination with its reverb, suggests that if the other instruments were playing in the same room, the sound would be extremely muddy with no sense of separation.

Summing up, in the first appearance of the full band in the 2015-version of "Die Young" the global sound-space consists of a small performance space, in which most of the instruments are located, and two larger spaces – the vocal space, which is deep and wide, and the lap steel space, which is deep, narrow and low. There is a clear sense of separation between the instruments, which leaves some holes in the texture, giving a sense of a quite large listener kinesphere. The short but prominent reverb of the performance space functions as a sonic marker of intimacy, as it is based on our experiences with small and cozy spaces, while the long reverbs used on vocal and lap steel are suggestive of longing or lament. This duality provides a contrast between two different narratives. The 2014-version has a similar location of sounds, with a few variations, but the perceived size of the sounds, combined with

the amount of reverb, results here in a dense space with few holes, if any. Still, there is some sense of sideways movement, provided by the very wide cymbals. Compared to the 2015-version, the listener kinesphere in the 2014 version is maybe larger in depth, but not in width. The spatial environment of the band, except for the lap steel, sounds rather anonymous, as the reverb is mostly masked by the instruments.

Last verse and chorus

In the third verse of the 2015-version, the steel guitar disappears in favor of a banjo entering quite far to the fore, on the far right side. As an acoustic instrument in front of a loud band, it has an almost unnatural proximity. Still, appearing as quite narrow, it does not occupy very much space. It rather functions as a kind of obbligato, a countermelody, against the lead vocal. The banjo shares the reverb of the performance space, but the ambience is quite prominent. This, in combination with the narrow sound and location quite high up in the sound-space, makes for an internal, local space around the banjo, giving a notion of a large sonic kinesphere for it. Furthermore, the proximate banjo reveals the use of technology, since it would have been impossible to achieve without the use of multiple microphones and/or multi-track recording. What adds an element of liveness to the banjo track is heard just before the second chorus, when something, probably the banjo player's chair, creaks. The banjo continues into the chorus, where also the lap steel comes back in, both playing throughout the song.

In the 2014-version it is the electric guitar that gives way for the banjo on the left side. The banjo is not as prominent as in the 2015-version, but still almost in the foreground and having the same musical role. In both versions it has its own space that tends to pull the listener towards it, because of its extreme horizontal position and surrounding holes. The lap steel continues on the right side. This, in combination with the loud and wide acoustic guitar, makes the verse more "floating" and less open sounding than in the 2015-version. The contrast between the verse and choruses is thus much smaller.

Guitar solo

The last part of "Die Young" is a guitar solo played over the chord progression of the verses. In the 2015-version the solo guitar replaces the electric guitar that has played up to that point. The rest of the band, except vocals, continues playing. The solo guitar has a much more distorted sound than the other electric guitar, and it is placed in the horizontal center. It takes place in a small internal space, a space with the same features as the performance space, but

its prominence suggests a smaller local space within the global space. It can be imagined as the rest of the band standing in a horseshoe formation to make room for this guitar space. The result of this is that even though the guitar is somewhat far from the listener (middleground), it can be clearly heard. Having the guitar at some distance tends to open up a space that can be imaginarily occupied by the listener, by enlarging his sonic kinesphere. There are no elements in the central foreground of the sound-space, so the listener is "invited" into the performance space.

The combination of the type of distortion and the roomy sound of the guitar gives associations to the alternative country band Wilco's debut album, *A.M.* (1995), especially the guitar solo on the track "Shouldn't Be Ashamed". There is thus a transtextual link to that song that emphasizes the genre context in which "Die Young" is set. The spatiality of the guitar resembles the usual characteristics of bad home recordings, and the bedroom/kitchen studio. This adds a sense of authenticity, sounding "un-produced", even though the difference between the local guitar space and the rest of the global sound-space suggests a constructed spatiality. The space in itself emphasizes, to some degree, a kind of joviality, inviting the listener to a domestic environment. On the other hand, it can also be associated with the frustration of being shut inside small rooms, mirroring the frustration of the singer. The interpretation of the solo guitar sound of the 2015-version thus provides an example of the difference between sonic markers and transtextuality. The reference to "Shouldn't Be Ashamed" is a transtextual one, since the guitar sound is typical neither for the genre nor the band, but is merely found in that one album. However, the guitar sound can also be interpreted as a sonic marker, as it clearly resembles common conceptions of DIY recordings. It could thus be called a sonic marker of roughness.

In the 2014-version, there is no distinctly separate guitar space during the solo. The solo guitar actually sounds like it is the same as the other electric guitar, but slightly more distorted. It is also panned to the same stereo location, thus doing neither more nor less to emphasize the text, other than clarifying its soloist role by coming farther to the fore of the sound-space.

Table 4.3.1 Main differences between spatial forming of the sound-space in two versions of "Die Young".

2014-version	2015-version
Intro and first verses	
Two acoustic guitars: One to the left and one spread out across the stereo plane. Both in foreground. Low placement. Spatial environment masked by guitars.	Acoustic guitar in near middleground and horizontal center. Narrow. Small spatial environment emphasized.
Harmonica is placed centrally, a bit behind the guitars. In large, not very distinct space.	Harmonica is slightly in front of the guitar, in same spatial environment.
Lead vocals further away than in 2015-version. Same space as harmonica.	Lead vocals in foreground. In same spatial environment as guitar and harmonica, but also in large space (spatial simultaneity).
Lap steel: Distant in a very deep space.	Lap steel: In foreground, in a deep space (separate local space). Reverb clearly standing out as effect.
Overall: Guitars and vocals uncoupled. Full stereo width employed. No clear sense of different spatialities.	Overall: Guitar and vocals in same position. Narrow sources, emphasized space ("singer-songwriter"). Distinct spatial simultaneity (small space, large space, lap steel space).
The band enters	
Acoustic guitar stays at the fore.	Acoustic guitar "pushed" back to far middleground. "Texture-filler".
Drums at far end of sound-space. Cymbals in background, very wide.	Drums at far end of local "performance space". Cymbals in background. Kick, snare and toms in middleground. Realistic width.
No clear separation in height between bass guitar and kick drum.	Bass guitar below kick drum, in middleground.
Electric guitar in far middleground, somewhat to the left.	Electric guitar in far middleground, far to the left.
Backing vocals: Male vocal to the left, in background.	Backing vocals: "Moving" from side to side, in background.
Overall: Different spatial environment seem to blend, no clear separation.	Overall: All instruments, except for vocals and lap steel, are in the same spatial environment (the "performance space"). The other spaces clearly stand out.
Last verse and chorus	
Electric guitar disappears in favor of banjo. The banjo is not as prominent as in 2014-version, but quite to the fore. In separate local space on far left.	Lap steel disappears in favor of banjo. Banjo is on the far right, in the foreground, in an internal local space.
Lap steel continues during verse and chorus. Contributes to a sense of "floating".	Lap steel returns in the chorus, while banjo continues.
Backing vocals. Male vocal left side, falsetto on right side. Background.	Backing vocals: same as previous chorus, but in far middleground.
Overall: Little contrast between verse and chorus.	Overall: Change of instrumentation provides contrast between verse and chorus. Change of texture.
Guitar solo	
Electric guitar "steps forward" from middleground to foreground, taking over the lead vocal's soloist role.	Electric guitar disappears in favor of a separate solo guitar. Also lead vocal disappears.
Overall: No further changes to the sound-space.	Overall: Band "makes room" for solo guitar in middleground (solo guitar space within the performance space).

This analysis has revealed a set of differences between the sound-spaces of the two versions of "Die Young", some of which I will repeat here. Firstly, the 2015-version highlights some of the contrasts of the song's arrangement to a larger extent than the 2014-version. This is most apparent when the instrumentation changes from just a singer-songwriter to a full band. Here, the contrast of two different performance conventions is emphasized in the 2015-version, whereas the 2014-version has a less noticeable transition. Furthermore, the different local spaces of the 2015-version seem more clearly defined than those of the 2014-version, setting up a clear distinction between the realistic performance space and the less realistic "effect" spaces (the lap steel space and the vocal space). These distinctions in space tend to emphasize the contrast between the intimacy of the performance and the "eternal questions" of the song's lyrics. The 2014-version's sound-space does not have such a clear separation between its local spaces, which seem to blend into each other.

Through this analysis I have demonstrated that an analytical model such as the sound-space can raise awareness of the effects of spatiality, and that this awareness can be helpful to producers and engineers for composing a mix. What's more, the sound-space can be used for purely analytical purposes, as it, in contrast to existing spatial approaches, includes a dimension of meaning, related to both the placement of the production in a larger context, and to transtextual references to the listener's previous experiences with space.

5 Concluding remarks

The aim of this thesis has been to elicit the ways in which a production-analytic tool, based on theories and analytical discussions of spatiality, can contribute to an improved understanding of how producers and engineers form a recording in terms of space. It has also demonstrated various ways in which a notion of recorded spaces can contextualize production, resulting in certain meanings and effects being bestowed the listener. Having discussed some of the concepts that I find relevant to the understanding of recorded spatiality from a record production perspective, as well as exemplifying these concepts through practice and an analysis of the outcome of that practice, I will now try to highlight the general insights that can be drawn from this thesis. Before concluding, however, I will briefly recapitulate the main points of the theoretical discussion, as well as discuss the outcome of the analysis.

Throughout this thesis I have discussed an array of different concepts related to the spatiality of recorded popular music. In chapter 2 I discussed Moore's sound-box and Danielsen's sound-room and how these work as analytical models. Their shared focus on the three-dimensionality of recorded music sheds light on important aspects of the complexities of popular music, thus contributing to Middleton's inquiry for studying popular music on its own terms. Being developed for the purpose of analyzing rather specific musical styles, the models differ somewhat. Moore applies the sound-box mainly to a discussion of 1970s and 1980s progressive rock, whereas Danielsen uses the concept of sound-room to analyze newer digitally mixed music, and lends thus more focus to the time domain. Another difference is that the sound-box is a tool in which sounds are "mapped", whereas the sound-room describes the recorded space as it appears to the listener.

I further discussed some of the ways in which these concepts can be expanded to apply to a production perspective. In that context I discussed existing spatial approaches to record production, including Mixerman's thoughts on spatiality in mixing, and Moylan's two-dimensional sound stage model. Mixerman's approach is based on his own experiences as a mixing engineer. Common to all models, also including Danielsen's and Moore's concepts, is that they do not, to any greater extent, discuss what the recorded spatiality can signify. I have suggested a few possible ways of including this to a spatial production-analytic model. For example, I discussed Doyle's study of echo and reverb in pre-stereo popular music, in which he revisits the myth of Echo and Narcissus to explain the origins of

echo as something "otherworldly". I also touched upon his discussion of Les Paul's recording techniques, exemplifying the correlation between spatiality and technological development (the proximate voice, overdubbing, and slapback echo). A notion of recorded spatiality that takes such considerations as these into account can explain why decisions made in production affect meanings in different ways.

Chapter 3 concerned the further discussion of the possible meanings that can be related to space. After a brief introduction to the perception of acoustics (spatial characteristics) through binaural hearing, I discussed the four dimensions of the recorded space, which are depth, width, height, and time. Width and, to some extent, depth can be said to be based on illusions that have become cultural conventions. I further discussed what Dockwray and Moore call the diagonal mix configuration, which, since the late 1960s, has been established as the normative mix configuration. The normative mix is a result of aesthetic and technological development. Being the normative way of distributing sounds across the stereo field, the diagonal mix has become a common denominator for nearly all popular music between 1970 and way into the 1990s, and for much of the music that is made today as well. This can be because it adapts very well to both stereo and mono playback, which was a challenge in the years of its development.

My concern in this thesis is what space can signify. As none of the other spatial concepts that I have discussed treat this aspect explicitly, I needed a new term encompassing both what the existing concepts deal with *and* a recording's meaning. Thus, I proposed the notion of a sound-space, which comprises the four-dimensional spatiotemporal organization of a recording, and the ways in which that organization affects the production of meanings in the recorded music. An important parameter of the sound-space is transtextuality, which suggests that a recorded space can contain references to previous recorded spaces, or to actual spaces. The latter presupposes a specific understanding of space as text. In that context I also discussed, based on Brøvig-Hanssen and Danielsen's research, the ways in which some spaces can appear as unrealistic or surreal, whereas other spaces either are more realistic, or are perceived as realistic because they have been naturalized. The introduction of Askerøi's term sonic markers further suggests a connection between the recorded space and larger narratives. This highlights the possibility, through the use of sound-space as tool, of emphasizing the interplay between the different narrative spheres of a song.

Lastly, I suggested a corporeal approach, based on Middleton's search for a theory of gesture, in which he discusses the corporeal aspects of Moore's sound-box. With the term kinesphere, derived from the Laban Movement Analysis, I argued for an understanding of

recorded spatiality based on imaginary movement. What I call the sonic kinesphere suggests that listeners, based on their previous experiences with being and moving in spaces, are involved in recorded spaces through transtextual references to actual spaces. Thus, the notion of sonic kinespheres provides an understanding of recorded sound to which listeners can easily refer (as everyone has had experiences of being and moving in space).

Chapter 4 consisted of a review of the practical work related to this thesis, as well as an analysis of the outcome of that work. Here the aim was to apply the notion of sound-space in a practical context. This approach, I concluded, is a form of practice-led research, as it integrates artistic practice in the research process. There are two outcomes of this that I will highlight here: Firstly, by applying the theoretical tools developed in this thesis on my recording it was possible to interpret relevant meanings of space in the 2014-version of "Die Young", although those meanings were not consciously produced. These could then be further developed in the 2015 version. Secondly, in the 2015-version, I based decisions regarding the forming of the sound-space (which included choice of microphone techniques and recording rooms, as well as the way the song was mixed) in an idea of how I wanted the spatiality to affect certain meanings in the song, in accordance with the theoretical work of this thesis. This resulted in a sound-space that substantiated the contrasting narratives of the song – those of the near and intimate, and of the remote and eternal.

The analysis of the differences between the two versions shows that theorizing and analyzing spatiality in record production is useful also for practice. After producing the 2014-version I became more aware of the potential for affecting meanings, and this awareness helped me achieve what I wanted in the production of the 2015-version. This was likely a result of my use of the sound-space as a tool for intentionally emphasizing certain meanings in the song.

During the recording and mixing of the 2015-version the sound-space model functioned as a tool with which I could leverage my decisions about forming the recorded space. It made me aware from the start of how the different spatialities of the recording could emphasize different meanings in the song. Becoming aware of how certain spatial characteristics can be associated with certain things (for example, intimacy, realness, nostalgia, or ability to move freely), I formed an idea of a sound-space, which I thought would go well with the song. This idea worked as a framework for my decisions regarding recording and mixing techniques. A similar spatiality, and thus similar interpretations of the result, could perhaps have been accomplished without the use of sound-space as a tool, but then as a result of more intuitive decisions. However, the sound-space tool made it easier to

achieve what I was after, and provided a plan for a spatial totality, from which I could form the local spatialities in a way that would evoke or emphasize the desired associative meanings.

Regarding the 2014-version of "Die Young", the analysis also reveals that, despite the minimal focus on transtextual relations in the production, the final sound-space substantiated certain parts of the song's narrative also in this version. This is probably because one tends to relate sound to meanings whether one is conscious about such connections or not.

Regarding the relationship between the two versions, the comparison revealed several differences which affect the listener's associations to each version as well as the historic narrative in which the song is set. The overall sound-space of the 2014-version is characterized by its lack of separation and contrast, with the different spatialities seeming to blend into each other. The location of the lap steel guitar far behind the rest of the band foregrounds the recording technology, as such a placement would not have been possible in an actual spatial situation. In the 2015-version, the different spaces are more clearly separated. Additionally, the sounds also within these spaces seem more separated than in the 2014-version. This provides a sense of openness. Furthermore, the separation, together with little difference between the distances of the different instruments, as well as a short reverb, suggests that the performance takes place in a small "everyday" room. Lastly, the long reverb that is used on the lap steel in the 2015-version does not push the instrument unnaturally far back in the sound-space, but it stands out as an effect (which has been naturalized after over 70 years of use in that particular context). This lap-steel sound provides associations to a larger narrative of lament. Through juxtaposing these two spaces within the global sound-space, I thus managed to emphasize an important aspect of the song: the duality that comes from the contrast between the up-tempo, "joyful" melody and the sad content of the lyrics. Overall, the sound-space of the 2015-version appears as more realistic than that of the 2014-version, which was also desired, given this thematic orientation.

The use of sound-space as a guiding analytical tool in the production process further elicited the different ways in which the listener is "invited into" the sound-space in terms of their corporeal experience. This last point is heard especially during the guitar solo of the 2015-version, where, as previously described, the electric guitar sounds like it is somewhat backgrounded in the sound-space while other instruments seem to make a horseshoe formation around it. In the 2014-version, the only change to the spatiality during the guitar solo is that the electric guitar moves a bit forward. What's more, the "openness" of the 2015-version's space, in contrast to the much more dense space of the 2014-version, tends to afford

a sense of (imaginary) bodily movement, as attention shifts between different points in the sound-space.

Summing up, the space is formed differently in the two versions to emphasize different aspects of the song. There are potentially many narratives going on in a recording, and regarding the two versions of "Die Young", listeners will probably hear the song differently and in accordance with their experiences and needs. However, one could say that there are two different "main narratives" highlighted in the two versions. In the 2014-version there is the narrative of the song's lyrics, coupled with a sound-space characterized by a lack of contrasts, suggesting some kind of "floating" state of mind, related to the protagonist's lament. The 2015-version, on the other hand, emphasizes a narrative related to the performance, in accordance with the genre traditions to which the song is related, as well as the way the band wants to appear to their audience. At the same time the less realistic local spaces, for example the reverberant vocal and lap steel spaces, suggest a larger narrative related to the eternal longing of the story's protagonist (which is also typical for the genre). The combination of a spatial simultaneity (different spatial situations acting simultaneously) and multiple spatial settings (changing of the spatial settings during the song), and the contrasts these brings to the sound-space, tend thus to emphasize the contrast in the song between everyday intimacy and eternal questions of life and death. All this supports the conclusion that the theoretical studies and the analyses in this thesis have led to a better understanding of these aspects of the two versions of the song, as well as improving the actual forming of the recorded spaces in a way that gives them even better chances of reaching the listener.

The aim for this thesis has been to investigate the possible meanings of the spatiality of recorded popular music. As a model to investigate this from a production perspective, I have developed the notion of the sound-space, building on existing spatial models for popular music analysis, while to a greater extent including the possible associations (for the listener) linked to the spatial forming of a recording. Sound-space can thus be regarded as a contribution to the discussion of the complexities of popular music, which sets this thesis in a tradition of popular music studies, and particularly the more recent field of record production studies. The analysis of the practical work related to this thesis demonstrates that the sound-space can be a useful tool not only for working with spatiality, but also for including an analytic perspective to record production, in which the producer and/or engineer takes the role of a critical listener. Furthermore, the sound-space can also be a tool for raising

awareness of how the spatiality of a recording contributes to the forming of text and narrative, as well as the corporeal involvement of the listener.

A challenge of actively applying the sound-space as a tool in record production could be that such considerations are not necessarily prioritized in an actual recording situation. A range of unexpected situations can occur during a studio session (for example technical difficulties, artists not agreeing to the approach, etc.), which can make it difficult to conduct the idea of a sound-space as intended. Nevertheless, it is my belief that an idea of the forming a sound-space, and its possible meanings, can be a good starting point when one embarks on a record production, as it can function as a reference for critical listening in the studio.

As suggested there are several possible directions in which one can take a thesis like this one, and the delimitations I have set have necessarily excluded many possible approaches to the subject. A further investigation of the corporeality aspect of recorded spaces would be particularly interesting, as it is relevant to how listeners perceive music. A possible approach could be empirical research on the correlation between sound-space and bodily movement through motion capture analysis.

I also discussed challenges related to the investigation of my own practical work, one of the potential problems being that the findings might be highly saturated by my own opinions and idiosyncrasies. It could therefore be interesting to investigate the ways in which other producers and engineers relate to spatiality, and whether such an investigation could substantiate, or refute, what I have argued in this thesis. This could have involved qualitative interviewing of a selection of producers and engineers, and possibly also observations of their working methods. To see the effects of how different producers and engineers form the sound-space it could also be useful to analyze different music by the same artist working with different producers/engineers. Lastly, it would be interesting to look at the ways in which (if any) studio musicians can make use of the sound-space when choosing sounds and style of playing in the studio.

Bibliography

Askerøi, E. (2005). *Man/Machine: Mot en åpning av soundbegrepet* (master's thesis). University of Oslo, Oslo.

Askerøi, E. (2013). *Reading Pop Production: Sonic Markers and Musical Identity* (doctoral dissertation). University of Agder, Kristiansand.

Barry, E. (2015). Mono in the Stereo Age, in P. Théberge, K. Devine, & T. Everrett (eds.), *Living Stereo: Histories and Cultures of Multichannel Sound* (pp. 125-146). New York, NY: Bloomsbury.

Brolinson, P. E., & Larsen, H. (1981). *Rock…: aspekter på industri, electronik & sound*. Falköping: Esselte Studium.

Brøvig-Hanssen, R. (2013). *Music in Bits and Bits of Music: Signatures of Digital Mediation in Popular Music Recordings* (doctoral dissertation). University of Oslo, Oslo.

Brøvig-Hanssen, R., & Danielsen, A. (2013). The Naturalised and the Surreal: Changes in the Perception of Popular Music Sound, *Organised Sound*, 18/1, 71-80.

Byrne, D. (2012). *How Music Works*. Edinburgh: Canongate.

Camilleri, L. (2010). Shaping Sounds, Shaping Spaces, *Popular Music*, 29/2, 199-211.

Chion, M. (1994). *Audio-Vision: Sound on Screen*. (C. Gorbman, trans.). New York, NY: Colombia University Press. (Original work published 1990).

Clarke, E. F. (2010). Rhythm/Body/Motion: Tricky's Contradictory Dance Music, in A. Danielsen (ed.), *Musical Rhythm in the Age of Digital Reproduction* (pp. 105-120). Farnham: Ashgate.

Crane, L. (2015). Glyn Johns, *Tape Op: The Creative Music Recording Magazine*, 109, 46-68.

Danielsen, A. (1993). *'My name is Prince' – en studie i Diamonds and Pearls* (master's thesis). University of Oslo, Oslo

Dockwray, R., & Moore, A. F. (2010). Configuring the Sound-Box 1965-1972, *Popular Music*, 29/2, 181-197.

Doyle, P. (2005). *Echo and Reverb: Fabricating Space in Popular Music Recording 1900-1960*. Middletown, CT: Wesleyan University Press.

Frith, S. (1986). Art Versus Technology: The Strange Case of Popular Music, *Media, Culture & Society*, 8/3, 263-279.

Frith, S., & Zagorski-Thomas, S. (2012). Introduction, in *The Art of Record Production: An Introductory Reader for a New Academic Field* (pp. 1-9). Farnham: Ashgate.

Genette, G. (1997). *Palimpsests: Literature in the Second Degree*. Lincoln, NE: University of Nebraska Press.

Hartmann, W. H. (2013). *Principles of Musical Acoustics*. New York, NY: Springer.

Izhaki, R. (2012). *Mixing Audio: Concepts, Practices and Tools*, 2nd ed. Oxford: Focal Press.

Jensenius, A. R., Wanderley, M. M., Godøy, R. I., & Leman, M. (2010). Musical Gestures: Concepts and Methods in Research, in R. I. Godøy & M. Leman (eds.), *Musical Gestures: Sound, Movement, and Meaning* (pp. 12-35). New York, NY: Routledge.

Laban, Rudolf (1966). *Choreutics*. London: MacDonald and Evans.

Lacasse, S. (2007). Intertextuality and Hypertextuality in Recorded Popular Music, in A. F. Moore (ed.), *Critical Essays in Popular Musicology* (pp. 147-170). Aldershot: Ashgate.

Mäkelä, M, Nimkulrat, N, Dash, D. P., & Nsenga, F.-X. (2011). On Reflecting and Making in Artistic Research, *Journal of Research Practice*, 7/1, Article E1.

Middleton, R. (1990). *Studying Popular Music*, Milton Keynes: Open University Press.

Middleton, R. (1993). Popular Music Analysis and Musicology: Bridging the Gap, *Popular Music*, 12/2, 177-190.
Mixerman (2010). *Zen and the Art of Mixing*. Milwaukee, WI: Hal Leonard Books.
Merriam-Webster (2016). Space, *Merriam-Webster* [online]. Retrieved from: http://www.merriam-webster.com/dictionary/space. Date accessed: April 19, 2016.
Meyer, J. M. (2009). *Acoustics and the Performance of Music: Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instrument Makers*, 5th ed. New York, NY: Springer.
Moore, A. F. (2001). *Rock: The Primary Text*, 2nd ed. Aldershot: Ashgate.
Moore, A. F. (2007). Introduction, in *Critical Essays in Popular Musicology* (pp. ix-xxii). Aldershot: Ashgate.
Moore, A. F. (2012). *Song Means: Analysing and Interpreting Recorded Popular Song*. Farnham: Ashgate.
Moylan, W. (2012). Considering Space in Recorded Music, in S. Frith, & S. Zagorski-Thomas (eds.), *The Art of Record Production: An Introductory Reader for a New Academic Field* (pp. 163-188). Farnham: Ashgate.
Moylan, W. (2015). *Understanding and Crafting the Mix: The Art of Recording*, 3rd ed. Burlington, MA: Focal Press.
Oxford Dictionaries (2015). Corporeal, *Oxford Dictionaries* [online]. Retrieved from: http://www.oxforddictionaries.com/definition/english/corporeal. Date accessed: October 23, 2015.
Oxford Dictionaries (2016). Fidelity, *Oxford Dictionaries* [online]. Retrieved from: http://www.oxforddictionaries.com/definition/english/fidelity. Date accessed: April 19, 2016.
Rossing, T. D., Moore F. R., & Wheeler, P. A. (2002). *The Science of Sound*, 3rd ed. San Francisco, CA: Addison Wesley.
Rumsey, F. (2001). *Spatial Audio*. Oxford: Focal Press.
Schaeffer, P. (2004). Acousmatics, in C. Cox, & D. Warner (eds.), *Audio Culture: Readings in Modern Music* (pp, 76-81). New York, NY: Continuum. (Original work published in 1966).
Schafer, R. M. (1994). *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester, VT: Destiny Books.
Smalley, D. (1997). Spectromorphology: Explaining Sound-Shapes, *Organised Sound*, 2/2, 107-126.
Smalley, D. (2007). Space-Form and the Acousmatic Image, *Organised Sound*, 12/1, 35-58.
Théberge, P., Devine, K., & Everrett, T. (2015). Introduction: Living Stereo, in *Living Stereo: Histories and Cultures of Multichannel Sound* (pp. 1-34). New York, NY: Bloomsbury.
Tingen, P. (2005). Steve Albini: Sound Engineer Extraordinaire, *Sound on Sound* [online]. Retrieved from: https://www.soundonsound.com/sos/sep05/articles/albini.htm. Date accessed: April 19, 2016.
Walther-Hansen, M. (2015). Sound Events, Spatiality and Diegesis: The Creation of Sonic Narratives in Music Productions, *Danish Musicology Online*, special edition on sound and music production, 29-46.
White, G. D., & Louie, G. J. (2005). *The Audio Dictionary*, 3rd ed. Seattle, WA: University of Washington Press.
Zagorski-Thomas, S. (2010). The Stadium in Your Bedroom: Functional Staging, Authenticity and the Audience-Led Aesthetic in Record Production, *Popular Music*, 29/2, 251-266.

Zak III, A. J. (2001). *The Poetics of Rock: Cutting Tracks, Making Records*. Berkeley, CA: University of California Press.
Zeiner-Henriksen, H. T. (2010a). Moved by the Groove: Bass Drum Sounds and Body Movements in Electronic Dance Music, in A. Danielsen (ed.), *Musical Rhythm in the Age of Digital Reproduction* (pp. 121-139). Farnham: Ashgate.
Zeiner-Henriksen, H. T. (2010b). *The "PoumTchak" Pattern: Correspondences Between Rhythm, Sound, and Movement in Electronic Dance Music* (doctoral dissertation). University of Oslo, Oslo.

Videography

Leader, G. (Producer), & Szalapski, J. (Director). (2005). *Heartworn Highways* [DVD]. United States: Navarre Corporation.
Music City Roots (2013). Daniel Romano "Hard on You", *YouTube* [online video clip]. Retrieved from: https://www.youtube.com/watch?v=6uZtCCm-xnA. Date accessed: April 19, 2016.

Discography

Adams, Ryan (2011). *Ashes & Fire*. Pax Americana Recording Company.
Daughter (2013). *If You Leave*. 4AD Ltd.
Dingus (2015). Die Young [digital single]. Safe & Sound Recordings.
Dingus (forthcoming August 2016). *Dingus*. Safe & Sound Recordings.
Jones, Sharon & The Dap Kings (2014). *Give the People What They Want*. Daptone.
Red Hot Chili Peppers (1999). *Californication*. Warner Bros. Records.
Romano, Daniel (2011). *Sleep Beneath the Willow*. You've Changed Records.
Talking Heads (1978). *More Songs About Buildings and Food…* Sire Records.
Who, The (1971). *Who's Next*. Track Records.
Wilco (1995). *A.M.* Reprise Records.
Wilco (1999). *Summerteeth*. Reprise Records.
Williams, Lucinda (1998). *Car Wheels On a Gravel Road*. Mercury.

Appendix: Recordings

CD

Track 1: Dingus (2015). “Die Young”. Recorded 2014 (version 1).

Track 2: Dingus (forthcoming August 2016). “Die Young”. Recorded 2015 (version 2).