Space in Electroacoustic Music: Composition, Performance and Perception of Musical Space

Frank Ekeberg Henriksen
Doctor of Philosophy
City University
Department of Music

July 2002

Contents

Introduction **15**

Approach 16

Musical space 17

Outline 20

I Musical space – perceptual basis **23**

1 Intrinsic space **25**

1.1 Auditory perception of individual sounds and events 25

1.1.1 Event recognition 25

1.1.2 Source recognition 28

1.2 Attributes of intrinsic space 29

1.2.1 Magnitude 29

1.2.2 Shape 31

1.2.3 Density 32

2 Extrinsic space **33**

2.1 Spatial hearing and acousmatic listening 33

2.2 Attributes of extrinsic space 35

2.2.1 Direction 35

2.2.2 Distance 38

2.2.3 Movement 39

2.3 Categories of sounds based on extrinsic space 40

3 Spectral space **43**

3.1 Theoretical background 43

3.1.1 Perception theories 43

3.1.2 Spectromorphological theory 45

3.2 Conclusion 47

II Musical space – composition, performance, perception **49**

4 Composed space **51**

4.1 Space as compositional element 52

4.1.1 Structural functions 53

4.1.2 Space on a macro-level: Spatio-structural orientation 54

4.1.3 Space on a micro-level: Spatio-structural functions 59

4.1.4 Significance of spatial detail 62

4.2 Stereo and surround-sound composition 62

4.3 Case study: Spatio-structural analysis of *Intra* 63

5 Listening space **71**

5.1 Sound and the acoustic environment 72

5.1.1 Case study: *Lorry Red Lorry Yellow*—a sound installation 75

5.2 Listening in private and public spaces 78

5.2.1 Private space 78

5.2.2 Public space 79

5.3 Sound systems 80

5.3.1 Headphone listening 80

5.3.2 Stereophony 82

5.3.3 Surround-sound systems 84

5.4 Case study: *Terra Incognita* 90

5.5 Sound diffusion 97

5.5.1 The sound diffusion system 97

5.5.2 Diffusion strategies 99

5.5.3 Case study: Diffusing *(dis)integration* 104

6 Perceived space **111**

6.1 Space as communication 112

6.1.1 Proxemics 112

6.1.2 Territoriality 114

6.2 Space as an element of expression and communication in electroacoustic music 116

6.2.1 Spatio-musical expression 116

6.2.2 Spatio-musical experience 118

Conclusions **123**

Bibliography **127**

Appendix I

Programme notes **135**

(dis)integration (1998) 135

Ebb (1998) 135

Intra (1999) 136

Itinera (1999-2001) 137

Terra Incognita (2001) 137

Appendix II

Programme notes for *Lorry Red Lorry Yellow* **139**

Lorry Red Lorry Yellow (2000) 139

Appendix III

Spatio-structural analysis of *Intra* **141**

Appendix IV

Diffusion score for *(dis)integration* **155**

List of Figures

0.1 Overview of the levels of musical space. 18

1.1 Intrinsic space. 29

2.1 Head-related system of spherical coordinates (From Blauert (1997) Figure 1.4, p 14. Courtesy of MIT Press.) 34
2.2 Extrinsic space. 35

4.1 Composed space. 52
4.2 Time and amplitude view of *Intra*. 64

5.1 *Lorry Red Lorry Yellow*. Covered-up loudspeaker, north-east wall. 76
5.2 *Lorry Red Lorry Yellow*. North wall. 77
5.3 *Lorry Red Lorry Yellow*. South wall. 77
5.4 Stereophonic loudspeaker array. 82
5.5 5.1 surround-sound configuration. 85
5.6 Monitor layout for the composition of *Terra Incognita*. 92
5.7 Time and amplitude view of *Terra Incognita*. 93
5.8 Assumed loudspeaker configuration for the diffusion of *(dis)integration*. . . 105

6.1 Overview of space in electroacoustic music. 124

List of Works

(dis)integration (1998). Two-channel stereo. Duration 9:54.

Ebb (1998). Two-channel stereo. Duration 10:43.

Intra (1999). Two-channel stereo. Duration 12:35.

Itinera (1999-2001). Two-channel stereo. Duration 18:10.
Part I: *Trans/Trance*. Duration 8:28.
Part II: *kernelMotion*. Duration 9:42.

Lorry Red Lorry Yellow (2000). Sound installation, audio excerpt. Duration 71:50.

Terra Incognita (2001). Ambisonic surround-sound. Duration 13:56.

Terra Incognita (2001). Two-channel stereo version. Duration 13:56.

Acknowledgements

This thesis presents the results of doctoral research conducted at the Department of Music at City University over the four year period 1997-2001.

I would like to thank my supervisor, Professor Denis Smalley, for sharing his insights into musical thinking and composition and guiding me through this project. I would also like to thank Dr. Simon Emmerson for his general availability and helpfulness.

Thanks to my parents, Grete and Tore Henriksen, for offering support and help whenever needed. Thanks most of all to my partner, Betsy Schneider, whose intelligent remarks, words of encouragement and emotional support helped me through all the ups and downs that accompanied these years of intense study.

This work would not have been completed had I not had financial support. I would like to thank the Research Council of Norway for awarding me a Doctoral Fellowship from January 1999 to June 2001. Without it the work would not have been of the same scope. I would also like to thank the Worshipful Company of Saddlers, whose Robert Kitchin Award I received for three years, 1997-2000. Many thanks also to O. Kavli and Knut Kavli's Fund for awarding me a research grant in 1998.

Abstract

This thesis concerns *space* as an essential element of expression and communication in electroacoustic music. It shows that *musical space* is a complex term which refers to many different aspects of composition, performance and perception of electroacoustic music. It is argued that space is a compound musical element which can be integrated into the compositional structure to a degree where space becomes the primary carrier of meaning in the work, and that the creation and interpretation of this meaning is a result of learned cultural aspects of interpersonal communication in terms of personal space and territoriality. Furthermore, the close relationship between electroacoustic music composition and technology is acknowledged, and the influence of available technology on aesthetic choices and decision making with regard to spatial composition and performance is taken into consideration.

The structure for the investigation is based on a model of musical space comprising three basic levels: 1) spatial properties of individual sounds in terms of *intrinsic space*, *extrinsic space* and *spectral space*, 2) the spatial arrangement of individual sounds and events into a *composed space* which is played in, and becomes affected by, the *listening space*, and 3) the *perceived space*, which constitutes the listening experience of the combination of composed space and listening space. A framework for describing and analysing spatial elements in electroacoustic composition is proposed.

The discussion and findings are largely based on my experience as a listener, composer and performer of electroacoustic music, and in addition finds support in research on auditory perception, particularly Jens Blauert's work on spatial hearing and Albert Bregman's *auditory scene theory*, as well as Denis Smalley's *spectromorphological theory*, James Tenney's writings on perception-based music listening and analysis, and Edward T. Hall's investigations into space as an element of non-verbal communication.

Introduction

Space is an essential dimension of human experience. In our daily lives we move around in relation to objects and other people and hear sounds in a multi-dimensional sound field. The significance of any given sound depends on where we hear it coming from. How we interpret distance cues and directional cues is essential to our survival and to our orientation in our surroundings. In its broadest sense space permeates every aspect of our life.

The implementation of space in musical composition—the spatial arrangement of sound materials in real and virtual spaces—is the manifestation of space as a primal element of expression and communication in music. Music cannot exist independent of space, but awareness of space as a fundamental musical element is nevertheless limited among the general music audience and even among many music practitioners. This is precisely due to its ubiquitous nature. The spatial influence is present in all listening, and space therefore tends to be taken for granted. It is only paid particular attention when spatial aspects of the music listening situation are something out of the ordinary. Surround sound and multi-loudspeaker sound diffusion are examples of spatial features which are unique to electroacoustic music. The inclusion and manipulation of recordings of sonic environments and environmental sounds are equally striking phenomena of the electroacoustic genre in which space is central.

How the meaning of spatial information in music is understood is determined by the deeper cultural knowledge and experience of spatial communication from everyday life. Patterns of interpersonal communication, the experience of rural and urban life, the architectural environment in which we live as well as how space is represented in language are all contributing factors in the shaping of our ability to interpret spatial cues from our surroundings. This knowledge is so deep-rooted that the individual remains mostly unaware of its influence on the interpretation of sensory input, and is often not aware that important spatial information is present at all. Nevertheless, spatial cues are constantly being processed in our encounters with our surroundings, and play a vital part in all our activities.

Approach

My discussion of space in electroacoustic music is from two perspectives. On one hand I discuss space from the point-of-view of a composer and performer of electroacoustic music. Spatial considerations in the composition process—the choice and arrangement of sound material in terms of spatial characteristics—are fundamental to the creation of the musical work. The composer's awareness (or lack of awareness) of space as a significant communicational factor in music is evident in the way spatial elements are integrated into the structure of the work. The size and layout of virtual spaces, the use of distance and movement, the integration of familiar environmental cues and the nature of spatial interrelations among the sound materials are powerful and flexible tools for musical expression.

Auditory space cannot, however, be perceived without anything in it. Successful spatial composition is therefore not possible without considering how the spectromorphological and associative qualities of the sound material affect the perception of musical space as a whole. A composition is most powerful when the combined forces of all the musical elements are used consciously and with great care. Thus, a deep awareness of the intimate relationship between space and sound is necessary in the composition process.

The life of the musical work is not fulfilled unless it is made available for others to hear. Concert performance, radio broadcast and record distribution represent the mediating link between composer and audience. When the finished work has left the composition studio, it is likely to be played in a variety of listening environments on a range of different types of playback systems. The spatial experience of music is quite different in a concert hall with a large multi-channel loudspeaker system surrounding the audience compared to solitary listening on headphones or on a two-channel stereo system at home. The potential of the sound system and the room in terms of flexibility to be adjusted to spatial elements composed into the work are of concern to the composer and performer of electroacoustic music with regard to how these elements may come across to the listener.

The listening circumstances are therefore central aspects in a discourse on musical space, and represent the other perspective of my discussion. Listening is involved in all stages of composition, performance and appreciation of music, but takes on a different form depending on where in the musical-communicational chain it takes place. The composer listens repeatedly and with great scrutiny on several levels during the composition process, and possesses detailed knowledge about the sound material and its organisation that is not possible for other listeners to gain. In contrast, the trained listener may be able to hear structural connections which can only be revealed by someone who has a certain distance with regard to the work, while the casual listener may hear external references which are most apparent to someone without great technical knowledge of the genre. Training and experience in listening to electroacoustic music determine the ability to reveal connections and aesthetic significance of space on various structural levels in the work. Regardless of the listener's musical background, however, it is the cultural knowledge of spatial communication which guides the interpretation of the spatio-musical elements perceived in listening and the spatial aspects of the listening situation as a whole. All the senses take in and process spatial information, creating a state of mind that varies with different surroundings and situations.

Musical space

A main objective for this thesis is to reveal and discuss the many facets of space in electroacoustic music. The terms "space" and "musical space" are often used in musical-theoretical discourse as if there is *one* space whose definition is commonly understood[1]. However, as I will show, "space" is a very complex term in the context of electroacoustic music, where it refers to many different things that for the most part can be discussed as if they were separate entities, but which in reality are intertwined and cannot be experienced in isolation.

Framework for investigation

Before going into a detailed discussion of the various components of space in electroacoustic music, it is helpful to outline what these components are and how they connect. As a framework for my investigation I have arrived at a basic model of musical space comprising three levels:

1. On the highest level is the listener's interpretation of the musical space of the composition as experienced in the listening situation. This is what I call *perceived space*.

2. On the middle level are the spatial characteristics of the musical composition itself—the *composed space*[2]—and the influence of the *listening space* upon the musical composition during playback and performance.

3. On the lowest level are spatial characteristics of the individual sounds that make up the composition. These characteristics are discussed as *intrinsic space*, *spectral space* and *extrinsic space*.

Figure 0.1 outlines the components of musical space.

[1] Harley (1994) gives a comprehensive review of the meaning and history of the concept of space in music and related areas.

[2] *Composed space* is a term coined by Smalley (1991). It is derived from Chion's (1988) notion of *espace interne*, referring to the internal organisation of the compositional elements in the work as they are recorded on to a fixed medium.

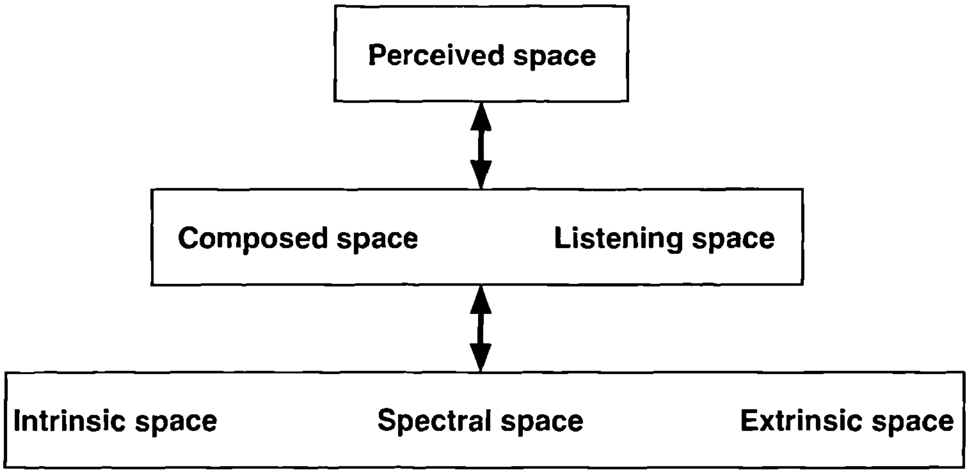


Figure 0.1: Overview of the levels of musical space.

Individual sounds

The lowest level of my model of musical space concerns individual sounds and sound events. My discussion of these is largely based on perception theory and theories on spatial hearing. The assumption is that these spatial properties can be perceived by listeners regardless of musical training as they are studied outside any musical context. But, as Deutsch (1982) and Bregman (1994: 703-4) report, there are variations as to how listeners interpret cues for segregating sounds, even on the most primitive levels of perception. This means that in ambiguous situations different listeners hear different things. In my discussion, however, I choose to simplify the matter slightly, and refer to perception in general terms.

Intrinsic space has to do with the sound *as* space. It deals with spatial components inherent in the individual sound in terms of perceived size, density and spatial shaping over the course of its existence. Intrinsic space is a somewhat abstract entity, and may not be immediately obvious to all listeners, but is, nevertheless, important and influential on the spatial experience. *Spectral space* is not a space in the acoustic meaning of the word, but is based on a psychological sense of height attributed to sounds based on pitch or prominent timbral components. Sounds are perceived as being high or low relative to the pitch or timbre of some reference sound. *Extrinsic space* has to do with the sound *in* space. It refers to the space surrounding the sound as a result of its spatial behaviour and interaction with the environment in which it exists. Extrinsic space is perceived in terms of movement, distance and direction of sounds. Location in physical and/or virtual space is based on extrinsic space.

Composed space and listening space

On the next level are composition and performance. The *composed space* is the spatial arrangement of individual sounds and events into a musical context. Composed space is made up of the intrinsic, extrinsic and spectral spaces of the chosen sounds and their spatial interrelations. My investigation into composed space focuses on structural functions of space in the work, both in terms of the spatial information in the sound material itself and in terms of virtual space and spatial discourse made up by the arrangement of the sound material.

The *listening space* is the spatial setting in which the work is heard. This comprises the physical listening environment with all its acoustical characteristics, the type and arrangement of the sound system as well as the listening position relative to the loudspeakers and to the physical boundaries of the listening environment. There is a substantial variety of physical spaces and sound systems in use for playback of electroacoustic music. The differences in spatial potential between headphones, two-channel stereo and the variety of multichannel and surround-sound solutions used for playback and performance affect the music in different ways. Unless the work is created for a particular space and sound system known to the composer, it is difficult to predict with any high accuracy how the spatial elements of the work will come across in listening spaces outside the composition studio. An awareness of these factors is therefore crucial for composers and performers of electroacoustic music.

Perceived space

The final level in my model of musical space has to do with the perception and experience of space in a broader context of music listening and experience. *Perceived space* is based on the interaction between listening space and composed space as experienced by the listener. Here I refer to *aesthetic* perception as opposed to the low-level everyday perception that decides how we see and hear the world and recognise objects and situations in our daily lives. Aesthetic perception has to do with how something appears in an aesthetic context, and determines the aesthetic experience of a work of art. In the specific case of spatial listening in music, it has to do with the perception and interpretation of space as an element of musical expression and communication, and with the integration of space with other musical elements into an aesthetic whole.

Musical training and experience in listening to electroacoustic music can increase one's ability to connect and interpret the structural relevance of spatial information in the work. Psychological and sociological influences regarding personal space and territoriality, both as features of the electroacoustic work in performance and of the listening situation as a whole, come into play as key components in the perception of space as a communicational element in electroacoustic music.

Outline

The thesis is divided into two parts. The first part concerns the three elements on the lowest level of my model of musical space—that is, the spatial features of individual sounds and events. This provides a perceptual basis for the discussion in the second part, which is concerned with the composition, performance and perception of spatial characteristics of electroacoustic music.

Chapter 1 is devoted to the concept of intrinsic space. It includes an overview of the perceptual processes involved in recognising and segregating sounds and sound events, mostly based on Bregman's auditory scene theory, before discussing in detail the components of intrinsic space. Chapter 2 deals with extrinsic space. An overview of relevant aspects of spatial perception and localisation is included, with particular emphasis on Blauert's work on spatial hearing. The chapter concludes with a categorisation of sounds based on extrinsic space. Spectral space is the subject of Chapter 3. The discussion here draws on research concerning the perception of elevation of sounds based on spectral components, where Blauert's work again is central. In addition, Smalley's notion of spectral space, which forms a part of his spectromorphological theory, is reviewed. In conclusion it is argued that the experience of spectral space is based on a combination of innate factors and learned aspects in listening.

While the first part of the dissertation concentrates on spatial features of the individual sound, the second part focuses on space in the context of musical composition, performance and listening. Chapter 4 deals with the composed space, and discusses space as a structural element in music. A *spatio-structural* approach to describing and analysing space in electroacoustic music is proposed, and is demonstrated in an analysis of my work *Intra*. The listening space is treated in Chapter 5, where I first review some relevant issues of acoustics before considering different types of sound systems and looking into the spatial characteristics of private and public listening environments. This chapter also includes a section on sound diffusion, where I outline eight categories of spatial characteristics of electroacoustic works which are likely to be affected by the listening space and can be controlled in diffusion. Finally, in Chapter 6 perceived space is examined. This chapter has a special focus on the socio-psychological notions of personal space and territoriality and their influence on the musical experience. In particular, Hall's work on *proxemics* provides a background for the discussion. The chapter concludes that the understanding and interpretation of spatio-structural elements in music and the effect of spatial aspects of the listening situation as a whole are largely based on culture-specific learning related to space as a component of interpersonal communication.

Part I

Musical space – perceptual basis

1 Intrinsic space

It is commonly accepted that sounds can be described in terms of pitch, length, loudness and timbre. These are perceptual attributes which have quantifiable counterparts in frequency, duration, amplitude and spectrum, respectively. Largely due to its limited notatability, timbre has only relatively recently become a compositional element of similar status as pitch and rhythm. The notion of *space* as an inherent aspect of the individual sound, however, has barely begun its entry into musical thought. Extensive manipulation and use of both timbre and space on the level of the individual sound are the forte of the electroacoustic genre, and were not feasible as compositional means until the advent of the electroacoustic music studio around the middle of the twentieth century.

The notion of *intrinsic space* is based on the perception of internal spatial components inherent in individual sounds and sound events. It is necessary, therefore, first to outline the process of perceiving sounds and sound events before discussing the inherent spatial characteristics of the individual sound.

1.1 Auditory perception of individual sounds and events

Perception is a form of problem solving where we first segregate and organise the sensory information into separate coherent events and then try to discover the source of these events. The result of this detection work depends on the nature of the specific information that is the focus of attention as well as the context in which it is heard. All the steps involved in the perception process interact and influence each other in our interpretation of the acoustic environment. Two fundamental activities involved in the perception of sounds are *event recognition* and *source recognition.*

1.1.1 Event recognition

The nerves in our ears are constantly firing off impulses. Not only are these impulses triggered by external fluctuations in air pressure, but random unprovoked neural firing is also mixed in to the flow of signals being passed on to the brain (Gleitman 1991: 161). The auditory organs themselves are not capable of discriminating between wanted and unwanted signals. The neural pattern transmitted to the brain therefore resembles that of a spectrogram (Bregman 1994: 7). When looking at a spectrogram, it is almost impossible to separate the images of individual sounds because all the sonic events that are captured by the system are superimposed. Audition is able to decompose this complex description of the acoustic environment and distinguish individual sound events. This is done on the basis of differences in time and intensity of the primary components of the auditory input. These time and intensity differences provide us with enough information to recognise beginnings and endings, pitch, timbre, location and movement of sounds.

Event recognition refers to the recognition of a sound, or a group of related sounds, as one coherent event that stands out from the background. *Auditory scene theory* seeks to explain this recognition process in terms of how the physical acoustic events come to be represented perceptually as audio *streams* (Handel 1989; Bregman 1994: 10). The term "stream" is preferred in auditory scene theory because, in addition to single-sound events, it includes events that consist of more than one individual sound, such as footsteps or raindrops. The recognition of discrete streams is based on relations among acoustic properties of the auditory input. Auditory scene theory describes the relations of these properties on the basis of the Gestalt principles of perceptual grouping (Bregman 1994: 196-203):

1. *The principle of proximity*: elements that are close together in space or time are likely to originate from the same event.

2. *The principle of similarity*: sounds of similar timbre, frequency or intensity are likely to be perceived as belonging to the same event.

3. *The principle of good continuation and completion*: sounds emanating from the same event tend to be continuous in some way or another.

4. *The principle of common fate*: sounds that follow the same trajectory in terms of frequency, intensity and rhythm tend to be perceived as belonging to the same event.

The principles of perceptual grouping represent what Bregman terms *primitive segregation* in auditory scene analysis (Bregman 1994: 39). These benefit from the relative constancy of the sonic environment. Based on the uniformity of the behaviour of related acoustic components, these components are grouped and recognised as distinct sound events. Primitive processes are regarded as innate and basic to all hearing as they are not under voluntary control and do not involve learning (Bregman 1994: 667).

The listening process is also organised by more refined knowledge and experience. Knowledge of sounds is structured into particular classes of information, which are controlled mentally in units called *schemas* (Bregman 1994: 397). Familiarity and regularity in the sonic world are dealt with by schemas. They are voluntarily controlled and employed whenever attention is involved, for example in active music listening and when something specific is being listened for. Schemas direct the listening process in that hearing a sound from one particular class may create the expectation that other sounds that are related and belonging to the same class will follow. This way the listening mechanism is prepared for certain types of sounds, and the introduction of sounds "foreign" to that particular context may have the effect of surprise and require more effort in the recognition process. When such expectations are met, however, the event recognition process takes place rapidly and most efficiently. Schema-based processes can look at longer time spans than primitive processes, but the number of events that can be attended to at any one time is limited[1] (Bregman 1994: 399; Gleitman 1991: 248).

This leads to the conclusion that perception organises the acoustic information in two overlapping time spans. There is the short time organisation (the primitive processes) which involves interpreting the composite acoustic wave at any one moment, and the longer time organisation (the schemas) which involves interpreting the acoustic wave as it changes over time, each moment being a small part of the larger pattern. These two processes support each other. What is eventually perceived as one coherent event is therefore influenced by what happened before and what happened after each moment in time (Handel 1989).

Thus, in order to recognise sound sources and events, the auditory system takes advantage of cues found in the context in which the sounds are heard (Handel 1989). The perceived qualities of a sound event can therefore change with varying contexts even if its physical attributes remain the same. With speech, for example, the context in which the language sounds are heard determines our ability to divide the stream into words and sentences. When analysing speech by looking at a visual representation of its waveform, we see an uninterrupted acoustic stream, the only silences being in connection with stop consonants and when the speaker is out of breath. The recognition of the individual words as distinct sound events is done on the basis of the context in which the language sounds are heard, and depends on the listener's fluency in the particular language. When encountering an unfamiliar language it is generally impossible at first to recognise individual words, and the acoustic stream sounds relatively continuous. Even language sounds that are known from our own native language are often difficult to pick out and recognise due to the unfamiliar context. After a while patterns begin to emerge, and eventually the acoustic stream is decomposed into meaningful linguistic components.

[1] This has to do with attention span, which in turn has to do with the capacity of short-term memory. The number of items that can be held in short-term memory at any one time is reported to be "7 plus or minus 2" (Gleitman 1991: 248).

1.1.2 Source recognition

Recognition of the source of a sound is often affected by higher-level knowledge and expectations. There is both a top-down and a bottom-up process involved in recognition (Gleitman 1991: 223-24; Blauert 1997: 410). Generally, the perceptual system starts out with both a stimulus and a hypothesis, representing the bottom-up and top-down aspects respectively. Features of the stimulus are analysed and tested against the hypothesis, which is either confirmed or is signalling a need for further investigation. If sufficient correlation between stimulus and hypothesis is not found, a new hypothesis is considered and tested. This process continues until the sensory input is recognised or we settle with an unsolved problem. When we hear familiar sounds, this process takes place unconsciously and at a very high speed. However, with unknown sounds we become aware of the steps involved in the recognition process, and the decision whether to carry on the investigation or accept that we cannot detect the sound source is often a conscious one. The top-down and bottom-up theories of problem solving make sense because without top-down processing memory, experience and expectation would have no effect, and without bottom-up processing there would be no effect of the external sensory information, and hence no perception (Gleitman 1991: 224; Blauert 1997: 410).

When hearing new sounds in a natural context, the input processed by several senses combine to form the knowledge of the sound as it is represented in memory. There are very few known non-electronic sounds of which we have not at some point seen the source[2]. Vision, therefore, plays an important role in the recognition of sounds. Seeing the source as it sounds makes the sound easier to remember than if only hearing is involved in the learning process. Exciting the sound by direct bodily contact, such as playing an instrument or splashing in water, involves touch, and perhaps also smell, which adds further knowledge about the sound source. The context in which the sound is experienced is significant in forming the knowledge and point of reference associated with it. For example, recognising the sound of a car as "a car" is based on knowledge about the object "car"; its typical size and shape, the material it is made of, what it is that makes the car generate sound and in which environmental contexts cars typically are found. By incorporating car sounds into a musical composition, a certain environment is implied and understood based on the knowledge that, for instance, cars are generally found outdoors and are mostly used for transportation and travel. With known sources, such associations are often triggered unconsciously and become aspects of the sound that are taken for granted. This is reflected in language, where people may respond to car sounds with phrases such as "I hear a busy road" or "I hear a fast car." These statements are based on contextual knowledge about the source associated with the sound, and are thence not attributes of the sound itself. The former phrase is a response

[2]For the blind, touch replaces vision as the dominant sense besides hearing.

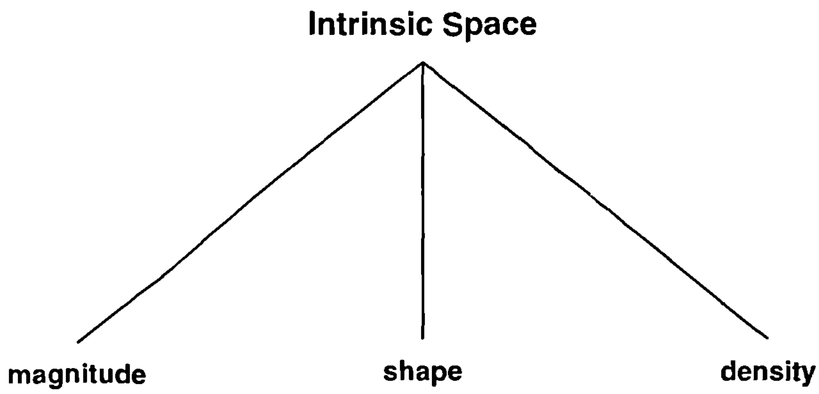


Figure 1.1: Intrinsic space.

to hearing the sound of many cars in the same spatial direction and associating it with the knowledge that cars are driven on roads which are occasionally crowded. The latter phrase refers to the sound of a raced car engine which changes spatial location at a high speed, and is based on the knowledge that cars are vehicles capable of rapid movement.

Experienced listeners are often able to suppress the source recognition process and instead fully concentrate on the spectral and spatial qualities of the sounds and their relative organisation. The success of this listening strategy often depends on the sound material in the composition. For example, in an electroacoustic work consisting mostly of heavily processed or synthesised sound material, an easily recognisable natural sound sticks out, and its source immediately comes to mind. In this way, a single sound can completely change our perception of preceding and subsequent events in the piece. In a composition consisting entirely of recognisable sounds it can be easier consciously to ignore the sounds' origins in the listening process, and instead concentrate on their spectral and spatial shaping and structural relationships. Composers of electroacoustic music generally have the ability to suppress source identification, but may instead, often involuntarily, identify the processing technique or synthesis method used in creating the sound.

1.2 Attributes of intrinsic space

There are three main attributes in the description of intrinsic space: *magnitude*, *shape* and *density* (see figure 1.1).

1.2.1 Magnitude

Most individuals will agree that different kinds of sounds appear to be different in size. Early perception studies into spatial attributes of individual sounds concentrated on the phenomenon of size perception, usually termed "volume" or "tonal volume" in the research literature (Boring 1926; Hollander and Furness 1995). "Volume" in common language most often refers to *sound pressure level* in connection with sound systems and music listening. I therefore prefer the term *magnitude* for this particular aspect of individual sounds. Magnitude, in the context of intrinsic space, is a perceptual entity that is affected by a great number of variables related to circumstance, source, spectral makeup and room acoustics, and is therefore difficult to quantify.

In terms of sound spectrum, it is particularly intensity and the amount of low frequency energy which contribute to the sound's magnitude. Magnitude appears to expand as the pitch goes down and when intensity increases (Boring 1926; Hollander 1994; Hollander and Furness 1995). There are several reasons for the influence of low frequency content on the perception of magnitude. One important factor has to do with the knowledge associated with the sound in relation to sound source and typical context, as addressed above. Experience from our physical surroundings tells us that low frequency resonance is generally found in large objects (Handel 1989; Wishart 1996: 191). We therefore tend to associate low frequency sounds with objects that take up a large amount of space or with forces that cover large areas, such as thunder or earthquake rumbling. The connection between frequency and magnitude is discovered at a very early age. I have observed young children (under the age of three) talking in a high-frequency voice when describing little things, and lowering their voice when referring to big things. They also referred to a high-frequency voice as "a *little* voice."

Low frequencies are, practically speaking, radiated almost equally in all directions (Winckel 1967: 151), and therefore cause a greater number of reflections from walls and other objects in the listening environment than do sounds of higher frequency. This *indirect sound* adds energy to the acoustic information at the listener's position, resulting in increased intensity which in turn leads to the impression of greater magnitude. The indirect sound arrives from other directions and at different times than the direct sound. This leads to a sense of spaciousness as the sound seems more spread out in both time and space. Also, due to their long wavelength, low frequency waves tend to bend around the head. The wavelength is longer than the distance between the ears, and the head does not cast an acoustic shadow as it does with high frequencies that prevents sound waves coming from the side to reach the other ear (Handel 1989; Kendall 1995: 25). This contributes to the relatively poor localisation of low frequency sounds and makes them appear to fill the listening space to a greater extent than do sounds of higher frequency.

Duration is another important factor in the perception of magnitude. Perrott et. al. (1980) have observed that stable sounds can seem to grow in magnitude as they are being listened to. In this investigation, listeners were presented 5 kHz tones at stable amplitude for a duration of five minutes. The subjects reported the sound to increase continuously in magnitude. One of them described the sound as gradually expanding until it completely filled his head. The researchers were able to determine the rate of expansion to be consistent across frequencies (Hollander 1994). In a musical context attention is more likely to drift between sounds, and a stable tone such as a drone or a pedal point will only seem to be growing when it is particularly attended to and concentrated upon. Nevertheless, duration does contribute to increased spaciousness in that it gives the sound time to interact more with the acoustic environment and blend with reflection and reverberation.

When the sound source is recognisable, the judgement of magnitude is often influenced by knowledge about the source. The association of the typical size of the perceived source suggests the magnitude of the sound and, in a musical context, also gives the listener an idea about the size of the virtual sound field of the composition in which the sound is placed. However, knowing the size of the source can sometimes lead to false judgements. As one study has shown, listeners expect there to be a difference between male and female hand clapping (Repp 1987). Because males generally have larger hands and arms, their clapping is expected to be slower, louder and of lower pitch than female clapping. The fact is that there are no gender differences in hand clapping since the sound has to do with how one hand hits the other.

Similarly to vision, there is a certain sense of size consistency in auditory perception. A sound appears to be moving farther away when its intensity and high-frequency energy decrease and, for enclosed acoustic environments, reverberation increases. In this situation the sound's magnitude is likely to remain constant as the initial close-up perception of the sound provides a point of reference with which the later added distance cues are compared. However, size constancy may only be perceived when the sound is moving into the distance. Judgement of magnitude of distant sounds or of sounds moving from the distance is difficult unless they are known to us and we already have the necessary point of reference in terms of magnitude.

1.2.2 Shape

The shape of a sound has to do with how its spatial properties change over time. This shaping is perceived on the basis of amplitude fluctuations in the spectral components that make up the sound. The energy distribution in the sound's spectrum can change over the course of its existence, resulting in constantly varying magnitude. Temporal variations in complex spectra can even cause several audio streams to appear as if the sound splits during its existence.

The *overall* amplitude envelope of the sound is perhaps most *directly* influential on its spatial shape as it only contributes to the expansion or diminution of the magnitude over time, and not to any change in perceived spatial placement or distance as variations in spectral distribution might do. Attenuation only of high frequency components in the sound, for example, is more an indication of increased distance than on a change in magnitude, and, hence, does not necessarily contribute to the perception of the sound's shape.

1.2.3 Density

Density has to do with the compactness or solidity of the sound. A sound of high density seems hard and impenetrable. It is difficult to hear *into* the sound, and it seems spectrally closed in addition to not being very spread out in space. Conversely, we can encounter sounds that seem to enclose a space, so that they are perceived as resonant or hollow in their quality.

Density can also be interpreted on the basis of associations with the sound's source or the physical gesture behind its excitation. Thus, a sound with a short decay, or one consisting only of an attack, can be interpreted as having a solid body, or its source can appear to be dampened due to some kind of human gesture.

2 Extrinsic space

We base our aural orientation of the world on information derived from the interaction between sounds and their surroundings. This information enables us to locate sounds and objects around us in relation to our own position in physical space. In electroacoustic music, sonic interaction can be of a virtual kind as it is when spatial cues are composed into a work, and it can be of an actual kind, such as the acoustic interaction between the sounds emitted from the loudspeakers and the physical listening environment.

Extrinsic space is my term for the spatial properties that provide listeners with information about the *direction*, *distance* and *movement* of sounds. Here I assume that the sounds are heard in a *sound field* where they can be localised relative to a specific listening position. The notion of extrinsic space therefore concerns the sound *in* space. The perception of extrinsic space is based on spectromorphological information as well as on spatial information resulting from the sound's interaction with the environment in which it exists. Extrinsic space is an inevitable part of sounds which are recorded in an acoustic environment, in which case the recording captures information about the sounds' location relative to the microphone as well as the acoustic characteristics of the environment in which the recording was made. Recorded sounds often carry additional information about the wider sonic context in which they are situated—aspects which can be highly referential.

2.1 Spatial hearing and acousmatic listening

In our everyday auditory experience, the signals that reach the two ears are rarely identical. This is mainly due to differences in how sound waves are reflected and diffracted in the environment, but is also due to asymmetries of the torso, head and pinnae, and to inconsistencies in humidity and temperature in the sounds' travelling towards our ears (Handel 1989). The environment in which sounds are listened to and the number and placement of the sound radiator(s) and reflective and absorbant surfaces relative to the listening position are influential factors when it comes to locating sounds. Thus, in a concert situation each listener will receive a different signal and hear different things in terms of space.

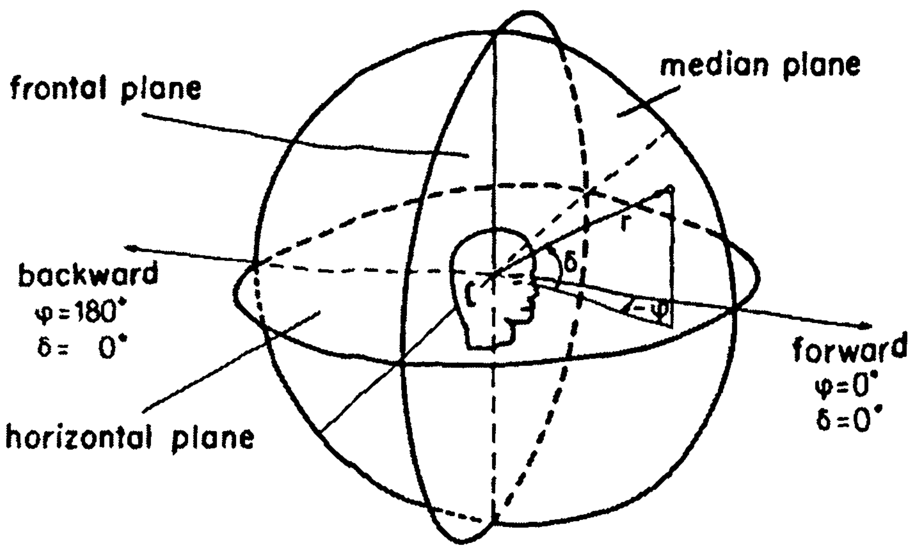


Figure 2.1: Head-related system of spherical coordinates (From Blauert (1997) Figure 1.4, p 14. Courtesy of MIT Press.)

Acousmatic music[1] is listened to over loudspeakers, and although the sound *sources* are thereby hidden from the view of the audience, the sound *radiators* (the loudspeakers) can normally be seen. However, as the majority of the music we hear today comes out of loudspeakers we have come accustomed to "seeing through" and largely ignoring the loudspeaker as the actual sound-emitting source. This, I believe, is partly because most of us are used to the phantom images of stereo systems, and we therefore expect the sound to appear from somewhere in between the loudspeakers rather than from the points where the loudspeakers are placed. Similarly when watching television: even though the loudspeaker[2] is normally positioned to the side of the image, the location of the sound usually matches the location of the corresponding visual event. With the eyes closed, the auditory event shifts to the side where the loudspeaker is positioned. Expectations of this kind guide the top-down processing of directionality information, and can steer the location of an auditory event closer towards where it is expected to be or, in the case of audio-visual stimuli, towards where the corresponding visual event is located (Blauert 1997: 193). One must therefore distinguish between the localisation of the sound radiator and localisation of the perceived auditory event. In everyday situations the two normally coincide, but when listening to music over loudspeakers the contrary is more often true. Most audio reproduction techniques

[1]The particular genre of electroacoustic music where all the sound sources are hidden from the view of the audience has been termed *acousmatic music*. The term has connotations to a certain compositional tradition and listening attitude.

[2]A single loudspeaker to the side of the screen is still the most common audio facility on television sets.

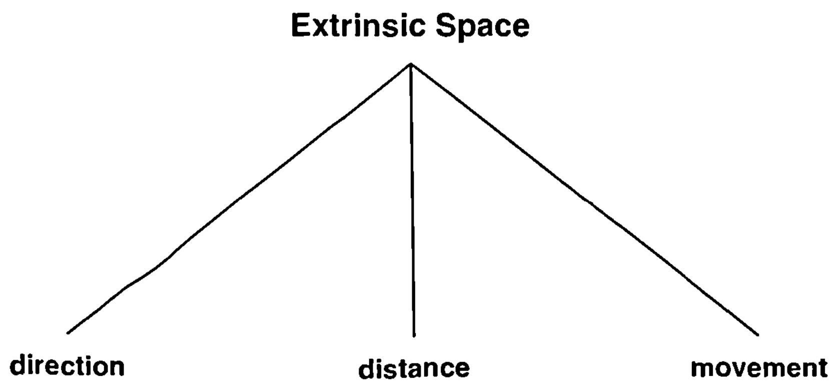


Figure 2.2: Extrinsic space.

rely upon a successful illusion of placement and movement of auditory events between and away from the loudspeakers rather than pointing the loudspeakers out.

Vision is an important aid in shifting the localisation of auditory events, as a shift in location is actually shown to occur when the eyes are moved (Blauert 1997: 196). Spatial acuity in vision is much higher than it is in audition, and the most accurate localisation takes place when there is an agreement between visual and auditory events. This indicates that vision and audition process spatial information in the same area of the brain, and support each other in the process of spatial mapping (Auerbach and Sperling 1974; Hollander 1994). Such a view is supported by Warren (1970) who found that localisation of invisible sounds is much more accurate with the eyes open than with the eyes closed, and is also better in the light than in the dark. Since there is a considerable drift of the eyes in the dark, this may contribute to the relatively poorer sound localisation under such conditions (O'Connor and Hermelin 1978: 46).

2.2 Attributes of extrinsic space

2.2.1 Direction

When judging the direction of a sound, the auditory system takes advantage of time, intensity and spectral information in the sound signals. *Interaural time differences* (ITD) and *interaural intensity differences* (IID) are the basis for directionality judgements on the left/right axis, while the spectrum-based *head-related transfer functions* (HRTF) enable the auditory system to determine whether the sound is in front, above, below or behind. The perception of direction is based on complex combinations of the information from these three phenomena, which can all be manipulated electroacoustically.

When listening to music on loudspeakers, ITD and IID are fundamental to the localisation of sounds in the stereo field and to the sense of spaciousness in stereophonic material. ITDs refer to differences in arrival times of the sound signals at the two ears, while IIDs are dissimilarities in sound pressure level between the signals at the two ears. A sound arrives sooner and is more intense at the ear nearer to the sound radiator. Even the slightest asymmetry in the sound's path from source to ear leads to components of the sound to travel different distances to the two ears, resulting in dissimilarities in the phase angle between the ear input signals. As a signal moves from directly ahead towards the side, the ITD, also referred to as *phase delay*, increases from 0 to approximately 650 μs (Kendall, Martens and Decker 1991: 66; Kendall 1995: 27; Blauert 1997: 143). In stereo listening, if the spectral components of the signal at one loudspeaker are delayed by different amounts than those at the other loudspeaker, one hears a clear widening of the sound in the stereo field (Blauert 1997: 146). When the wavelength equals the distance between the ears, the phase angle is the same at both ears, and ambiguity in localising the auditory event occurs. ITD works best as a spatial cue for frequencies below 800 Hz[3], above which the effect of displacement based on phase angle alone decreases significantly. It has no effect above 1600 Hz because the head acts as an acoustic barrier for short wavelengths (Kendall 1995: 31; Blauert 1997: 148-9). If the general arrival time of the signal at one ear is more than 1 ms earlier than that at the other, then the sound is in most cases localised at the direction from where it first arrives. The later arriving signals at the other ear are largely ignored in the localisation process[4]. This is what is termed the *precedence effect* (Blauert 1997: 204; Pierce 1999: 92-93), and is a commonly utilised phenomenon in electroacoustic sound reproduction for horizontal placement of sound in a stereo sound field.

At high frequencies, time differences provide little information, and lateral localisation is therefore largely based on intensity differences between the two ears that vary with angle of arrival of the sound at the head (Holman 2000: 206). IID increases from 0 to around 20 dB as the sound moves horizontally from directly in front towards the side (Kendall, Martens and Decker 1991: 66). Thus, the signal is heard at only one ear when the intensity differences between the two ear input signals are greater than approximately 20 dB. When a signal of relatively long duration is stronger at one ear than the other, the auditory event shifts towards the centre due to a natural adaptation and fatigue of the more strongly stimulated ear. It takes some time, usually a few minutes, to retain normal sensitivity after such stimulation (Gleitman 1991: 167; Blauert 1997: 163).

Because differences in time and intensity only provide cues for localising sound on the left/right axis, spectral information becomes crucial in order to sense whether the sound comes from behind, in front, above or below. The head-related transfer function is a spectral

[3] An 800 Hz tone has a wavelength corresponding to approximately half the distance between the ears.

[4] However, the later arriving signals are not suppressed entirely, as the earlier discussion (p 34) on cognitive shifts in localisation shows.

profile created by the sound reflecting off the listener's upper torso, shoulders, head and pinnae. HRTFs are therefore unique for each individual, and are different for every distance and direction of sound arrival and different for each ear. It is the pinna, due to its convoluted shape, which contributes most strongly to the HRTFs by filtering sound signals and creating resonances that vary with the direction and distance of the sounds. This filtering effect is particularly significant for frequencies higher than 4000 Hz (Rogers and Butler 1992: 537).

Although the transfer functions differ considerably from individual to individual, certain trends in the spectral pattern resulting from the filtering in the pinnae have been observed in an attempt to arrive at "a set of idealised transfer functions that will provide the best possible image of sound direction for the general population" (Kendall, Martens and Decker 1991: 68). Generally, individuals localise sounds most accurately with their own transfer functions, but it has been found that some pinnae provide more accurate cues for localisation than others (Kendall, Martens and Decker 1991: 70). The use of generalised HRTFs and binaural recording technology is at present most successful for headphone reproduction, where acoustic crosstalk between the two loudspeaker channels and natural reverberation from the listening environment and the effect of the listener's own HRTF are eliminated.

The accuracy in determining direction is not the same for the whole listening sphere. The lower limit of what has been termed *localisation blur* (Blauert 1997: 37), the resolution with which listeners determine the spatial location of a sound source, is about 1° in the forward direction (O'Connor and Hermelin 1978: 45; Blauert 1997: 38; Holman 2000: 207).[5] The direction straight in front of the listener is the region of the most precise spatial hearing in both the horizontal and vertical directions. In the horizontal direction, localisation blur increases with displacement of the signal to either side, with a maximum at the right angles to the facing direction where the minimum audible angle is about 10°. Behind the listener the minimum angle is near 6° (Kendall 1995: 32). The resolution of vertical localisation is lower, with a maximum accuracy in front of about 4° and overhead of about 22° (Kendall 1995: 32; Blauert 1997: 44). Localisation blur is different for different sound signals depending on their loudness, duration and spectral content, so that the actual localisation blur depends on the type of sound. In particular, localisation of low frequency sounds is much less accurate than it is for sounds in the mid- and high-frequency range.[6] Furthermore, localisation differs (within limits) from individual to individual, and can vary over time, partly as a result of learning and experience. Accounting for variations with different sounds, Blauert (1997: 40-41) outlines localisation blur behind the listener to be approximately twice the value for the forward direction and between three and ten times the

[5] Kendall (1995: 32) has found the minimum audible angle to be "2 degrees or less, depending on the exact nature of the experimental task."

[6] The most sensitive natural sound material (under ideal listening conditions) is found to be male speech (Holman 2000: 207-208).

forward value for the sides.

Duration plays a role in determining direction in the sense that it may give the listener time to resort to head movements in order to more accurately locate the auditory event. Not only does this lead to a greater amount of information from variations in transfer functions and the interaural differences on which we base our judgement, but facing towards the sound also decreases localisation blur, making more accurate localisation possible. Head movements are particularly important in determining the front/back localisation of sounds, especially when they are located near the median plane where interaural differences do not provide sufficient information (Blauert 1997: 44).

2.2.2 Distance

Distance judgements are based on loudness, spectral cues and degree of reverberation. As a sound source moves away from the listener, the overall intensity in a free field decreases at a rate of the inverse square of its distance[7]. Hence, a doubling of the distance corresponds to a physical reduction in amplitude of 6 dB (Ahnert and Steffen 1993: 167). However, since loudness is a perceptual entity, the experience of distance does not necessarily correspond to the results of physical measurements. Moore (1991: 99-100) reports that, especially for unfamiliar sounds, subjective impression of distance is approximately proportional to the inverse cube of the distance[8] so that a doubling of distance leads to a reduction in intensity of about 9 dB.

Source recognition is therefore a crucial factor in distance hearing. When encountering familiar sounds or sounds of a familiar category[9], spectrum rather than loudness becomes the primary cue for judging distance (Chowning 1999: 271-72). With known sounds, we are generally familiar with the timbral differences between a soft and a loud excitation of the sound, and can judge, based on the combination of loudness and spectrum, whether the sound is excited near by or far away. Human speech at a normal level can generally be localised with quite good accuracy by most listeners. Localisation is, however, found to be considerably better when hearing the voice of someone familiar compared to that of an unfamiliar person (Blauert 1997: 44-45).

When determining distance, the auditory system takes into account that an increase in

[7]This is the *inverse square law:* $I \propto 1/d^2$, where I is intensity and d is distance. This is equivalent to $A \propto 1/d$, where A is amplitude and d is distance (Moore 1991: 99; Chowning 1999: 269).

[8]The inverse cube law is expressed in the form of $I \propto 1/d^3$, which is equivalent to $A \propto 1/d^{1.5}$, where I is intensity, A is amplitude and d is distance (Moore 1991: 100).

[9]We tend to categorise sounds in memory as metallic, hollow, wooden, vocal, etc. (Kendall 1991: 71). In the sound recognition process, an unfamiliar sound is put into the appropriate category based on the familiarity of the morphology and spectral qualities of the sound. Any sound with a sharp attack and a rapid decay, for example, may be identified as "a hitting-sound," even if it is obvious that it is synthesised and no hitting actually takes place when producing the sound (Risset 1996: 31).

distance makes high frequencies diminish more rapidly than lower frequencies due to air friction and absorption. According to Blauert (1997: 118), air-related high-frequency attenuation only becomes an issue for sounds more than approximately 15 m away. However, the previous section's discussion on HRTFs indicates that spectral distortion of sounds in the pinnae is taken advantage of also when determining shorter distances.

In addition to loudness and timbre, reverberation is an important element of distance hearing. The effect of reverberation on distance perception has to do with the proportion between direct and indirect sound at the position of the listener (Chowning 1999: 272). The location of the auditory event becomes less precise when the time-interval between direct sound and reverberation decreases and the intensity of the reverberation relative to direct sound increases. As is known from real-world experience, little or no reverberation indicates that the sound is located near by, while a great amount of reverberation leads to a diffuse sound field of largely indirect sound, which informs us that the sound source is located far away. Reverberation also tells the listener that the sound is excited in an enclosed space, and a spontaneous image regarding type, size and acoustical properties of this space is formed. This phenomenon has been termed *spatial impression* (Blauert 1997: 282), and is central in perceiving virtual spaces in electroacoustic music where there is no visual information. The boundaries of an auditory space defined by spatial impression indicates outer limits to the possible distances of auditory events located in this space.

2.2.3 Movement

Movement, either of the sound source or the listener, involves time varying information from IIDs, ITDs and HRTFs. A great amount of spatial data is processed in relation to auditory motion, especially if the listener (for example during exploratory head movements) and the sound source are moving at the same time in varying directions relative to each other. An important cue in the perception of motion, in addition to the principles outlined above, is the *Doppler effect*, which describes the pitch shift that occurs when the distance between sound source and listener varies. It is the rate at which the distance changes that determines the size of this pitch shift. For example, when a sound source is moving at a constant velocity along a path which passes some distance in front of the listener, the increase in frequency is at its maximum as the source approaches from a great distance. The frequency increment decreases as the source approaches, and reaches its true value as the source passes in front of the listener. As the sound source moves away from the listener, the frequency decreases rapidly at first, and then more slowly as the source travels into the distance. The Doppler effect is a significant cue in the perception of motion, since a sound source will only have to move at a speed of 2.1 km/h (0.58 m/sec) to cause a perceivable change in frequency (Handel 1989: 99). The effect is the same whether it is the sound source or the listener that is moving.

Electroacoustic manipulation of auditory space can exceed the limits of perception, for example in terms of how rapid changes in position of sound events can be detected. Blauert (1997: 47-48) has found that a full cycle of left-right alternation must take at least 172 ms and front-rear alternation 233 ms for the auditory system to follow accurately the moving sound. For circling sounds, as the speed of the moving sound increases beyond that of accurate detectability, the auditory event is first reported to oscillate between the left and right sides, and after further increase in speed to become spatially stable, located approximately in the middle of the listener's head (Blauert 1997: 47).

2.3 Categories of sounds based on extrinsic space

As the discussion above shows, localisation of sound events depends on a great number of acoustic factors that affect the sound waves as they travel from the sound radiator to the eardrums. In order successfully to place sound events away from the loudspeakers and to create convincing virtual spaces in composition, it is important to be aware of these factors.

The categorisation below is grounded on spatial characteristics that are fundamental to localisation as well as to how these characteristics influence electroacoustic music listening. I am only considering individual sound events here; they are not assumed to be in any structural relationship at this point.

Stationary sounds

- *Directional sounds* are sounds that can easily be localised in physical space. These are typically sounds of little magnitude and of high density. It is easier to pin-point the location of sounds of mid- to high frequency than low-frequency sounds. Also, placement of the sound source relative to the listener is an important factor due to the varying degree of localisation blur in different directions. Sounds that are meant to be accurately localised, therefore, need to be placed in front of the listener, where localisation is most precise.

- *Non-directional sounds* are diffuse sound-regions which occupy wider areas of the space (Kendall 1995: 24). There are various degrees of non-directionality: some sounds fill the entire space and can be experienced as "coming from everywhere," while others can be localised as coming from specific areas in the listening space. "Sound-clouds" and sounds of great magnitude are typically non-directional.

Localisation blur can be a factor when determining whether spatially ambiguous sounds are directional or non-directional. It is difficult to be absolute in categorising stationary sounds based on directionality when encountering borderline cases.

Moving sounds

- *Sounds coming from or moving into the distance.* This implies changes in distance cues: intensity, degree of reverberation and spectral and temporal distribution.
- *Sounds moving from left to right or from right to left (laterally moving sounds).* When this entire movement takes place either to the left or to the right of the listener, the sound has to move a greater distance to be perceived as changing location because of the higher localisation blur. A change in location is easier to perceive if the sound moves through the median plane, especially to the front of the listener.
- *Sounds moving from front to back or from back to front* are detected almost entirely on the basis of spectral changes. Variations in spectral quality are perceived as the sound moves; sounds are brightest when in front. Furthermore, if the movement takes place overhead and/or through the frontal plane, it goes through the regions of the least precise spatial hearing and the movement therefore needs to be of sufficient range in order to be detected.
- *Elevating sounds* are perceived to be rising or falling relative to their onset location. When listening to a sound source that is not physically moving, elevation is closely connected to the metaphor of high and low in connection with pitches. Often under such conditions, it is the change in spectral distribution that causes the sound to be perceived as moving up or down in physical space due to frequency-dependent reflections in the pinnae. Cultural factors, such as the notation of high frequencies above lower frequencies in traditional Western music notation and the use of the words "high" and "low" in describing such sounds, may amplify the sense of elevation.
- *Dispersing sounds* begin as directional sounds and spread out in the sound field during their existence. Thus, the shape of the individual sound is changing over time.
- *Converging sounds* begin as non-directional sounds and become directional by gradually turning denser and often smaller.

3 Spectral space

Spectral space spans the lowest to the highest audible frequency. It is a vertical space in which sounds are localised on the basis of spectral emphasis, such as pitch or nodal spectrum[1]. *Physical* localisation of the sound source is largely irrelevant for the notion of spectral space as it is a psychologically and psychoacoustically based sense of elevation and vertical placement. Sounds in spectral space, therefore, cannot be pointed to and localised physically in the same way as can sounds in acoustic space. In that sense it is a metaphorical space. Nevertheless, it is highly influential in the spatial experience of electroacoustic music, and must be considered in any investigation into musical space.

3.1 Theoretical background

3.1.1 Perception theories

The link between frequency and localisation in spatial hearing has long been known. Early studies[2] into directional hearing found that high tones are localised spatially higher than low tones, confirming the experience many listeners have in regard to perceived elevation of (pitched) sound events. The early experiments into this phenomenon only made notice of frequency-related elevation *in front* of the listener, (probably) because the subjects knew that the only actual sound radiators were in the front, and thus became biased towards frontal localisation (Blauert 1997: 114). Later studies, with hidden sound sources or dummy sources behind the subjects, have shown that a sound emitted from a stationary source can, in fact, appear to come from anywhere between below the frontal horizon to behind the listener depending on the spectral makeup of the sound. Furthermore, a static sound emitted from two loudspeakers level with the listener's ears can appear to vary in elevation as the angle

[1]*Nodal spectrum* is a term originally coined by Pierre Schaeffer (referenced in Smalley (1986: 67)). It refers to sounds with a strong spectral emphasis, but without a clear identifiable pitch.

[2]Urbantschitsch, V. (1889): "Zur Lehre von den Schallempfindungen." In *Phlügers Arch.*, 24. Pratt, C. C. (1930): "The spatial character of high and low tones." In *J. Exp. Psychol.*, 13. Trimble, O. C. (1934): "Localization of sound in the anterior-posterior and vertical dimensions of "auditory" space." In *Brit. J. Psychol.*, 24. All quoted in Blauert (1997: 106).

between listener and loudspeakers in the horizontal plane is altered. Blauert (1997: 219) gives a brief description of what has been termed the *elevation effect*:

> In auditory experiments using [the standard stereophonic loudspeaker array] the auditory event frequently appears at a certain angle of elevation δ rather than in the horizontal plane. If the subject moves toward the loudspeakers in the plane of symmetry of the array, the auditory event becomes elevated by a greater angle; when the subject is exactly midway between the loudspeakers, the auditory event appears directly overhead ($\delta = 90°$).

The elevation effect is explained on the basis of the spectral components of the sounds in question. Blauert refers to related research[3] which has shown changes in IID and ITD that are similar when a listener turns the head through an angle α relative to a two-loudspeaker arrangement as when a single sound source is elevated through an angle $\delta = \alpha/2$ (Blauert 1997: 219). Because the elevation effect also occurs without turning the head or displacing the sound source, the explanation must be related to the sound spectrum. From what has been said earlier about HRTF and the dependency on the sound spectrum for accurate localisation, it makes sense that variations in spectrum will result in variations of the localisation of the sound events and vice versa. This view is supported by Butler (1973) who found evidence for the influence of the spectral content of a stationary sound source on perceived elevation of the auditory event. Butler presented his subjects with a series of sounds with the same pitch, but with different timbre. The sounds were radiated from five (one at a time) randomly activated loudspeakers located at different heights straight in front of the listener. The presented timbres had their formant frequencies at 630 Hz, 1600 Hz, 2500 Hz and 6300 Hz, respectively. The tendency in listener response was that timbres with formants centred around the first three of the test frequencies were localised in the lower, middle and higher regions, respectively, while the judged localisation of the highest frequency timbre coincided largely with the actual localisation of the loudspeaker it was emitted from. Butler comments that the frequency 6300 Hz is known to be among those essential for accurate vertical localisation of sounds, and this result was therefore expected (Butler 1973: 257).

The spectral modifications of the sound caused by the filtering and the resonating effects in the pinnae result in peaks and notches whose centre frequencies vary with the direction from where the sound arrives (Rogers and Butler 1992: 537; Blauert 1997: 310-11). In other words, different directions of arrival cause certain frequencies to be boosted while certain other frequencies are attenuated. With narrow-band signals at least, this can be turned

[3] de Boer, K. (1946). "The formation of stereophonic images." In *Phillips Tech. Rev.*, 8. de Boer, K. (1947). "A remarkable phenomenon with stereophonic sound reproduction." In *Philips Tech. Rev.*, 9. Wendt, K. (1963). *Das Richtungshören bei der Überlagung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie.* Dissertation. Aachen: Technische Hochschule.

around to say that the emphasis of certain frequencies will lead to the perception of certain directions of arrival, as Butler's (1973) study shows. Peaks at certain key frequencies[4] have been found to be particularly influential on elevation (Rogers and Butler 1992: 536; Blauert 1997: 107-16, 311):

- at 250-500 Hz the sound is most frequently reported to originate in the front;
- at 1,000 Hz the sound is reported to come from behind;
- at 4,000 Hz from the front;
- at 8,000 Hz from overhead;
- at 12,000 Hz from behind;
- at 16,000 Hz from the front.

Thus, peaks in the frequency spectrum cause the sound to "move" in an arc above the listener's head between the front and behind. In order to achieve the effect of, for example, upwards elevation in spectral space, one needs to filter out (using a notch filter) the frequency portion of the sound spectrum associated with lower spatial regions. However, the success of this process depends on the type of sound, since not all spectral cues associated with a specific spatial region are concentrated within a single narrow frequency band (Rogers and Butler 1992: 545). Furthermore, Blauert (1997: 108) points out that the width of the directional bands vary among individuals, although the centre frequencies seem to be common.

3.1.2 Spectromorphological theory

Denis Smalley's *spectromorphological theory* includes an investigation of the concept of spectral space (Smalley 1986, 1997). Smalley is particularly concerned with musical experience and structural relationships in sound-based—as opposed to note-based—musics. At the core of his theory is the interaction between sounds (spectra) and their temporal alteration (morphology). In Smalley's theory, the term "spectral space" refers to the sense of a vertical space whose boundaries are defined by the sounds which occupy it (Smalley 1997: 121).

In spectromorphological theory, the sounds perceived in spectral space span the continuum between the *note* on one extreme and *noise* on the other. The presence or absence of pitch in spectral discrimination is a significant factor, and gives rise to two note-views: one view is the traditional emphasis of pitch over timbre, while the other view magnifies

[4]Termed *covert peak areas* in Rogers and Butler (1992), and *directional bands* in Blauert (1997: 108).

or "looks into" the note in order to make apparent the spectral components inside it. These two note-views represent external and internal *spectral focus*, respectively. In relation to notes, Smalley makes the distinction between *intervallic pitch* and *relative pitch*. The former implies the presence of more than one note and relates to the traditional use of pitch where intervals and pitch relationships are central, while the latter refers to contexts where pitches and distance between pitches are more diffuse and cannot be precisely placed in spectral space. Noise, on the other extreme of the continuum of spectral space, can occur in many different guises. It can span narrow or wide frequency bands, be of varying degrees of density, be of texturally different spectral and associative qualities and be coloured or resonant, for example. Smalley defines two noise-views: the first, *granular noise*, concerns qualitative attributes, while the second, *saturate noise*, refers to density. These two are not distinct, and represent two ways of describing non-pitched sounds.

In spectromorphological theory, the occupancy of spectral space in a musical composition is defined in relation to three basic reference points, *canopy*, *centre* and *root*. These reference points represent the outer limits and centre region of spectral space in an electroacoustic work, and form a *pitch-space frame* (Smalley 1986: 79). Canopies and roots are structural reference points which can act as goals or departure points for musical textures. The boundaries of the pitch-space frame become clear during the course of listening to the work; only very rarely is the full spectral range presented at the opening moment of an electroacoustic composition.

Smalley explains the sense of elevation in spectral space as a combination of analogical and actual aspects, where *source bonding*[5] is an important factor. Spectral height, he says, is related to high pitches being regarded as physically smaller than low pitches and therefore not rooted. Moreover, high registers can be more easily localised than sounds in the lower registers and are more spectrally mobile, something which Smalley speculates has an analogy with flight (Smalley 1997: 122). Thus, spectral space is perceived on the basis of a combination of learned musical and associative aspects. Smalley makes no mention of more innate factors such as those outlined in the previous section.

Trevor Wishart (1996: 191-92) also describes the connection between pitch and elevation as metaphorical. Wishart questions the notion of "high" and "low" in relation to pitched sounds on the basis of the paradoxical *Shepard tone*—the tone or tone sequence that is experienced as simultaneously going up and down[6]—and concludes that the association of high-frequency sounds with high physical localisation is not absolute. Although the

[5] "[...] 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." (Smalley 1997: 110.)

[6] The Shepard tone can be created by adding together a number of sine tones in octave relation while separately controlling the amplitude of each tone. The sine tones are made to gradually descend in pitch while at the same time the relative strength of these spectral components is gradually moved towards the higher frequencies (Risset 1991: 150).

Shepard tone is not a phenomenon of natural origin, it is a cause for Wishart to pose the question why we do not consider a descending tone going upwards instead of downwards. He, like Smalley, refers to flight as a possible explanation. The environmental metaphor of tonal "highs" and "lows" are, according to Wishart, based on the experience of airborne creatures having high-frequency voices and earth-bound creatures having deeper voices due to the size of their respective sound producing organs (Wishart 1996: 191). He underscores the significance of this spatial metaphor in music, and points in particular to instrumental music where metaphorical space has often been exploited by utilising the wide range of pitches available in the orchestra. An example here is the sense of open space created by orchestrating a high register melody together with a low bass figure and not having anything in the intervening registers.

3.2 Conclusion

The discussion above indicates that there is a psychophysical basis to the experience of spectral space in listening as well as a strong metaphorical side. In the context of musical space, perceptual research based on largely unrealistic laboratory experiments[7] can only serve as a foundation on which to build and apply knowledge from music listening and other relevant activities. Spatio-musical experience is a complex phenomenon which depends on a great number of factors. Perception is, nevertheless, at the base of the sensory input, and cannot be ignored.

There is no doubt that the metaphorical nature of spectral space is connected to associations with our spatial environment and real-world experiences. The knowledge that large objects produce low-frequency sounds and tend to be heavy and less likely to elevate than smaller objects, amplifies the perceived vertical displacement caused by the spectral makeup of the sound. Intrinsic space is therefore influential in that it is mostly sounds of great magnitude which tend to be perceived as heavier and more "earth-bound". However, there are sounds of great magnitude, such as thunder or the sound of aeroplanes, which are normally perceived as being elevated, but in this case it is not the sound itself, but its *source* which is known to be raised.

All sounds have spectral content, and have therefore a place in spectral space. Parallel to the categories of directional and non-directional sounds in extrinsic space, sounds in spectral space span an area of a certain width, or *spectral range,* of the space. Thus, sounds can be *concentrated,* as is sound material with a definite pitch, or they can be *diffuse,* as are noise-based sounds. In between the spectro-spatial extremes of (single-harmonic) pitch

[7]Based, for example, on single-sound stimuli, monaural listening, (near) anechoic listening environments and restricted head movements.

and (white) noise is a continuum of sounds of varying spectral range, a notion which corresponds to the continuum between note and noise in spectromorphological theory. Heavily noise-based sounds may span most or even all of the spectral space, so that a sense of space in the meaning of "distance" or "openness" in the spectral composition of a work may not be present in the listening experience. However, the concept of spectral space still remains helpful as a tool for describing and analysing electroacoustic music in terms of spectro-spatial density and spectral range.

When localising sounds in extrinsic space, listeners determine distance and placement relative to their own position in that space. In regard to spectral space, there is no such obvious physical point of reference. The judgement of a single isolated sound being "high" or "low" must, obviously, be based on comparison with something. Furthermore, since the measure of what is high and what is low seems to be fairly constant, the reference point must be within the individual rather than something external. It is therefore not far-fetched to suggest that the "height" of single sounds in spectral space is judged in relation to the fundamental pitch of one's own voice.

Part II

Musical space – composition, performance, perception

4 Composed space

Listeners bring with them spatial knowledge acquired from real-life experiences. This knowledge becomes the basis for the way in which they perceive and interpret spatial information in electroacoustic works. At the most fundamental level, perceptual mechanisms are the same for all listening. However, during music listening one tends to be more attentive to the sonic information than in casual situations, and is more likely to employ listening strategies to detect relationships and connections between sound materials over a longer time span and within defined (physical) spatial boundaries. Cognitive processes based on the listener's background and musical training come into play, and are essential for the listening experience.

The use of loudspeakers in electroacoustic music presents potential for using space as a structural element to a greater extent and with more sophistication than is possible in acoustic music. Space as a means for musical expression and communication was not given much attention until the advent of acousmatic music, when placement and localisation of sound material became flexible and manipulable elements for composers to work with. The spatial possibilities open to the electroacoustic composer are significant: one can transport the listener to a great variety of virtual sound environments, expand the listening space beyond its physical boundaries, play with intimacy and remoteness, and utilise movement and direction. All of these factors can be integrated into the compositional structure.

Composed space refers to the composer's organisation of the sound material into a musical context. This is where spatial relationships among the sound material are set up and virtual spaces based on the sounds' intrinsic, extrinsic and spectral spaces are established. Composed space constitutes a temporal space in which spatial configurations are connected as the work progresses in a structural manner. Figure 4.1 represents a schematic overview of the spatial components of composed space.

The boundaries of a composition's virtual space(s) are defined by the extremes in physical placement in the horizontal, vertical and depth directions (i.e. intrinsic space and extrinsic space), by the highest and lowest perceivable frequencies in the work (i.e. spectral space), and by information derived from any reverberation and reflections heard in the work. In addition, if there are references to real environments, these can aid in defining spaces in the

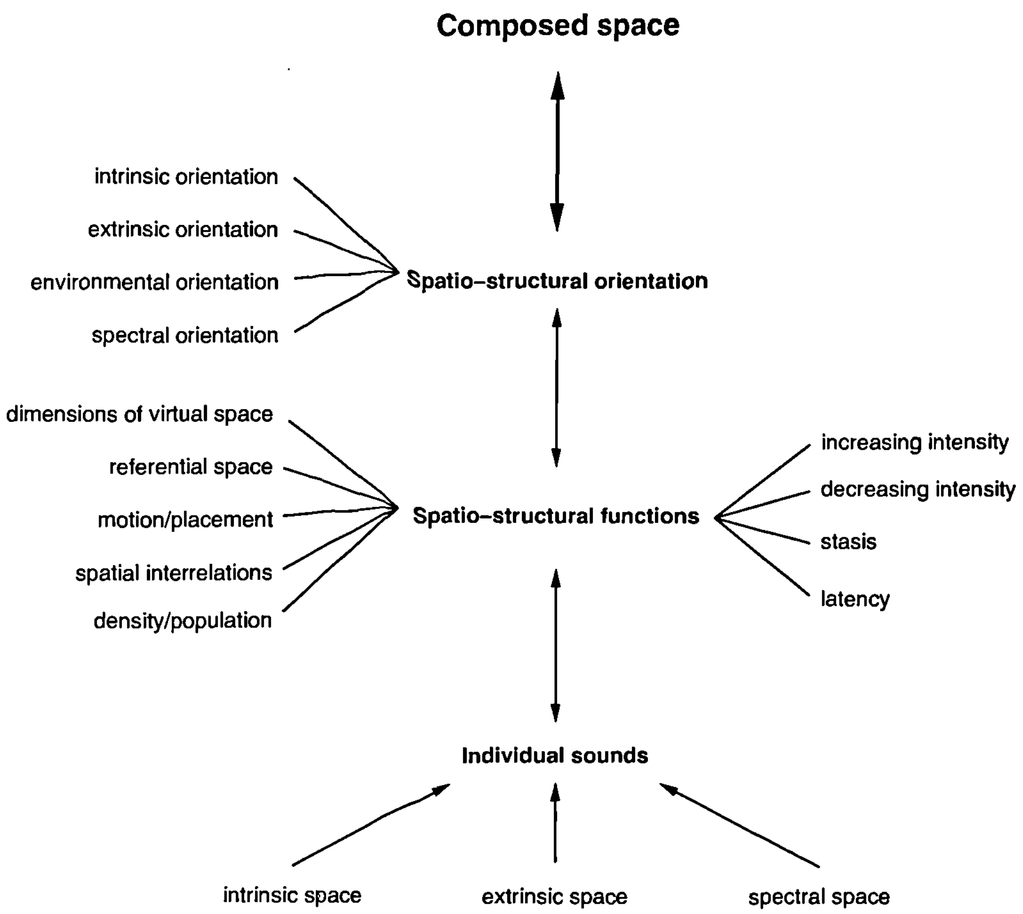


Figure 4.1: Composed space.

composition. Since music—and audition in general—is a temporal activity, the boundaries of virtual space set up in an electroacoustic work cannot be known at the onset of the piece, but can only be revealed during the course of listening. Virtual sonic spaces are by nature dynamic: they can be widely transformed during the course of the composition in ways that need not resemble any real-world spatial experience.

4.1 Space as compositional element

Unlike in casual listening circumstances, the position of the listener is normally fixed during attentive music listening. The spatial information available to the listener, apart from that obtained from exploratory head movements, is therefore entirely dependent on the cues composed into the work: the placement and movement of the sounds and the virtual space(s) they inhabit in the composition[1, 2].

One cannot manipulate a sound without concurrently altering its spatial properties. Sound processing and manipulation can affect qualities such as magnitude, shape, density, spectral distribution and spatial spreading of the sound material. The close relationship between sound and space manifests itself in a tendency for the spatial unfolding to follow the unfolding of the sound material at large in a work, so that, for example, a gradual move towards shorter sounds with more definite attacks leads to an increased emphasis on spatial interrelationships among the sounds as opposed to emphasising their inherent spectral and spatial qualities.

4.1.1 Structural functions

Structure in the context of composed space refers to the temporal shaping of spatial parameters into paths of motion and growth. Motion and growth are important aspects of musical expression and directivity, and guide expectation and anticipation in listening. In a spatial context, motion can refer to actual movements of sound events in physical or virtual space (i.e. extrinsic space), to any kind of spectromorphological change (i.e. intrinsic space or spectral space), or to change in the configuration of spatial interrelations among the musical material (i.e. composed space). It is this temporal shaping of musical parameters as functions in the compositional structure which gives rise to variations in intensity and evolution in the music, and thereby forms the basis for musical expression.

Structural functions in music have four possibilities: *increasing intensity*, *decreasing intensity*, *stasis* and *latency*. Parameters of increasing and decreasing intensity function as processes of growth and decline, and give a sense of progression or recession in the music. Stasis and latency indicate unchanging degrees of intensity, but they have opposite functions: stasis implies stability and rest, while latency implies tension and leads to an expectation of change. Neither stasis nor latency necessarily assume a low degree of musical activity.

The network of sounds and events in a musical composition is ordered—sometimes hierarchically—in units of different time-spans on various structural levels in the work. The time-spans, or perceptual boundaries, of these structural units are determined by the nature of the sounds and their function in the composition. The number of structural levels may vary throughout the work so that no single structural level needs to span the entire duration of the piece. Furthermore, it is unlikely that all of the musical elements and structural units

[1] The influence of the listening space—the sound system and physical space in which the listener is situated as well as the listener's position relative to the sound image—will be discussed at a later stage.

[2] Music in combination with visual stimuli (instrumentalists, dance, film) is not considered here, as an investigation into the audio-visual space is an enormous field on its own with a large number of extra-musical factors, many of which might direct the spatial listening experience.

in a work change in the same intensity-direction or change at the same rate. Musical material can therefore form *intensity-layers* which complement or counteract each other in terms of intensity, and are of different structural significance within the specific context. Space can represent one or several of these intensity-layers alongside intensity-layers of other musical elements such as dynamics, tempo and spectral content. Intensity-layers work on several levels in a composition, from the most detailed to the more global level.

4.1.2 Space on a macro-level: Spatio-structural orientation

Spatio-structural orientation concerns spatial change on the larger scale in a work—that is, on a high structural level in the composition. I have identified four categories of spatio-structural orientation in electroacoustic works in relation to the large-scale compositional structure. These are *intrinsic orientation*, *extrinsic orientation*, *spectral orientation* and *environmental orientation*. Intrinsic and extrinsic orientation as well as spectral orientation build on my notions of intrinsic space, extrinsic space and spectral space, respectively, set forth in relation to individual sounds, and can be seen as extensions to these into larger-scale structural contexts. Environmental orientation builds on extrinsic space, but concerns the wider spatial context of environmental cues and referential sound material composed into the work.

Several spatio-structural orientations can exist simultaneously in a work, but the main focus tends to be on one spatio-structural orientation at a time. Nevertheless, a work rarely stays within a single spatio-structural orientation all the way through, but varies between orientations from section to section[3]. This concept only applies to space on a macro-level, as variations in spatial focus on a more detailed level can occur quite frequently without necessarily having an effect on the overall spatio-structural orientation of the work.

Intrinsic orientation

For *intrinsic orientation*, temporal change in spectral distribution and spatial shaping inherent in the sound material carry the spatial development of the work. The focus is on the intrinsic space of sounds. External spatial relationships among sounds in terms of movement and placement become less prominent. However, extrinsic space does play a certain role in works of intrinsic orientation where an emphasis on spatial depth and distance is common. Generally, once the spatial relationships among the sound material are established, these relationships tend to change slowly, if at all, so that focus remains on the intrinsic space of the sound material.

[3]Sections identified through spatio-structural analysis do not necessarily have to correspond with sectional division based on criteria other than space. The usefulness of the notion of intensity-layers becomes apparent in this circumstance.

Non-directional sounds are common in this category of work. There is a tendency to employ sounds without clear attacks. The attack is a very important cue in sound localisation as well as in source identification, and without it sounds are likely to become non-directional. This, combined with a widespread use of relatively long and fairly slow moving sounds, diminish the amount of information in extrinsic space and directs attention towards intrinsic space and spectral distribution. Long time-intervals between the onsets of sounds contribute to reducing the significance of relationships in spatial placement among sounds, and focus on intrinsic space is thereby maintained.

Stretched sounds are often used in works with an intrinsic orientation. Barry Truax points to granular time-stretching as "the single most effective approach... [for] ...shaping the space inside the sound" (Truax 1998: 142). Time-stretching techniques can direct attention towards the sound's intrinsic space by emphasising and enhancing its magnitude and spatial shaping over time. Such techniques can also make the sound's inner spectral components become more apparent to such an extent that they can be heard as individual streams[4]. With granular time-stretching in particular, it is relatively straight-forward to decorrelate the signal in two or more loudspeaker channels in order to spread the sound spatially in a stereo or multi-channel space. This can be done with a high degree of control in order to make the spreading more or less subtle and more or less gradual.

The slowing down of a sound by means of time stretching often makes it easier to hear how its shape varies through time. An even more dramatic spatial effect is achieved if downward transposition is used to slow down sounds as this adds to an increase in magnitude when the concentration of spectral energy moves down in frequency. Moreover, due to limitations in frequency range in the current standard of digital audio technology[5], a sound transposed down by merely changing the sample rate may sound duller because the highest frequencies originally present in the digitised sound, which earlier were at the upper limit of our hearing, now are lower, leaving nothing in the very highest audible range. Depending on the nature of the sound, its relative loudness and the context it is in, this can make it appear muffled or more distant. An example of this can be heard at 3'21" in the middle section of my work *(dis)integration*, which begins with a medium crash cymbal transposed down several octaves by changing the sample rate. The resulting sound is more reminiscent of distant rumble or thunder than a close-up human gesture.

Intrinsic orientation is not restricted to long sounds. Short, repeated material can put the same emphasis on intrinsic qualities as longer sounds provided that the repetitions occur

[4] Compares to Smalley's (1997: 119) notion of internal spectral focus (see section 3.1.2).

[5] I refer to the most widespread digital audio standard here, which at the time of writing is based on the Compact Disc specification of 16-bit 44.1 kHz agreed upon by Sony and Philips in 1979. Digital Versatile Disc Audio (DVD-A), which was launched in 1995, and Super Audio Compact Disc (SACD), introduced in 1999, offer a sampling frequency of up to 192 kHz, and have therefore the potential of reproducing frequencies much higher than the limit of human hearing.

sufficiently closely in time and space so that their relative extrinsic placement and movement do not demand too much of the attention.

Structural variation in intensity in intrinsic orientation is based on changes in density and spectral spreading of the sound material and on how and to what extent the material fills the sound field. Increasing intensity can for example be a result of increasing magnitude or a widening of the stereo field.

Extrinsic orientation

Works of *extrinsic orientation* have a particular focus on placement and movement of the sound material. Attention is drawn towards spatial relationships *among* sounds rather than towards intrinsic spectral and spatial variations of individual sounds. In some ways, extrinsic orientation is the opposite of intrinsic orientation, but combinations of the two are not uncommon.

Extrinsic orientation is based on my notion of extrinsic space, and is defined by attributes that emphasise and "clarify" the extrinsic space of the sound material. For example, a short sound with an unambiguous attack (i.e. reverberation and diffusion are low) can quickly and easily be localised, while other perceptual processing of the sound, such as source recognition, may require more time and cognitive effort. In a musical context with many events rapidly occurring, there may not be time to process cognitively other attributes of the sound, and localisation thus becomes the dominant element. Most often in works of this category, short similar sounds are grouped together to form compound sound events (streams), and as such the source recognition and frequency detection processes are not based on the information from a single short sound, but becomes an interactive process that takes place over the course of the event. Even so, it is location which tends to be detected first, but may come in the background during the cognitive processing of other attributes if the sound is not immediately recognised and the investigation process becomes a conscious activity. When the sound has become familiar, localisation again might take precedence.

Although accurate localisation is often a concern in extrinsically oriented composition, it is not always the specific locations of the individual sounds that are significant. It is rather the distinction between locations that are important, and thereby the establishment of spatial relationships among the sounds. As the sound field becomes more densely populated, distinguishing between individual sounds becomes increasingly difficult and can reach a point where the extrinsic spatial relationships among them are lost. When this happens, the sound material is more likely to be appreciated for its intrinsic qualities as a fluctuating sound mass than for the extrinsic qualities of its individual components.

Sections containing rapidly moving longer sounds represent a border case between intrinsic orientation and extrinsic orientation. Movement and placement are important and help define virtual spaces in the composition, but intrinsic qualities also attract attention as the ear is given time to investigate intrinsic space in the sound material. Attention may drift between intrinsic and extrinsic spaces when listening to such material.

Intensification in works of extrinsic orientation may be a result of, for example, increasing speed of movement of the sound material, increasing density of sounds and events in the sound field, expansion of the virtual space by extending the range of spatial movement or placement, or placing the sound events in closer proximity to the listener.

Environmental orientation

Works of *environmental orientation* make use of recordings of recognisable natural or cultural environments, or simulations thereof, which may trigger knowledge the listener has from past experience. Environmental cues are composed into the work in order to create real, unreal or surreal sonic environments (Wishart 1986: 47-48). I will term such a sonic environment a *referential space*. Familiar environmental cues can indicate the possible dimensions of the virtual space of the work as well as the listener's position (point-of-view) relative to it. Outdoor environmental cues in particular can suggest a much larger space than can be implied by enclosed spatial environments or human-made sounds alone. An example is the opening of my composition *Ebb*, which begins with a recording of calm water lapping which quickly, but gradually, turns into a huge ocean surf, and thereby sets the scene of a vast virtual space.

Environmental references presented in acousmatic contexts can provide associations not only with physical spatial settings, but also with other properties related to these settings, for example of social, psychological or historical kind. Environmental sounds are particularly effective with regard to metaphorical content due to the meaning they inherently carry from the real-world context in which they were captured. Barry Truax states that "environmental sound acquires its meaning both in terms of its own properties and in terms of its relation to context" (Truax 1996: 52). He argues that one cannot escape the inherent contextual meaning of environmental sounds, and that this represents a major obstacle in their use in music. Risset (1996: 37) agrees that the inclusion of environmental recordings in electroacoustic composition represents a dilemma, particularly because "the richness and strength of identity" of natural sounds make them difficult to transform without reducing their content and quality. Yet, recognisable sounds of natural origin are frequently found in electroacoustic works, especially in acousmatic compositions descended from the tradition of *musique concrète*—often with great success.

A unique feature of the electroacoustic medium is the possibility to *recontextualise* sounds, i.e. add, change or create new meanings to sounds by putting them in different contexts. Merely placing a natural sound in a musical context gives it a role that it did not have in its original environment. Sounds which are unlikely to be heard together in nature can be juxtaposed into what can be described as an ecologically unrealistic landscape, but which may be perceived by most listeners as "real". If, on the other hand, it is obvious that the sound sources cannot realistically exist together or they display "impossible" behaviour, listeners are likely to experience the composed space as surreal (or "unreal") (Wishart 1986: 47). It is then the inherent meaning in the sound material which gives rise to the sense of surrealism through its function in the musical context. The point of using recognisable environmental sounds in musical compositions must be because their inherent referential nature and associative power are desirable and relevant in the compositional context, and as such, I will argue, they do not represent any form of obstacle, but, rather, a great resource for artistic exploration. Below (in section 4.3) is a spatio-structural analysis of my work *Intra*, in which there are several examples of the utilisation of the inherent meaning and real-world references of natural sounds and their integration into an underlying compositional idea and musical structure.

The spatial behaviour of natural sounds can be such that compositions of environmental orientation can have an emphasis on either intrinsic space or extrinsic space or both, although this may be overshadowed by the semantics of the sound material.

Spectral orientation

The category of *spectral orientation* includes compositions where the sounds' place in spectral space is the main concern. Physical placement and intrinsic space become less important in the spatial structure of the work. On the extreme of this are compositions where intervallic relationships among pitched sounds is the dominant organisational principle, and where timbral qualities and other spatial attributes have little or no significance as structural means. The compositional organisation of such works tends to have more in common with note-based instrumental and vocal music than the sound-based structural principles typical of most electroacoustic music. The spatio-structural approach may therefore not be of much use when considering compositions of this kind. In tonal music, though, the metaphorical function of space[6] remains subordinate to the structural principles of tonality. Composers of electroacoustic music are free from the constraints of tonality and the limitations of acoustic instruments and human performers, and can treat pitch and frequency content in radically different ways, allowing for variations, combinations and exaggerations of the material which are impossible in acoustic music. The metaphorical use of spectral space in electroacoustic music therefore has more possibilities than in other musics[7]. How-

[6]Often referred to as *pitch space* in connection with instrumental and choral music.

[7]The visual human presence and the symbolism associated with the sound, playing technique and physical presence of the different orchestral instruments can only partially be exploited in electroacoustic music.

ever, because any sound may be used in electroacoustic works, real-world references can be used directly in the form of recorded sound instead of pitch-based metaphors. Abstractions into pitch-space may sometimes be desirable within the compositional context, and there is no reason why they should not then be used. A referential space in the form of real-world sounds will most likely fall under the category of environmental spatio-structural orientation.

Timbre-based works consisting of sounds that are leaning towards "pitchness" represent a border case of spatio-structural orientation. The timbral qualities of such sounds are obviously very important and may be appreciated in terms of intrinsic space, but the sounds are compositionally organised according to their place in spectral space. A clear vertical layering of the sound material in spectral space is common in this kind of composition.

Works of spectral orientation can be of strong intrinsic or extrinsic orientation. Pitch-based compositions are more likely to be extrinsically oriented, while non-pitch based spectrally oriented works can belong to either category.

4.1.3 Space on a micro-level: Spatio-structural functions

On a detailed structural level, relationships among sound materials are established on the basis of the nature of the sounds themselves—their spectromorphological qualities as well as associative qualities regarding source and context—and on the basis of the sounds' spatial behaviour relative to each other and to the space they are in.

The virtual space in which sounds are situated is an important factor in connecting the musical material structurally. Sounds that are not perceived as related may not inhabit the same virtual space. The structural relationships among the sounds must then be considered on the basis of the relationships among the spaces in which they are situated. Different virtual spaces can be juxtaposed in electroacoustic music, and a juxtaposition of spaces into a multi-spatial sound field can be a powerful structural device. Virtual spaces composed into the work can change and undergo transformations during the course of the work. This can take the form of gradual changes, abrupt switching between spaces, a sudden juxtaposition or a switching off of different virtual spaces. The composer thereby has the possibility to create effective contrasts as well as structural layering of virtual spaces.

Dimensions of composed space

Composed space is a three-dimensional space where the sound material can be placed in the horizontal, vertical and depth directions. The vertical and depth directions of an electroacoustic sound field are not always of a physical kind, as is the case in conventional two-

two-channel stereo[8]. In this case, the vertical direction is defined by placement in spectral space and by associations with height in referential sound material. Similarly, depth—the sense of space extending beyond and behind the loudspeakers—is based on distance cues that are either captured in the recording of the sound material or incorporated into the work during the composition process.

Spatial depth and distance are central elements in defining the dimensions of virtual space. With conventional mixing techniques, the listener's physical distance to the loudspeakers defines how close to the listener sounds can be placed. Techniques that make sounds "stick out" from the loudspeakers can help expand this dimension in the direction of the listener, and allow for a sound placement that can be quite intimate[9]. Distance, intimacy and the expansion of musical space into the listener's psychological private sphere can be a powerful means in musical expression and communication, particularly in connection with referential sound material.

With conventional mixing techniques, the horizontal direction is defined by the distance between the loudspeakers, although techniques exist which psychoacoustically expand the width of the sound field beyond the span of the loudspeakers[10]. Combined with placement in the depth dimension, the boundaries outlined by the loudspeakers can be blurred and the sound material can appear to inhabit a larger space than the span between the loudspeakers suggests. However, if the loudspeakers are placed close to side walls, these may act as perceptual boundaries in the horizontal direction and counteract any virtual extension of the width of the sound field.

Motion

Spatial movement of sounds is an important intensifying device in electroacoustic music. Direction and speed are the two components of spatial movement. Different directions of movement can provoke different psychological reactions in the listener, and thereby be of different structural significance in the compositional context. For example: sounds that move away from the listener provoke different reactions than sounds moving towards the listener, sounds moving upwards lead to different associations than sounds that are falling, and sounds that are orbiting around a virtual centre have a different function from sounds that are oscillating from side to side. In a configuration of several moving sounds, sequences of spa-

[8] To take it further: in the case of a single loudspeaker, the horizontal direction is also virtual, as illusions of a horizontal space wider than the loudspeaker cone can be made. A single-loudspeaker playback system, however, is unrealistic for serious music listening nowadays.

[9] Certain manipulation of phase can make sounds appear in front of the loudspeakers, provided that the listener is positioned in the exact centre of a stereo array.

[10] *QSound* by QSound Labs Inc., for example, is one of the most widespread technical solutions for expanding the stereo image into a quasi 3-D sound field. This particular technique is based on frequency dependent adjustments of the differential phase and amplitude of the two channels of a stereo source.

tial counterpoint and other combinations of different movements can effectively underline or counteract the spatial element as well as other musical elements, and act as intensifying or deintensifying devices in the compositional structure.

The fundamental directions of possible movement correspond to the three physical dimensions of space: horizontal, vertical and distance. The basic linear movements are:

- Left to right and right to left
- Upwards and downwards
- From the distance and into the distance

Combinations of these movements form diagonals and non-linear paths. Theoretically, any path of movement can be composed into a work, but not all sound paths are audible. This is both due to the varying accuracy of our spatial hearing in different directions and to the way sound radiates and interacts with the environment in which it exists. The preceding discussion on spatial hearing and localisation blur shows that there are perceptual limitations which must be taken into account in spatial composition, particularly in terms of extrinsic space. Elaborate and accurate placement and movement of sound events in the listening sphere, such as those outlined by Stockhausen (1961: 77-82) and Wishart (1996: 206-222), do not always make sense perceptually. Clearly, the type of sound involved is a determining factor, so that any absolute statements about what works or what does not work cannot be made. Generally, though, as has been pointed out earlier, pulse sounds can be localised much more accurately than static sounds. It is therefore no coincidence that so many demonstrations of three-dimensional sound systems are done with helicopter sounds.

Speed and energy

The distance and speed at which sounds move as well as the number of concurrently moving sounds contribute to the activity level and sense of energy in a work. These are important factors when considering intensity direction, for example. In terms of speed, a moving sound can have one of the following characteristics:

- constant speed
- accelerating speed
- decelerating speed
- irregular speed

A sound that moves at a constant speed has no change in energy input. Accelerating and decelerating sounds may change speed at a constant rate, and the rate of energy input or energy loss in this context is therefore likely to be constant.

The connection between speed behaviour and structural function is not always clear-cut. Increasing speed does not necessarily mean increased intensity, although this seems to be the norm[11]. The interpretation of any spatial activity must be carried out on the basis of the context in which the spatialisation takes place, and must be seen in connection with the nature of the sounds or sound configurations that are being considered.

4.1.4 Significance of spatial detail

Musical space is defined by the sounds in it and, hence, cannot exist on its own. Nevertheless, on a detailed level it is possible in specific contexts to compose sound materials in such a way that spatial organisation has structurally more significance than the choice of sounds. For example, when certain musical material is presented from a certain place in space and this place is used at key points in the composition, then the location of the sounds can be what attracts attention, and thereby becomes more important than the sounds that are in that location. This is difficult to maintain over longer periods of time, but can occur repeatedly throughout a work. The context in which spatial information is heard is crucial, and gives the listener cues as to how significant the spatial parameter is in the compositional structure.

4.2 Stereo and surround-sound composition

There are significant differences in the composition of stereo and surround-sound works. Apart from the technical issues of choosing among the various surround-sound formats, the composer is faced with a number of additional challenges when composing for more than two audio channels. A main reason for composing a surround-sound work must be one of artistic necessity where the compositional idea calls for more acoustic space for its realisation than two-channel stereo can deliver. Surround-sound works that in reality are pre-diffused stereo compositions may be good for private listening in the studio or at home, but not necessarily successful when transferred to the larger concert space. Some compositional ideas function best as stereo works which can then be further enhanced and spatialised in concert diffusion. Some practitioners of acousmatic music hold that concert diffusion is an extension of composition, and that the composition process is only complete when the work is realised in the concert hall (Barrière 1998; Boeswillwald 1998: 229).

[11] For example, increased speed in combination with a gradual reduction in loudness and attenuation of high frequencies indicates that the sound is moving out of the sound field, and thus reduces its contribution to the overall intensity level.

Whereas stereo relies fully upon the virtual space, where a multi-dimensional space may be implied and prepared to be realised in concert diffusion, surround-sound composition adds the dimension of a real space. A work composed for the full listening sphere has the potential of more complex spatial treatment than a stereo piece does. However, there is a danger of "overcomplexifying" the composition by thinking of different output channels and the three-dimensional space in a visual manner and work out spatial paths and trajectories in an absolute, quantified way. For instance, treating each point in a circular loudspeaker configuration as equal will only make sense in terms of perception if the listener is expected to move around within the circle. Otherwise, the aforementioned dissimilarities in localisation accuracy in different directions may obscure spatial detail.

Common delivery formats for music as well as the design of professional audio software and hardware are generally defined by the commercial music industry. The embracing of the widespread stereo format as the "fixed medium" of electroacoustic music has very much shaped the way composers think about their work in terms of composition and diffusion[12]. The relatively recent introduction of surround sound for home listening in the forms of SACD and DVD-A may challenge that way of thinking, and may instigate the development of the electroacoustic genre towards one where physical space becomes increasingly focal in the composition process. Although two-channel stereo is a central part of the specifications for these new audio formats, stereo is likely to be phased out as more and more commercial music releases are mixed in surround, and the record-buying audience becomes accustomed to the surround-sound listening experience. Chapter 5 deals with sound systems and recording formats in more detail.

4.3 Case study: Spatio-structural analysis of *Intra*

The idea for my acousmatic composition *Intra* comes from my general interest in psychology and my fascination for psychopathology in particular. Some reflection on the title will give a hint in this direction: the Latin word "intra" means "inside" or "within." The exploration of the human psyche in art has long been a strong interest of mine, and is a recurring theme in much of my work. However, it is more often a "hidden" structural aid during composition than a concept which is meant to be immediately apparent on the surface of the finished work.

The structural scheme for *Intra* is based on descriptions of some of the symptoms encountered in different stages of schizophrenia[13]. A key symptom of this psychopathological

[12]"Diffusion" is here meant to mean two things: the act of diffusing an electroacoustic work in the concert hall and the distribution ("spreading") of the music to the general public in the form of (at the time of writing) CDs and radio broadcast.

[13]From the Greek *schizo*, "split," and *phrene*, "mind." Refers to one of the main attributes of the disorder: a

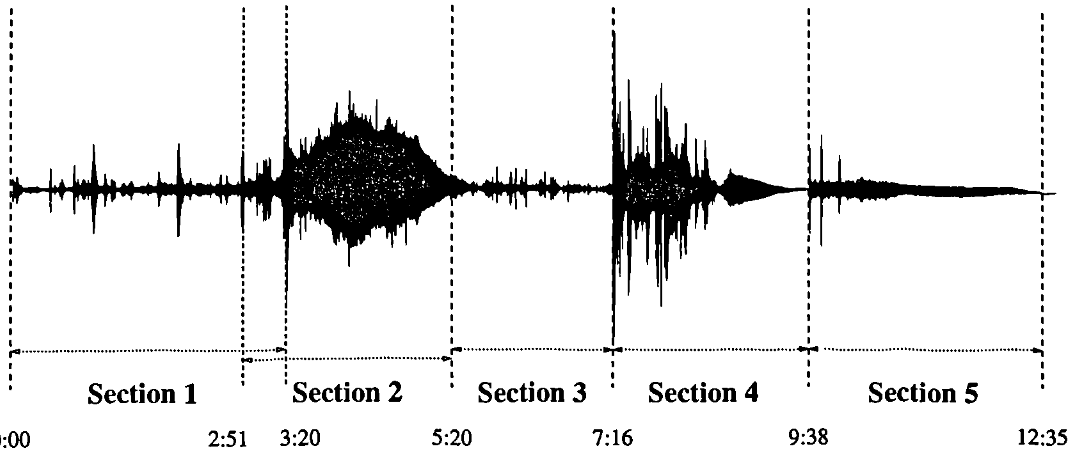


Figure 4.2: Time and amplitude view of *Intra.*

condition is a withdrawal from social relationships. Such a withdrawal may result in a lack of opportunities for *social-reality testing* in which one's own ideas are checked against those of others and corrected when necessary. The long term consequence of this behaviour is a condition where the schizophrenic can no longer distinguish between fantasy and reality, and has lost the ability to differentiate between internal and external worlds. Such a state often leads to hallucinations where the schizophrenic perceives experiences that occur in the absence of any actual sensory stimulation. It is fairly common for schizophrenics to "hear" voices that are not actually there and accredit them to known or imaginary persons or creatures. Another key symptom of schizophrenia is a disturbance of attention, which manifests itself in a difficulty in suppressing irrelevant external stimuli, so that the schizophrenic hears (and sees and feels) too much. Selective attention is very ineffective, something which leads to severe overstimulation. Noises seem louder, and it is difficult to discriminate foreground sounds from background sounds.

By utilising some of the spatial possibilities of the electroacoustic medium, I was able to translate this extra-musical idea into a useful model for composition. An obvious tool is the construction of virtual spaces to represent fantasy and reality as internal and external environments. This is a central technique on a high structural level in the work. However, there is not always a clear separation between the internal and external environments in this composition—they are sometimes combined and sometimes obscured by spatial components on other structural levels. Thus, the work is not based on a single spatio-structural orientation all the way through, but on shifting between orientations as well as combinations of different orientations.

A rough examination of a standard graphical representation of the work (fig. 4.2) sug-

fragmentation of mental functions. It does not—as is commonly believed—refer to a split personality.

gests that it can be divided into five sections. The sections alternate between soft and loud and span a great dynamic range. Listening to the piece in terms of overall spatio-structural orientation confirms that a division into five sections makes sense. The first section is of environmental orientation. There are several cues that indicate an external environment, although it is somewhat surreal. The second section is of intrinsic spatio-structural orientation. It starts at 2'51" and overlaps the first section by about thirty seconds. It has a transitional function, and shifts the point-of-view in the piece from an external, observing one in the opening section to a more introspective one in following sections. Section 3 begins at 5'20", and is a combination of extrinsic and intrinsic orientations. In relation to the structural model for the work, this section combines internal and external environments. The fourth section, from 7'16" to 9'38", represents an internal environment, and starts out in extrinsic orientation, but shifts gradually towards spectral orientation. The spectrally oriented structure continues in the last section. However, this section is somewhat ambiguous in respect to spatio-structural orientation as it also contains intrinsically oriented elements that are quite prominent throughout.

Intra is composed in two-channel stereo and intended to be performed on a multi-channel sound diffusion system. The total duration of the work is 12'35".

Low level spatio-structural analysis

Section 1

The first section spans the first 3'20" of the piece. The spatio-structural orientation of this section is environmental. The most easily perceivable environmental cues are perhaps non-melodic bird calls that occur frequently throughout the section, as well as a rhythmic frog sound towards the end. There is also a quasi-melodic event that occurs three times in this section that can, in the current context, be interpreted as a surreal form of bird call. The consistent environmental orientation in this section indicates the important function of environmental space in the work. Furthermore, its surreal elements suggest that the environment is not entirely external: fantasy and (perhaps) hallucination are involved as well.

From the onset of the section virtual space is quite wide and deep. In the horizontal direction the space spans the entire stereo width, and similarly extends quite far in the depth direction. However, the loudspeakers are not pointed out, thereby suggesting a vast virtual space that is not confined by the physical distance between the loudspeakers.

The vertical span, as defined by spectral space, is in the opening limited to a band in the mid- to high-frequency range. This expands stepwise downwards at two points: the first at about 0'30" when a low mid-range drone fades in, and the second point—which perhaps has more immediate impact—at 0'45" when a low bass figure enters.

The non-melodic bird calls that are heard throughout the section are particularly prominent in defining the boundaries of virtual space in the depth direction, and to a certain extent in the horizontal direction. The bird calls are stationary and directional, and occur in various locations spread out over a large area in virtual space. They provide directional cues in terms of spectral content as well as reverberation. Two noise-based drones contribute to the sense of depth as well: the first spans from 1'10" to 1'43", the other from 1'50" to 3'02". The second noise-based drone has more low-frequency energy and is of greater magnitude than the first, but is located further back and is therefore more influential in defining the depth of the space.

Throughout the section the bass is relatively unstable, and in that sense conforms to the underlying idea for the work. Its shape is quite uneven: it fluctuates in amplitude and varies continuously over a small frequency range, leading to a sense of constantly varying magnitude. Similarly, its spatial position alters in an unsteady, fluid manner. The most stable elements in this soundscape are the pitch-based mid-range drones from 0'30" and 1'10".

Considering the section as a whole, we see that its structural function is twofold: one is to establish environmental space as a central element in the work, and the other is to increase gradually the level of subjective intensity. What contributes to the intensification is, in particular, a gradual increase in the number of sounds in the sound field, but also the bass, which is especially prominent being at the low extreme in spectral space, becomes increasingly fluctuating in spatial movement as well as in loudness. Its magnitude is gradually (but not smoothly) increasing, and in that sense greatly contributes to the heightened density of the sound field.

Section 2

The intensification of the sound field in terms of density reaches its maximum in section 2. This section starts at 2'52" and ends at 5'20". It is the most dense part of the composition. It has the overall shape of a crescendo-decrescendo, and is dominated by a fairly broadband sound cloud which is present throughout the section. This noise-based sound cloud is used as a metaphor for the schizophrenic's loss of ability to discriminate between foreground and background and the resulting overstimulation of the senses.

The overall spatio-structural orientation of section 2 is intrinsic. The dense, broadband sound cloud fills the sound field in all directions and can only be appreciated for its intrinsic qualities. Other sounds that contribute to the overall spatio-structural orientation are stretched bell sounds and mid-range drones of definite pitches. The fluctuating bass gradually transforms into a steady low-pitched drone early in the section and provides the virtual space with a much more stable foundation than earlier. This low-pitched drone is present throughout this section and into the next. Together, these sounds—the drones and the sound cloud—form a very present background which almost drowns the foreground material.

The sounds in the foreground all have human characteristics. The first sign of human gesture is at 3'20". Although the sound source is not recognisable, this sound has a sense of "forced" gestural movement that stands out in contrast to the more natural or organic sound-world that so far has characterised the piece. This "human" sound is located relatively close in the foreground, and consequently indicates a shift in the listening point-of-view from one of distant observation established in the first section to a more intimate, participatory one. The most prominent feature of this particular sound is its extrinsic space, which manifests itself in a close-up placement in the sound field as well as in a wide horizontal movement of the sound.

The footsteps that appear at 3'42" are an important element in the work. The change in perspective anticipated by the human gesture at 3'20" is here confirmed. This is the sound of someone walking, but the footsteps are not moving spatially in the sound field. They are located in the centre, and appear to be heard from the point-of-view of the person walking. This section provides a transition from external world to internal psychological environment, and therefore changes the perspective towards an introspection. Another striking element is a voice that first appears at 3'57". Because this voice is moving horizontally across the sound field it cannot come from the person who is walking. The shift in perspective towards an internal psychological environment suggests that it is a voice *imagined* by the walker. Again, this refers to the extra-musical idea behind the work.

Section 3

The voice elements continue into the next section, which starts at 5'20". Section 3 is a mixture of internal and external environments. Its external cues include variations on some of the elements of the environment heard in the first section in addition to the sound of rolling rocks. In terms of internal cues, the voice elements continue, but they are more fragmented and more spatialised here than in the previous section. The spatial movements of the vocal sounds relate to the movements of the rolling rocks: sometimes they move together in the same direction, and sometimes in countermovements (simultaneous left to right against right to left movements, for example). The spatial relationship between these two sound configurations underlines the fusion of internal and external worlds at this point in the composition.

In the background there is a sound that can be interpreted as slow, heavy breathing, whose function is to emphasise the intimate human presence at this point in the piece. This sound also contains distance cues that maintain the sense of a large space.

This section combines extrinsic and intrinsic spatio-structural orientations. The directional moving sounds, such as the voice and the rolling rocks, contribute to the extrinsic orientation, while the stretched bells and the background breathing sound contribute to the intrinsic spatio-structural orientation. The overall intensity direction of this section is relatively flat (stasis). Its function in the composition is to provide stability and rest, and not create any anticipation of change in any specific direction, thereby establishing a certain listening mode which generates a state of calm. However, the duration of this part is of such a length that its overall structural function may turn from stasis to latency for some listeners, i.e. it moves towards an expectation of change. Nevertheless, there is no hint or forewarning as to when something may happen.

Section 4

In the light of the listening mode established in section 3, the sharp attack at the onset of section 4 comes unpredictably and has a strong impact. The level of spatial activity here is high, and makes this part the most energetic in the work. Most of the section is characterised by directional sounds that are distributed all over the virtual space. It is a very dense part in terms of sound population, but it is generally not difficult to distinguish the sounds from each other due to their directionality and separation in spatial distribution. There is a clear focus on spatial movement and placement, and the spatio-structural orientation of the beginning of this section is therefore extrinsic.

The sound material is organised in spatial layers in the depth direction, and for the most part there is little movement between the layers. However, there are a few sounds that move from the distance and up close and some that make wide spatial rotations in the virtual space. It is characteristic for this composition that there are relatively few wide spatial gestures. Such a stable spatial context, where the emphasis is more on placement than motion, heightens the effect of the few broad spatial movements that are there. To go back to the schizophrenia-idea, this refers to one way of dealing with overstimulation: some schizophrenics can maintain a specific posture for hours in an attempt to keep the sensory stimulation stable. Because, if one moves, then the senses receive constantly varying information which can lead to severe disturbance to an already overstimulated mind. It is such a severe disturbance that section 4 portrays.

Considering section 4 as a whole, it is apparent that its spatio-structural function is one of decreasing intensity. The number of sounds populating the sound field is gradually reduced, and as the section progresses the virtual space transforms into a smaller, more intimate space. Towards the end of the section, as the sound field becomes increasingly sparse, the spatio-structural orientation shifts from extrinsic to spectral as the section ends with a pitch-based harmonic structure that continues into section 5.

Section 5

The last section of the piece is in a sense a variation of the opening section. The stretched bells and the non-melodic bird calls from the opening of the piece are present and help define the size and nature of the virtual space. As in the beginning there is a stepwise widening of spectral space, starting here with a downward expansion at the very onset of the section at 9'38". The second attack eight seconds later widens the space upwards when the stretched bells enter, and the third attack 20 seconds into the section introduces a noise-based cloud which fills the mid- to high-frequency band and establishes the spectral height of the space. There is a variation on the fluctuating bass line, which here appears higher in pitch and smaller in magnitude, and is therefore less dominant in the sound field than its incarnation in earlier sections. Below the bass line, with a somewhat distant location, there is an unpitched low-frequency drone which goes through the entire section and eventually fades out as the piece ends. This drone is quite stable and provides a firm foundation for the section.

This section is a combination of spectral orientation and intrinsic orientation. There is a clear vertical layering of the sound material in spectral space, but some of the sounds contain distinct intrinsic spaces that may demand attention. There are sounds with strong extrinsic presence, but in the spatial context these are not powerful enough to influence the overall spatio-structural orientation.

This is the last section of the piece, and its overall direction of intensity is a decreasing one.

Conclusion

In my discussion of the spatial organisation of *Intra*, I have shown one example of how space can be the fundamental structural device in electroacoustic composition. The framework of spatio-structural analysis is primarily a descriptive tool, but represents an approach which can be useful for the composer in developing an awareness of spatial potential in sound material, as well as aid in the organisation of composed space. A complete analysis of a musical composition, however, needs to be grounded on a consideration of all the parameters that are present in the work. The context in which the spatial information is heard is crucial as it provides cues as to how to interpret the significance of the spatial parameter in the compositional structure. Spatio-structural analysis must therefore be used in combination with frameworks set up for investigating other musical parameters, such as pitch and temporal organisation, in order to be in a position to examine all the aspects of the work.

5 Listening space

New halls for listening must be built to meet with demands for spatial music. My idea would be to have a spherical chamber, fitted all round with loudspeakers. In the middle of this spherical chamber, a platform, transparent to both light and sound, would be hung for the listeners. They could hear music, composed for such adapted halls, coming from above, from below and from all directions.

Karlheinz Stockhausen (1961: 69)

The physical space in which the music is played has a significant influence on the performance and the experience of music whether it is acoustic, acousmatic or amplified acoustic. While instrumentalists can adjust tempo, phrasing and loudness in order to compensate for the acoustic conditions of the room, electroacoustic music produced on a fixed storage medium is less flexible in that sense, and is particularly fragile with regard to spatial influence because space often has an important function in the structure of the work. Thus, the spatio-structural intelligibility of the work is often dependent on a successful interaction between the spaces composed into the work and the space in which the work is heard. The electroacoustic genre depends on loudspeakers for its existence, and the type, placement and number of loudspeakers and their matching with the listening environment are a major influential factor on the performance practice and on the listening experience of electroacoustic music.

The room and the loudspeakers operate as one acoustic system. The term *listening space* in this context, therefore, refers to the combination of the physical space where the work is presented and the sound system on which it is played. Differences in spatial potential between headphones, stereo systems, surround-sound and multi-channel sound diffusion systems present different possibilities and problems as to how the music is distributed in the room, and ultimately how the musical work is experienced. The best sonic results come from the best combination of music, room and sound system. Whether listening takes place in a private environment or in a public space often dictates the possible choice(s) of listening position relative to the sound system, and influences localisation of sound events as well as the experience of other spectral and spatial attributes of the work. The distinction between public space and private space is therefore important in this discussion.

5.1 Sound and the acoustic environment

When listening in a room, the sound pattern reaching the ear is always different from that emitted from the source. The listening environment acts upon the sound in different ways, resulting in variations in spectral, temporal and spatial distribution. The acoustic influence of the room is always present to a greater or lesser extent, and is a factor that the composer must bear in mind when composing works for public performance. The acoustic influence varies with different rooms, and can therefore not be thought of as an absolute entity in the composition process unless the work is to be created for a specific, known listening space. Even then the acoustic conditions are likely to vary with the size of the audience, and the performance of the work must be carried out accordingly.

The three main phenomena that act upon the sound in the acoustic environment are *reflection, diffraction* and *absorption.* The degree to which each plays a part varies with different physical spaces, and can be used to musical advantage with a sensible arrangement of the loudspeaker system and the listening position(s). Much acoustic research has been published which shows the complexities involved in sound propagation in enclosed spaces. A comprehensive review of that area, however, is beyond the scope of this thesis. I will only outline a few of the matters of room acoustics which are of direct concern to composers of electroacoustic music, and which are relevant for the following discussion on concert performance and listening in private and public spaces.

Reflection

Most of what we hear in our daily existence are reflected sound waves. We are most used to listening in reverberant spaces, and a certain amount of sound reflection is therefore essential in order for the auditory information to sound "right", and in order for our interaction with the acoustic environment to feel comfortable and natural. Everybody who has been in an anechoic chamber has experienced the stifling feeling of not having the aid of reflection to carry the voice when trying to speak, and not being able to use acoustic feedback to monitor one's own speech rate and speech power.

Reflection leads to an increase in total sound energy in the room due to the creation of multiple representations of the sound wave. The number of reflections depends on the number of surfaces in the room as well as the distance between the surfaces (White 1975). Reverberation time[1] depends on the size of the room, the acoustic characteristics of the

[1] Reverberation time is commonly defined as the time it takes from the moment the sound is switched off until

building materials of the floor, ceiling and walls, and on the amount and construction of any furniture and other decorations in the room. Because the reflected sound waves travel different distances, they arrive at the listener's position at different times. Reflections that arrive within 20 ms of the direct sound are perceived as part of the original sound wave, and will therefore directly influence the spectral quality of the sound. Reflections travelling within 50 ms to the listener's position are heard as reverberation, whereas intervals longer than 50 ms result in discrete echoes. In rooms with high ceilings, the first reflections to reach the listener are those which bounce off the walls. Reflections bouncing off the side walls lead to an experienced enhancement of the width of the sound field, and the result is a subjective sense of spaciousness generally regarded by music audiences as a desirable acoustic quality for concert halls (Rasch and Plomp 1982: 145; Adams 1998: 319).

However, too much reverberation comes at the expense of clarity. Clarity is the ability to distinguish spectral and temporal detail in the musical sound. Electroacoustic music sometimes requires a higher degree of clarity in the performance space than the case is for instrumental music. One reason is that the often dry acoustics of the composition studio may entice the composer into implementing sonic detail that is dependent on a low- or non-reverberant listening environment in order to be heard. On a similar account, reverberation composed into a work in response to a dry composition environment may clash with the natural reverberation generated in the physical space and, in extreme cases, result in a spatially detailed work presenting itself as an undefined sound mass. A negative impact of a reverberant space on the listening experience can to an extent be minimised by adjusting the distance between loudspeaker and listener so that the first sound waves to reach the listener are predominantly direct sound. With a multi-channel sound diffusion system it is usually desirable to have loudspeakers placed at different distances to the audience in order to use the acoustic characteristics of the room in a controllable and flexible way to the advantage of the music by enhancing and expanding the effect of distance and closeness that may be composed into the work. Reverberation also contributes to a reduction of the effect of dynamic range in the music, and active level adjustments during concert performance of the work may be necessary in reverberant spaces.

Naturally, whether the effect of reflection is a positive or negative one depends on the nature of the particular work. Reverberation leads to a blurring of attacks and prolongation of the sounds. This is what may conceal musical detail, but it can also create an increased sense of spaciousness that is desired for certain types of electroacoustic compositions. The diffusion of sound in a reverberant space can make the music sound "from all directions" and fill the space to submerge the listener.

Very little reverberation in the listening space can lead to the loudspeakers being empha-

---

the sound pressure level in the room is reduced by 60dB (Winckel 1967: 66; White 1975: 88; Handel 1989).

sised[2]. This can happen when the composed space does not incorporate any reverberation and there is no aid from the room in spreading the sound by reflecting off surfaces. However, this is an extreme which also depends on the extrinsic space of the sound material as well as the listener's position relative to the sound system. If the listener is positioned in the exact centre in front of a stereo pair of loudspeakers, then a short reverberation time contributes to a more accurate stereo image. If, on the other hand, the listening position is even slightly off-centre, a displacement of the stereo image is likely to occur. Asymmetry in the amount of reflected sound from the boundaries near one loudspeaker compared to that of the other can similarly shift the position of the stereo image due to a difference in acoustic amplification around the two loudspeakers. In this case, other undesired effects such as an unbalanced audio spectrum[3] can also occur.

Diffraction

The wavelengths of sounds within the range of human hearing vary between about 2 cm and 20 m. Such dimensions correspond to the sizes of most indoor objects as well as door openings, windows and hallways. *Diffraction* is when sound waves "bend" or spread out when they hit obstacles and cut through openings. This may result in modifications of the sound spectrum and spatial distribution of the sound (White 1975: 45; Benade 1990: 193-94).

Diffraction depends on the ratio of wavelength to object size. Diffraction of light is practically imperceptible to the naked eye because its wavelengths are very small (less than 10^{-4} cm), resulting in sharp shadows. A similar shadowing of sound only occurs for the highest audible frequencies, and—in enclosures of normal dimensions—not at all for frequencies below 250 Hz (White 1975: 45). Furthermore, composite sound waves do not diffract uniformly: low harmonics may bend around the object, while higher harmonics may be refracted in different directions.

We perceive a substantial amount of the sound energy by diffraction. Diffraction, together with reflection, is an important contributing factor to the sense of spaciousness in a listening environment, and is a central element in sound localisation.

Absorption

Absorption is the opposite of reflection. A sound is absorbed when it enters a material in which the vibratory energy of the sound waves is transformed into heat by means of friction.

[2]This assumes a technique based on phantom images where sounds are deployed horizontally by means of altering the relative loudness of the signals fed to the loudspeakers (discrete-channel mixing technique).
[3]Reflections fed back to the loudspeaker diaphragm can to a certain extent affect its vibratory pattern.

The absorption factor[4] depends on the depth and internal structure of the absorbant material, and is frequency-dependent in that certain materials are particularly absorbant for certain wavelengths. High frequencies are generally absorbed first, while low frequencies (long wavelengths) are the most difficult to absorb. The most noticeable effect of absorption is a spectral colouring of the reverberation, fewer reflections and reduced reverberation time.

In rooms where there is a considerable difference in acoustic quality when the room is empty compared to the presence of an audience, this is generally due to an overall low level of absorption in the empty room. In "good" concert halls such a difference is minimised by, for example, installing upholstered seats with high backs and an absorption factor close to that of an average person. This way it will hardly matter (acoustically, at least) whether someone is sitting in the seat or not. In halls where the seats are not fixed, as is often the case in smaller venues, the seats tend to be made out of wood or plastic (light material which is easy to lift and move) which reflects sound when there is nobody sitting on them.

5.1.1 Case study: *Lorry Red Lorry Yellow*—a sound installation

Lorry Red Lorry Yellow is a sound installation created by myself in collaboration with composer Jo Thomas and photographer/writer Betsy Schneider. The work was commissioned by the Central School of Speech and Drama in London for a two week exhibition in their art gallery in January 2000. As the gallery was not acoustically isolated from the rest of the building, sounds from neighbouring rooms spilled over into the gallery to such an extent that it became necessary to incorporate them into the installation in order to eliminate their disturbance. The sound material we used in the work was all based on recordings of the environmental sounds of the building as they were heard in the gallery. Most of the sound material was manipulated in various ways, and the result was a sound pool of recognisable and unrecognisable sounds—all with their origin in the space where they would be played back. The playback was done from two continuously looping CDs of different durations, an arrangement which ensured that the combination of sounds was never the same. The installation played non-stop during the entire exhibition period, even outside of the gallery's opening hours. This made the reverse effect of that during the day: the sounds from the gallery spilled out into the neighbouring rooms of the building and became audible in the evenings when the general sound level in the building around the gallery was much lower—a sort of a spatial "mirroring" of the sonic image from the day time.

The physical space in which the work was installed has a reverberation time of approximately 5 seconds. We had four loudspeakers at our disposal, and to conform with the basic idea behind the work it was crucial to install the loudspeakers so that the sound emitted

[4] A value between 0 and 1 indicates how efficient the material is as an absorbant. An absorption factor of 1.0 means it absorbs all sound, while a factor of 0 implies that all sound is reflected (Mackenzie 1969: 134).

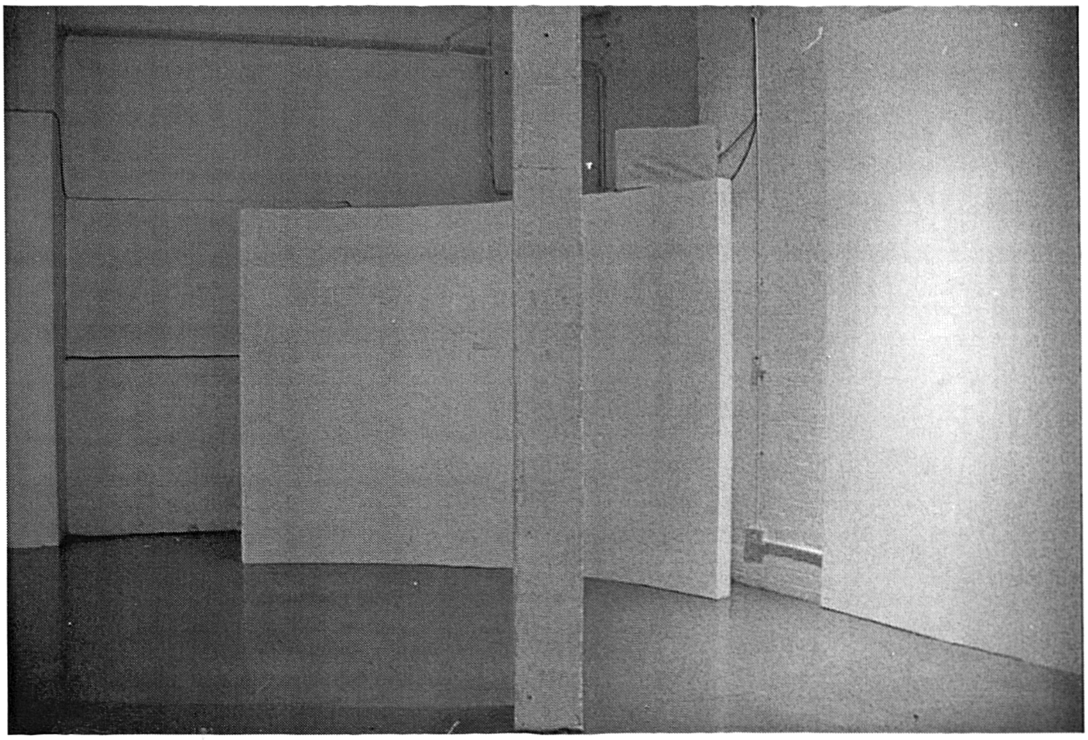


Figure 5.1: *Lorry Red Lorry Yellow*. Covered-up loudspeaker, north-east wall.

from them was as non-directional as possible. This was important in order to make our sound material blend with the sound environment in and around the gallery, and make the two sound worlds as equal as possible in terms of localisation from within the gallery. A long reverberation time is of great advantage in the creation of non-directional sound fields within an enclosure. The most crucial point in order to reduce spatial detectability of the loudspeakers was to install them away from ear height, preferably as high above the audience as possible. The higher up the loudspeakers are, the more of the sound will reflect off the ceiling first. This causes the the sound waves to spread out more before they reach the ears of the audience, and makes it difficult to identify aurally the loudspeakers as the actual sound radiators. Furthermore, placing the loudspeakers as far away from each other as possible helps fill the space and diffuse the sound to a great extent. In addition to this, we hid the loudspeakers by covering them up, and thereby diminished their visual presence, further removing attention away from them. The layout of *Lorry Red Lorry Yellow* can be seen in figures 5.1-5.3.

Figure 5.1 shows how the loudspeakers where covered with fabric of similar colour to the walls. This particular loudspeaker was located on top of a short, loose wall that was set up in

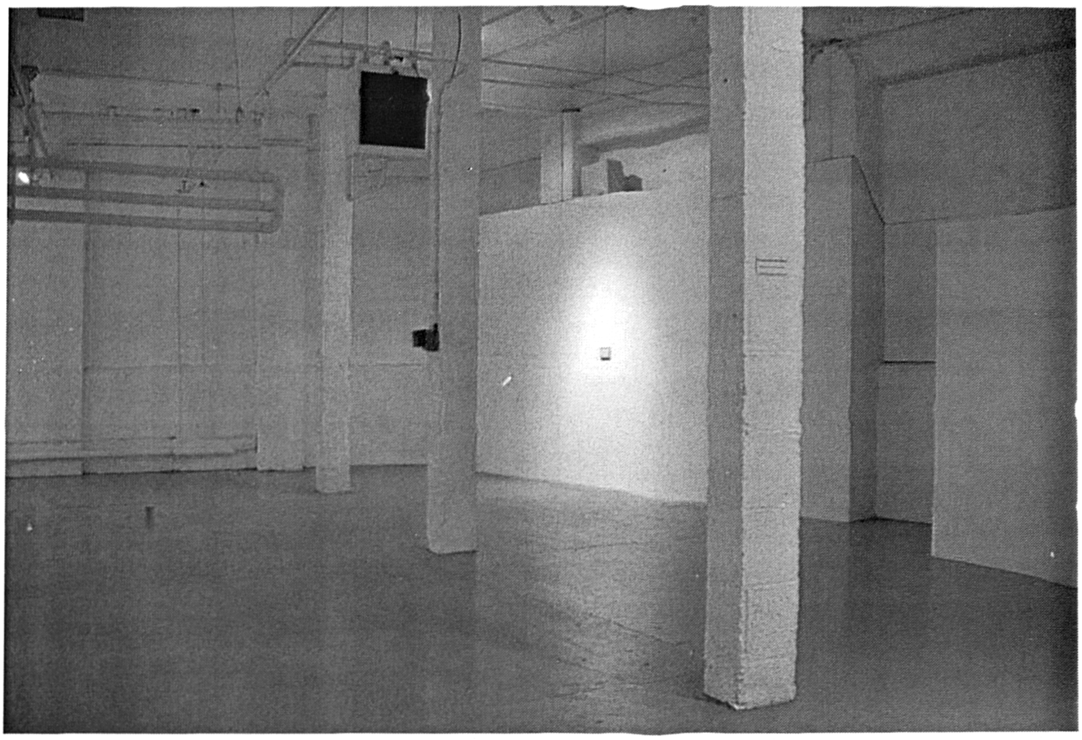


Figure 5.2: *Lorry Red Lorry Yellow*. North wall.

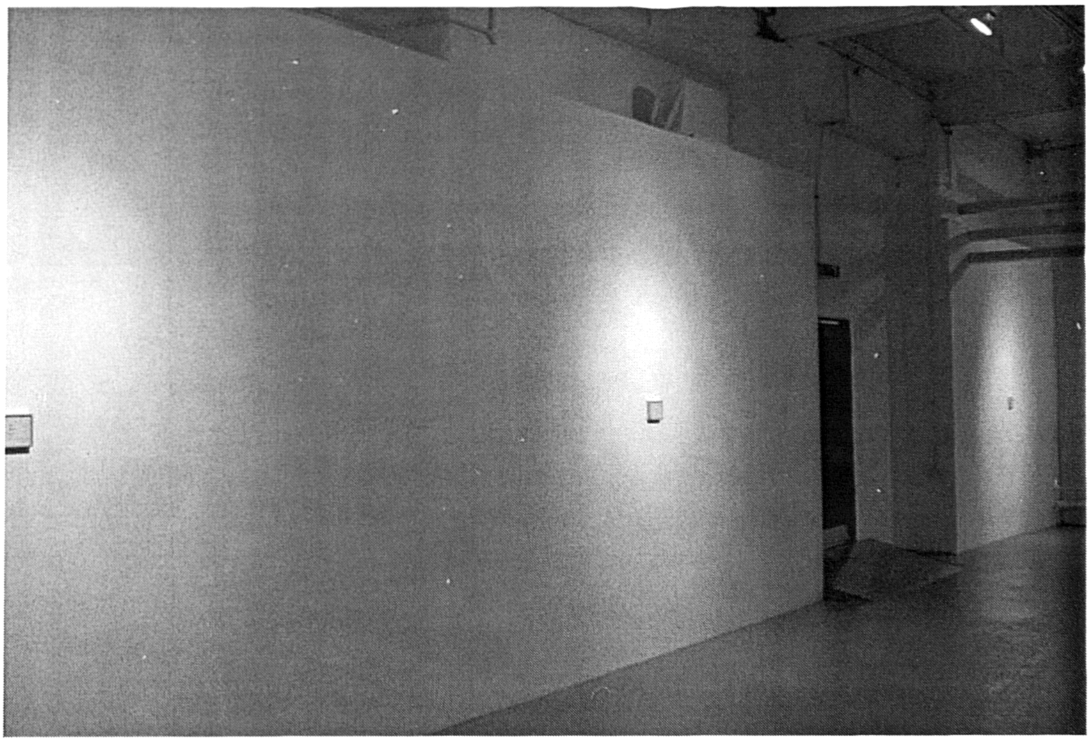


Figure 5.3: *Lorry Red Lorry Yellow*. South wall.

front of the entrance to the gallery toilet area. Recorded sounds from the toilets were among the sound material used in the installation, and the real sounds from the toilet area blended in with the recorded/processed material quite successfully. Figures 5.2 and 5.3 show the north and south walls of the gallery, respectively. Conveniently, both side walls of the gallery had a top "shelf" deep enough to allow the loudspeakers to be placed safely on them. The height of the walls was more than sufficient for placing the loudspeakers above ear level so that the directionality of the sounds emitted from the loudspeakers was minimised. Also in these two pictures one can see the visual part of the installation—a series of small framed pictures, each showing three words related to listening and the experience of sound.

5.2 Listening in private and public spaces

5.2.1 Private space

In the private setting where only one listener is present, this listener is generally free to choose a position relative to the loudspeakers. Assuming that the individual in question is enjoying the music at home, there is the additional freedom of moving furniture and loudspeakers around to set up a suitable listening environment for the available space. Of course, only the most dedicated (or privileged) music lover would arrange a room for the sole purpose of creating the best possible listening environment. More often the listening position and placement of the home sound system are compromised by other factors such as space restrictions, other functions of the room and visual-aesthetic considerations. It is difficult to generalise in a discussion on the private listening space because of the unpredictability of such factors. However, certain observations can be made.

Acoustically, the home listening environment can be quite advantageous for electroacoustic music. In most homes, the listening room also serves as the general living room and is furnished as such with curtains, carpets and upholstered furniture. The materials of such furnishings are often efficient sound absorbants, and can help reduce reverberation quite significantly. Since high frequencies are more easily absorbed than low frequencies, the spectral balance of the sound field favours the low end of the frequency spectrum. This reduces the brilliance of the treble and results in a certain "warmth" or richness in the bass. Furniture and other objects cause diffraction which in turn may increase the sense of spaciousness in the room, depending on the amount of absorption. The acoustic effect of living room layout and decoration varies with culture and fashion. For example, studies have shown that the average reverberation time[5] in Swedish living rooms is 0.55 seconds while it is 0.35 seconds in American and British living rooms (Toole 1998: 471). As a reference it should be mentioned that the minimum reverberation time accepted for extensive listening

[5]Measured in the mid-frequency band, 300-3000Hz.

tests[6] is 0.3 seconds. Anything below this amount seems unnatural and can make listeners feel uncomfortable (Toole 1998: 472).

The common distribution of furniture in a normal room is such that most of the sound absorption takes place in the lower third of the room. This means that there is more reflection higher up in the room. How high up the loudspeakers are placed, their distance from walls as well as their direction therefore determine the effect of reflections and absorption of the sound. The higher up the loudspeakers are placed, the more reverberation will be generated. Because low frequencies are the most difficult to absorb, these will mostly reach the listener as reverberated sound. Mid-frequencies will be heard as a mixture of reverberant, early-reflected sound bouncing off immediately-adjacent boundaries and of direct sound. High frequencies are mostly heard as direct sound. The direct-sound component is influenced only by listening distance and the orientation of the listener with respect to the loudspeaker. The reverberant sound field in the room is dictated by the overall acoustic properties of the room and by the directivity of the loudspeaker (Toole 1998: 470-71).

5.2.2 Public space

In a public listening environment the choice of listening position is often restricted. Although concert auditoriums are normally arranged so that the audience can enjoy the best possible listening positions, the individual listener may find the best seats occupied by other audience members, or in some cases even by a mixing desk. There can be a considerable difference in listening conditions between the "sweet spot" and other listening positions in a public listening space. The consequences of the listening position depends on the shape and size of the hall and the type and placement of the sound system. Any unfavourable listening conditions can often be rectified by a well-balanced installation of the sound system with the aid of sound absorbants and diffusors. However, equal listening conditions for every member of a larger[7] audience group is impossible to achieve. Some compromise must always be made in order to create a spatial average of the highest possible quality for as many listeners as possible.

With a great distance to the sound source, temperature and humidity variations and air movements create continuous variations in the signal reaching the listener. This can cause varying phase cancellations and spectral peaks, sometimes leading to the perception that the sound source is moving or that it seems to originate from a different place than it actually does. Localisation is therefore difficult over long distances. This is mostly a problem outdoors, but can also occur in large concert halls.

[6]Guidelines from *The International Electrotechnical Commission*. "Listening Tests on Loudspeakers." In Publication 265-13 (1985). Referenced in Toole (1998).

[7]What constitutes "large" in this case is relative to the physical room in which listening takes place.

The acoustic influence an audience as a group has on the individual listener in that group can be quite significant, including the ability to locate sound in the listening space. As mentioned earlier, head movements are a useful aid in locating sounds in space. This is because by moving the head one takes advantage of the extra sound information provided by acoustic cues such as reflections, which are copies of the original sound waves, but are of different amplitudes and originate in other places than the direct sound. The movement of the sound recipient is interchangeable with the movement of the sound radiator, so that even if our listener sits still, the movements of other listeners will disturb the paths of reflected waves, and the result will be a fluctuating sound field which provides a great number of cues for auditory location. At instrumental performances the movements of the musicians is a considerable aid in the sound localisation that is fully exploited by the auditory mechanisms, for instance to distinguish a soloist from the rest of the ensemble. Such an aid is not available when the music is distributed solely through loudspeakers, and with a small audience the amount of motion in the room is limited. In addition, excessive physical movement among the audience is a social behaviour that is not accepted in the concert tradition of Western art music. However, the motion needed to collect extra cues for localisation depends on the spectral content of the audio information, and even the smallest head movements may suffice. Nonetheless, during music listening we normally do not collate additional directional information in this fashion. Exploratory head movements, I have observed, are as good as non-existent among concert listeners, regardless of listening position.

5.3 Sound systems

5.3.1 Headphone listening

Headphones are the most private of listening spaces. The sound radiators are placed on the pinnae or just inside the ear canal. This minimises sonic disturbance from the environment and does not disturb the surroundings unless the listening level is excessively high. Using headphones as the sole sound radiator eliminates acoustic cues otherwise provided by reflection, diffraction and absorption of sound waves in the physical environment. The filter effect caused by the pinnae that varies with direction and loudness of the sound source is also absent, preventing another important aid in sound localisation and aural orientation. Furthermore, we do not only listen with our ears, but also with our body in terms of sensing low frequency vibration. This tactile sonic experience is not present in headphone listening. The effect of extrinsic space is therefore reduced to only the virtual extrinsic cues that are composed into the music.

A stereophonic image is based on differences in intensity, phase and time of arrival at the listener's ears, and requires therefore that the signal from each loudspeaker reaches both ears. This is particularly true with *intensity stereo* where the horizontal location of a sound in the stereo field is determined by the relative intensity between the two channels (Watkinson 1998: 8). With headphones there is no such acoustic "crosstalk" between the two sound channels as there is between loudspeakers in a room, and both ears thus do not receive the signal from both channels. Strictly speaking then, there is no stereophonic image on headphones (Watkinson 1998: 207).

When music intended for loudspeaker playback is listened to over headphones, the sound image appears stationary on an axis behind the eyes. There is no spatial depth as there is in a frontal stereo image. Interaural differences are fixed to what is present in the music, and there is no change in auditory information when the head turns as one expects under normal listening conditions. Spatial differences resulting from head movements are no longer present, because the sounds move as the head and eardrums move. Thus, there is no help from the environment and the pinnae in detecting the position of sound sources as we are used to from our everyday existence, and the only possible place the sound appears to come from is therefore inside the head. Blauert (1997: 116) terms this experience *inside-the-head locatedness*, which occurs when the distance to the auditory event is judged to be smaller than the radius of the head.

Spatialisation on headphones, especially in terms of sound placement outside the head, will be successful only if the music is mixed specifically for headphone listening[8], in which case the spatialisation will most likely suffer when the music is played back over loudspeakers. Reproduction of binaural recordings depend on complete channel separation, and headphones are therefore ideal for this kind of material. Even though binaural recordings are regarded as most accurate in terms of hi-fi sound and spatialisation, their use has not been widespread. The majority of music productions and radio broadcasts are intended for loudspeaker playback, and headphones are not a standard component in a home listening system. In concert situations headphones become expensive and impractical if more than just a handful of listeners is to be catered for simultaneously. It is therefore in more specialised situations that headphones have become the preferred sound system, such as in virtual reality systems and certain types of sound installations which are designed to be experienced in private.

[8] An exception is when employing a *shuffler*, a device which simulates the cross-coupling of loudspeaker listening by means of delay and filtering. The result is a frontal stereo image which is nearly as spatially realistic as conventional stereo loudspeaker listening (Watkinson 1998: 10-12, 207-08).

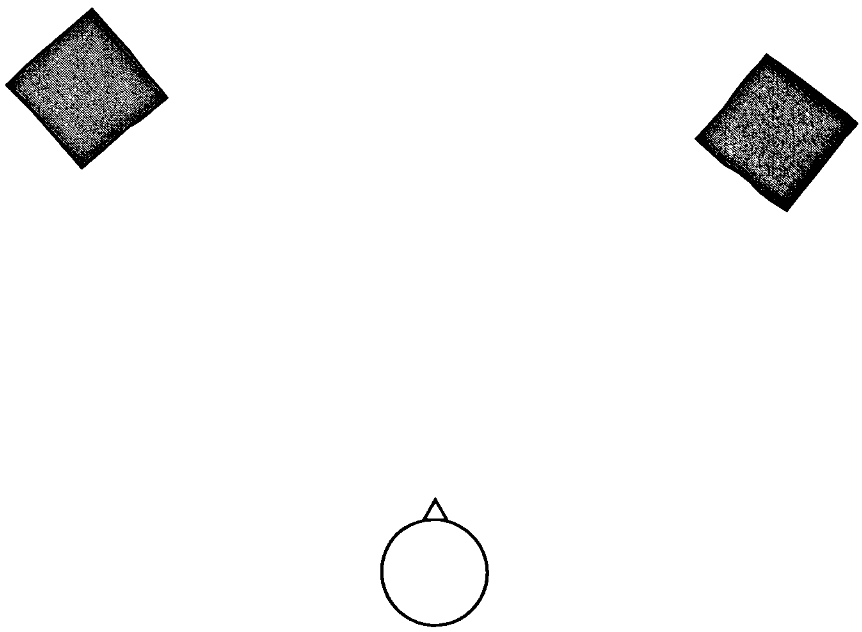


Figure 5.4: Stereophonic loudspeaker array.

5.3.2 Stereophony

The first documented demonstration of stereophonic sound took place when Clément Ader transmitted sound from the Paris Opera over a pair of telephone lines in 1881 (Malham and Myatt 1995: 58; Watkinson 1998: 196). However, it was not until the mid-1950s, when pre-recorded stereo reel-to-reel tapes became commercially available, that two-channel stereo became the dominant sound delivery format. Two-channel stereo is still the most common playback setup in private homes as well as the most widely used configuration in composition studios for electroacoustic music.

The stereo technique is based on the appearance of a *phantom image* between the loudspeakers in which the sounds are positioned. Horizontal placement and movements within this stereo field are achieved by altering the relative amplitude output from the two channels and/or manipulating phase and time differences between the two signals. The listener is positioned in front of, and facing, the sound field (see figure 5.4). The loudspeakers should ideally be placed at an angle of 55-65° at the listener's position (Toole 1998: 476). This angle influences the spatial resolution and the stability of the stereo image. Smaller angles narrow the stereo field and subsequently decrease the size of the sound field, resulting in a reduction of spatial resolution. Greater angles create increasingly unstable phantom images, often with the result that sounds are localised at the positions of the loudspeakers rather than between them, or that the sounds appear to jump arbitrarily from side to side. At 90° and beyond, stable central images in a standard stereo-setup cannot be formed (Malham and

Myatt 1995: 61).

Symmetrical placement of the loudspeakers relative to walls and listener position is crucial for the best possible reproduction of the stereo sound field. Equal distance from the two loudspeakers to the side walls is important in order to minimise inconsistencies in reflections and absorption which may result in spatial distortion of the stereo image. The quality of the loudspeakers is important, particularly in terms of linearity and resolution. Watkinson (2001: 721) denotes that "non-linearity in stereo has the effect of creating intermodulated sound objects which are in a different place in the image from the genuine sounds." This means that inaccuracies in the stereo system can lead to serious spatial artefacts in the reproduction of the stereo field.

At long wavelengths the two loudspeakers work in each other's near field and thus appear acoustically as a single radiator (Watkinson 1998: 103). Consequently, the low frequency output from the two loudspeakers cannot acoustically be out of phase, and on a properly installed system this interaction will enhance the bass. However, if the two loudspeakers are reversed in phase the bass frequencies will cancel each other out and the overall stereo image will be poorly defined.

The average home stereo system is not always designed with accurate sound reproduction in mind, and distortion both in terms of stereo imaging and frequency response is not rare. Phase problems are common, as well as a tendency to give priority to loudness over clarity in the lower end of the frequency range. Poor loudspeakers have their own characteristic "sound", and there is such a variety in these sonic characteristics among different loudspeaker models that an "average" low-quality loudspeaker does not really exist (Watkinson 2001: 722). Relying upon an arbitrary pair of low-quality loudspeakers in music production in order to get an indication as to how the mix may sound on an average home stereo system is therefore very risky.

Diffusing surfaces close to the listener may result in a smearing of the stereo image, so that sounds placed close to each other in the stereo field cannot be properly distinguished spatially (Toole 1998: 476). This may happen when the listener sits in a chair with a high back that extends above the shoulders, or in a listening position close to the back wall. Problems with this kind of near field reflections can only be rectified by treating the surfaces with absorbant material.

Because there is, strictly speaking, only one ideal listening position for stereo, a single pair of loudspeakers is not adequate for concert performance of electroacoustic music, especially if space has a significant structural function in the work. A stable central stereo image can only be experienced from a listening position in the exact centre relative to the two loudspeakers and at a specific distance so that the angle between listener and loudspeakers is within the optimal range. In a concert situation with an audience of some size, this is not possible to achieve for every listener. A listening position off-centre will lead to the listener hearing predominantly what is emitted from the nearer loudspeaker; a centred listening position, but one that is too close to the loudspeakers, can lead to the experience of a "hole in the middle" where all or most of the musical activity appears to take place on the sides; and a listening position too far from the loudspeakers can cause the stereo image to disappear in reverberation and phase cancellations, and be more or less experienced as a mono signal heard from a distance (Doherty and Berry 1998: 9; Harrison 1998: 14). A solution to this for concert performance of stereo compositions is to employ multi-loudspeaker arrays, something which has become standard practice in electroacoustic music concerts. Multi-loudspeaker sound diffusion will be discussed in section 5.5 below.

5.3.3 Surround-sound systems

5.1 Surround-sound systems

For home listening, the 5.1 surround-sound configuration has been established as the standard setup for surround-sound audio. This setup has its roots in cinema sound[9], and its adaptation to the home reflects the focus of mass market electronics manufacturers on complete home entertainment systems where audio and video are integrated. After the failure of quadraphony, it appears that the electronics manufacturers are wary about introducing a new multi-channel audio-only format for private listening[10]. Most casual listeners have never experienced surround-sound loudspeaker setups in concerts, and would consequently not see the point of adding more loudspeakers to their home music system. Surround sound has, however, been standard in cinemas for so long that anyone who has been to the cinema in recent decades has experienced it. Furthermore, home video delivery formats with surround sound have already been available for more than two decades, and the prospect of recreating the cinema experience at home has led many to install additional loudspeakers to their video setup. This has paved the way for the recent interest shown by the commercial music industry in producing and releasing surround-sound music to be played back over the same (or similar) sound system that was originally intended for home cinema. The fact that the DVD-Audio specification allows for video to be included on the disc shows the strong ties of the new surround-sound music format to home cinema.

[9]This configuration is sometimes referred to as *cinema style surround* (Malham 1998).

[10]It is interesting to note that the CD specification allows for four channels of uncompressed, full-bandwidth audio at half the playing time. However, to my knowledge, only one manufacturer ever made a CD player with four output channels, and a reintroduction of four-channel audio for home listening was therefore never bound to happen. Quadraphony was a failure partly because of technological limitations of attempting to deliver four channels of audio on two-track analogue vinyl discs (Holman 2000: 14). Digital formats, such as the CD, DVD or SACD, can store multiple audio channels without crosstalk or phase misalignment, and make true surround sound possible (Watkinson 1998: 14).

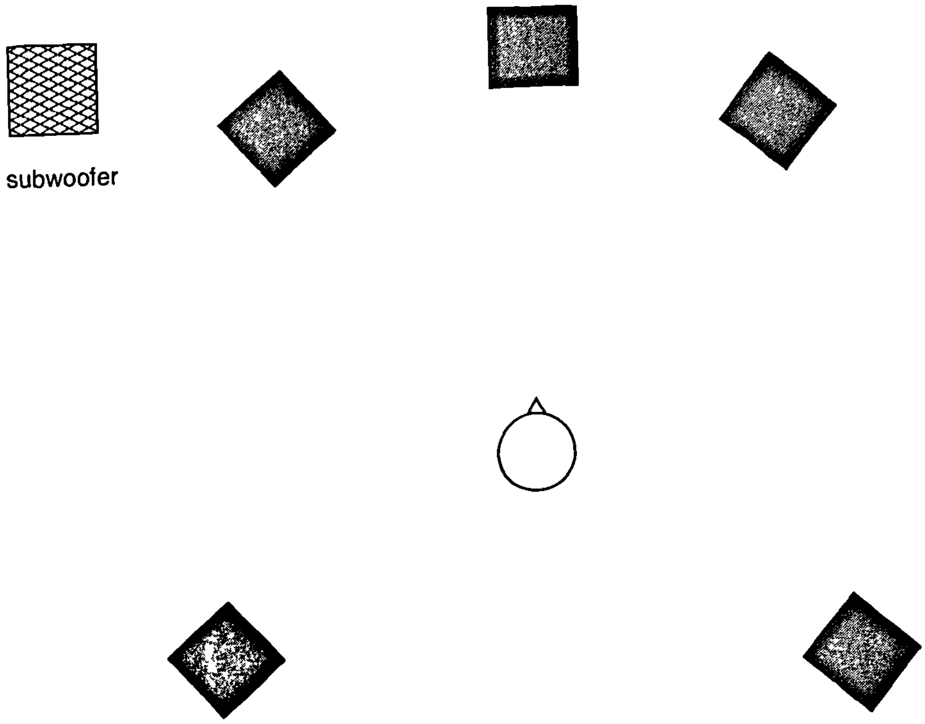


Figure 5.5: 5.1 surround-sound configuration.

The 5.1 surround configuration comprises three front loudspeakers (left, right, centre), two rear loudspeakers (left surround, right surround) and a subwoofer (the ".1" channel[11]) as shown in figure 5.5. This configuration was agreed upon by the *Society of Motion Picture and Television Engineers* (SMPTE) in 1987 as the minimum number of channels that will "create a sensation of surround sound" (Holman 2000: 16). It is based on the current practice in 70 mm film at the time[12], which employed five main channels and one channel for *low-frequency enhancement*[13] (LFE). The recommended[14] placement of the loudspeakers is different for video and audio. In cinema halls, the front loudspeakers are placed behind the screen[15], with the side loudspeakers at or near the edge of, but not outside, the screen.

[11] Actually, .005 is a more accurate number. This decimal number refers to the sample rate of the LFE channel, which is 1/200 of the principal sample rate. .1 was chosen to represent the system requirement more simply (Holman 2000: 66).

[12] Multi-channel audio in film dates back to Disney's *Fantasound*, which was a sound spatialisation system developed in 1938-41 for the film *Fantasia*. Three channels of recorded audio were played back over three front channels and two rear channels—a configuration close to the present day's 5.1 system (Malham and Myatt 1995: 59; Holman 2000: 12-13).

[13] Sometimes referred to as *low-frequency effects*.

[14] The *International Telecommunications Union* (ITU) have issued guidelines for loudspeaker placement. Although they do not consider themselves a standards-making body, their recommendations are highly regarded and become *de facto* standards in many countries, and are often followed by major international electronics manufacturers (Woodgate 1998: 579).

[15] A high-end cinema screen is perforated in a way which minimises high-frequency attenuation to an almost negligible level.

As long as the loudspeakers are reasonably directional, all frontal sound will then appear from inside the frame of the image. A television set, which is the most common "screen" for home cinema, is much narrower, and the loudspeakers must be placed outside the image frame: the centre loudspeaker above or below the television set and the left and right loudspeakers to the sides. The left and right loudspeakers for video sound are recommended to be at a maximum 45° angle at the listening position, while for music listening the recommendation is the same as for two-channel stereo, namely 60°. With the increasing interest in surround-sound music, I believe that the preferred loudspeaker placement for the home entertainment system is likely to become what is recommended for music, which favours a wider stereo image with a higher spatial resolution. This ensures some flexibility of the configuration, because most listeners will also use the same loudspeakers for conventional stereo playback.

Home video watching tends to be more of a social activity than music listening, and it is likely that some of the viewers/listeners are seated off-centre. The function of the centre channel is mainly to ensure that these listeners are able correctly to localise sounds which have a centred visual source, most notably dialogue. A listener seated in the exact centre would often not need (or notice) this channel because the sound would already be centred in the phantom image. However, the presence of a centre channel in the sound system creates a "hard" centre which makes possible a greater angle than 60° between the left and right loudspeakers without the frontal sound stage collapsing for the centred listener. The hole-in-the-middle effect is thus eliminated and the increased distance between the left and right loudspeakers increases the spatial resolution in the sound image and the general spaciousness and potential envelopment of the listener. For solitary listening a 90° angle (i.e. 45° between centre and left/right) would significantly increase the resolution of the frontal stereo image, and even 110°, which is within the ITU-recommended angle between the left and right surround loudspeakers, may be advantageous in terms of spaciousness. Such a wide angle between left and right reduces the distance between the front and the rear pairs of loudspeakers and allows for potentially smoother transitions between front and back. However, in order to achieve a perfectly smooth front-to-back transition by means of amplitude panning, it is crucial also to employ filtering in order to compensate for the differences in spectral sensitivity of the auditory system to sounds coming from the rear compared to those from the front. If this is not done, the sound will be experienced as initially being together in the front, then as some signal is fed to the rear loudspeakers, listeners may experience a "spectral split" in the sound where certain frequency bands are localised in the front and others in the back. The spectrum will merge again when the signal is only emitted from the back pair. The same type of split can be experienced with sounds that are panned in a fixed position between front and back.

A similar problem may arise in connection with the subwoofer when playing back bass sounds which contain higher overtones. If there is a bass management system present, something which is quite common in home surround playback systems, this extracts the bass frequencies from all the channels and sends them to the LFE only. The common cut-off frequency in such a bass management system is in the range of 80-120 Hz. The notion here is that low frequencies cannot be accurately localised, and a single dedicated bass loudspeaker therefore is all that is needed. In a medium-sized, slightly reverberant space this might be true. However, the subwoofer does not necessarily have to be placed on the centre line of the loudspeaker array (most often it is not), and in a relatively dead room the subwoofer is likely to be localised if it is not placed with utmost care. The spectral content of a sound below the cut-off frequency may then be localised at the position of the subwoofer, while the overtones are localised in one or more of the other loudspeakers. There are no guidelines as to where in relation to the five main loudspeakers the subwoofer should be placed. Its placement is a matter of experimentation where the ideal spot is a point in the room where it cannot be localised aurally from the main listening position. This is often achieved by placing it in a corner or near a wall so that the near-field reflections from wall(s) and floor help spread the sound and make the low frequency output of the sound system as non-directional as possible.

It is important to be aware of the spectral and spatial implications of the subwoofer at the listener's end when composing works for home surround listening. A bass management system is not particularly common in composition studios; the level and output to the subwoofer is more often controlled manually—if there is a subwoofer present at all[16]! The result of not having a subwoofer and/or a properly calibrated bass management system in the studio can be that the listener at home with a complete 5.1 system hears things in the mix that the composer did not even know were there.

Note that the 5.1 surround configuration does not deliver *true* surround sound, because all the five main loudspeakers are placed in the same horizontal plane. There is no overhead sound output, and in terms of space the 5.1 system is only two-dimensional. Furthermore, it is a discrete-channel system where mixing is done pair-wise and is therefore of the same principle as intensity stereo. This makes the 5.1 sound field just as fragile as two-channel stereo in terms of stability of the phantom images. A centred listening position in equal distance to all the five main loudspeakers is crucial for the success of the spatial image of the 5.1 surround.

[16]A full-range loudspeaker should be able to play back the actual full range of human hearing. However, such loudspeakers are often impractical to install in a studio environment due to their size (and cost). Most studio monitors therefore only go down to around 40 Hz, and without a subwoofer to manage the lower end of the frequency spectrum, a full octave of audio within the hearing range becomes inaudible to the composer.

Ambisonics

While 5.1 surround primarily is a home listening format with a specific layout, ambisonics is a more flexible, scalable approach to surround sound which can be employed both in a home environment and in concert situations. In contrast to 5.1, ambisonics does not specify a particular loudspeaker array, neither in terms of quantity nor placement, other than a symmetrical layout where the loudspeakers are evenly spread around a centre point. When preparing an ambisonics mix, the composer specifies direction and distance[17] of the sound relative to a reference listening position within the sound field rather than output channel of the playback system as is done in discrete-channel mixing techniques. The sounds' directional components are encoded in the mixing process as a set of spherical harmonics (Malham 1998: 175; Rumsey 2001: 116-17). The minimum set necessary for full three-dimensional sound consists of four signals, often referred to as first-order *B-format*, that are stored as four[18] regular audio channels. These four signals may contain a single sound in a single position or many sounds spread out all over the listening sphere. In any case, the B-format signal can be treated as a single entity and undergo transformations in order to, for example, zoom into areas in the sound field or rotate, tilt, mirror or filter the entire sound field without necessarily losing spatial resolution or stability (Malham 1998: 175). B-format signals can be synthesised or recorded, such as with the special SoundField[19] microphone. Several B-format signals can be mixed, and even non-ambisonics signals can be mixed in with the decoded signal without compromising the spatial quality of the sound field.

The ambisonics signal is decoded at the playback stage for the available loudspeaker configuration. Thus, the number and placement of loudspeakers is largely a matter for the reproduction end, and does not have to be specifically addressed by the composer during the creation of the work. The basic criterion with regards to loudspeaker configuration is that a minimum of four loudspeakers placed in a rectangle is required for two-dimensional horizontal surround (*pantophonic systems*), while eight loudspeakers in a cuboid is the minimum for full three-dimensional sound (*periphonic systems*). As a general rule, the more loudspeakers in the system, the more accurately and stable will the sound field be represented (Malham 1998: 176). It has been observed, however, that ambisonics decoding can cause anti-phase components from diametrically opposed loudspeaker pairs (Rumsey

[17]Distance information is not a part of the B-format coding *per se*, but is recorded or mixed in just as any other sonic information. Some ambisonics production tools implement distancing algorithms and let the composer specify distance together with direction for each sound or event.

[18]Second-order requires nine channels, while third-order calls for fifteen. A higher-order ambisonics signal improves the accuracy of the spatial image over a larger listening area (Rumsey 2001: 116-17).

[19]The SoundField range of microphones from Soundfield Research in Yorkshire is based on a design of four separate microphone capsules in a tetrahedral array which makes it capable of recording the full three-dimensional sound-sphere from a near-point source. The recording signal is output, via an electronic control unit, as B-format (or as a choice of more conventional formats, such as figure-of-eight or M/S stereo).

2001: 116). This can be advantageous for localisation at the point of the reference listening position in a small room, but less so in a large room and in situations where (most) listeners are positioned away from the centre. This phenomenon can similarly cause problems where loudspeakers are positioned in less than ideal locations. The effect of an asymmetrical or uneven loudspeaker layout can, however, often be rectified by the use of delays and level adjustments in order to make the loudspeakers aurally appear to be positioned in the right places (Malham 1998: 176).

The decoding is done by applying psychoacoustically optimised shelf filters above 700 Hz that compensate for the shadowing effect of the head, and an amplitude matrix that determines the correct level for each loudspeaker in the rig (Rumsey 2001: 112). The decoder needs to know where the loudspeakers are and where the listening position is in order to compute the best possible output signal for the particular space. Each individual loudspeaker is sent a specific combination of the B-format signals according to its position relative to the centre of the array (Malham and Myatt 1995: 64).

Unlike techniques based on pair-wise panning, all the loudspeakers in an ambisonics rig collaborate during playback. In other words, they are all "on" at all times in order to recreate the wavefront of the audio as realistically as possible around the listener(s). This means that an ambisonics system is quite efficient in terms of sound output. Small loudspeakers can do the job very well and collectively deliver sound levels comparable to much larger loudspeakers in a stereo setup, for example. Similarly, the joint effort of many small loudspeakers results in an enhanced bass output which is on par with bigger loudspeakers found in a stereo setup.

Although a properly set up ambisonics system can deliver a stable spatial image over a relatively large listening area, the optimal listening position with the most accurate spatialisation is still confined to a small spot in the centre of the loudspeaker array. However, there is no sharp division in spatial quality between the centre position and off-centre as with a phantom-image based technique. The further away from the centre position one is situated, the less accurate the spatial image becomes, but the stability of the spatial image persists. The gradual decrease in spatial accuracy out from the centre spot thus creates an illusion of a much larger "sweet spot" than in stereo or 5.1 surround.

The flexibility of ambisonics with regard to loudspeaker configuration has led to the development of specific decoding algorithms for standard configurations such as two-channel stereo and 5.1 surround systems. Two-channel *UHJ* is compatible with stereo and is intended for playback on a standard stereo pair of loudspeakers. Naturally, there is no height information, and the horizontal surround image is less accurate than it is with four or more loudspeakers. However, the stability of the sound field is better than that of intensity stereo, and mono compatibility, which is still important for broadcast audio, is very good for UHJ-decoded recordings. *G-format* is a reversible decoding format specifically developed for playback on 5.1 surround-sound systems. A conventional 5.1 recording is simply moved from the studio environment to the home listening system without going through any form of decoding[20]. This can be done because the 5.1 loudspeaker configuration is standardised with the loudspeakers placed more or less in the same positions in the studio and at home. Merely to carry the decoded ambisonics signal from the studio to the home, however, is not satisfactory for the serious listener, because the height component, which is special to ambisonics, and the flexibility in loudspeaker placement become lost. The reversibility of the G-format makes it possible to use a decoder to recover the original B-format signals and drive a different loudspeaker rig, which may include the height component. Thus, G-format makes available superior surround quality for 5.1 listeners without decoders, while at the same time maintains the flexibility of ambisonics for those with decoders and dedicated loudspeaker arrays.

5.4 Case study: *Terra Incognita*

In the early planning of my composition *Terra Incognita* I considered several possible surround-sound formats for which to compose the piece. My compositional experience so far had been with intensity stereo and pair-wise panning, and I feel I master the spatial possibilities of the stereo field quite well. Furthermore, many of the sound sources I wanted to use for this composition were environmental sounds recorded in stereo. The concept of the work, however, suggested a three-dimensional space where sounds can be independently moved around in the listening space. The virtual space of two-dimensional stereo became too limited. Because this piece would not be composed for a particular hall or sound system, I was not limited by any specific technological requirements. My choice of output format could therefore be made purely on artistic and practical grounds.

The technical complexity involved in adding audio channels beyond the two of conventional stereo has implications for the portability of the work to different listening spaces. The choice of playback format therefore directly affects the number of possible performances the work might receive. Discrete-channel compositions are restricted to specific loudspeaker configurations, and can only be performed successfully in listening spaces which can accommodate these requirements. Although 5.1 surround is used on a larger scale in cinema halls[21], it is spatially too restrictive and frail to be a suitable format for

[20]There are exceptions, of course. Dolby Pro Logic and Dolby AC-3 are but two examples of compression technologies which require players with appropriate decoders built in. These compression formats are, however, more common for video sound than for audio-only discs because they cause a degradation in sound quality.

[21]Loudspeaker channels are duplicated in order to cover the entire listening area in cinema halls. Delay lines are used extensively in large halls to ensure that localisation of the audio corresponds with the visuals

concert performance of electroacoustic music, and is best left as a small-room listening environment. Eight loudspeaker channels arranged in a circle around the audience is a relatively common format for multi-channel electroacoustic composition, which requires eight matched loudspeakers set up at equal distance from a centre point. This, in turn, is not a realistic format for home listening. What I wanted for this composition was a multi-channel format suitable for both home listening and concert performance, and which would easily let me convert the composition to two-channel stereo for maximum portability. Ambisonics proved to be the only technique which would provide this flexibility.

Technology

While software for audio processing and mixing in stereo is easily available, there are far fewer choices of computer based tools specifically for surround-sound composition. Most serious mixing software does to some extent accommodate multi-channel work, and some have specialised modules for 5.1 surround, but two-channel stereo remains the main target, and the user interface for these programmes tends to reflect this. Processing and mixing software for ambisonics is particularly hard to come by. However, because all my sound sources for this composition were in stereo, I could still use my stereo editing and processing software when preparing the sound material. It was on the mixing side that problems arose. The only mixing software I found that could provide a sufficient level of control and flexibility for this work was Richard W. E. Furse's *VSpace*[22], which is a scripting language for spatialisation and mixing of mono sound files into an ambisonics sound field. Being a scripting language, *VSpace* does not provide any graphical representation nor real-time rendering of the mix. The entire mixing process takes place in a text based environment, and adjustments and additions to the mix can only be heard after two stages of computer processing: rendering of the mix into a B-format file by executing the script, and subsequently decoding this file so it can be played back using a multi-channel sound file player.

Because I am used to instant feedback in the mixing process, not being able to work in real-time meant a significant change in work methods. As the mix script grew longer, the rendering and decoding processes took an increasingly long time. It became important therefore to plan each addition to the mix well and specify timing and spatialisation for several sounds before executing the script. Experimenting with different sound material and spatial configurations in a more or less spontaneous way was impractical due to the lack of

for listeners in different distances to the screen, and in order to prevent undesirable echo effects when duplicating channels. (In fact, cinema films are printed so that the sound is emitted one frame (42 ms) early, with the result that audio and visuals are in sync approximately 14.3 m from the loudspeakers at the screen (Holman 2000: 62).)

[22] *VSpace* runs on the Linux, SGI and Windows operating systems and can be acquired from `http://www.muse.demon.co.uk/mn_download.html`.

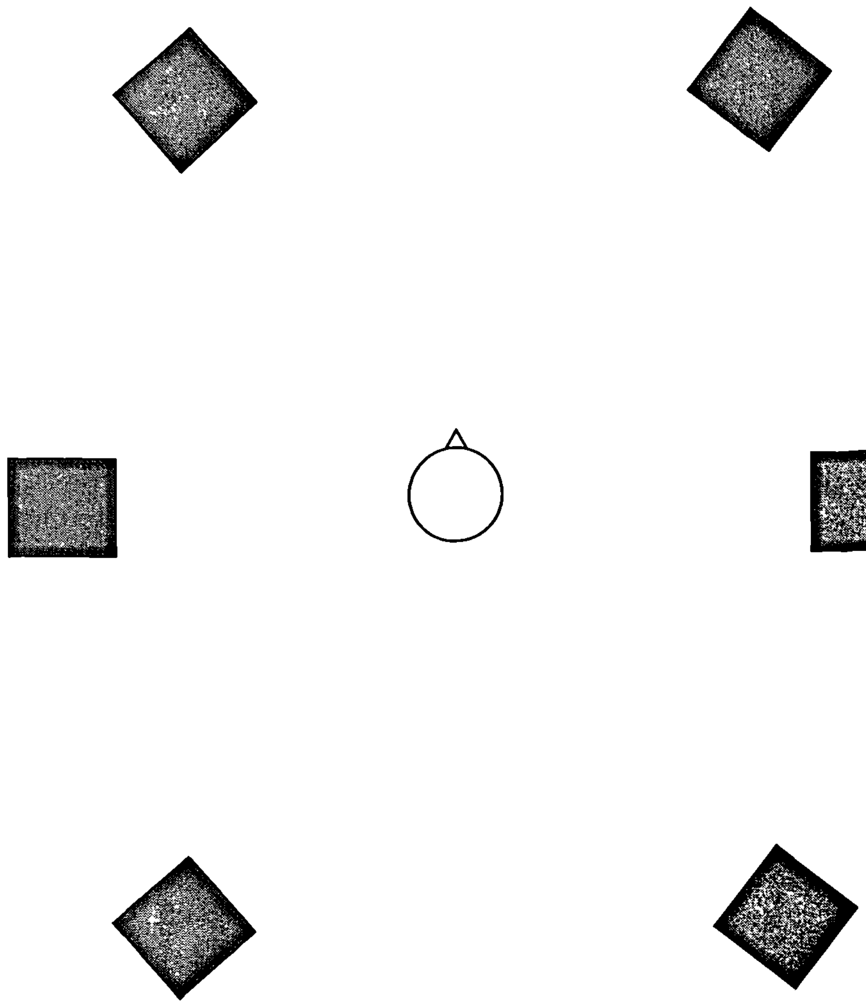


Figure 5.6: Monitor layout for the composition of *Terra Incognita.*

real-time response from the system. Furthermore, not having the visual aid of a graphical interface made the process of arranging and adjusting sounds temporally in composed space awkward and time consuming. However, the amount of spatial control available in terms of placement and movement of the sound material and reverberation and dimensions of the virtual sound field ensured a great amount of compositional flexibility.

A six-channel loudspeaker array was used for monitoring during the composition of the work (see figure 5.6).

Compositional considerations

The title, *Terra Incognita*, or "unfamiliar land", is based on the idea of using the three-dimensional sound field to create surreal referential spaces that change sometimes gradually,

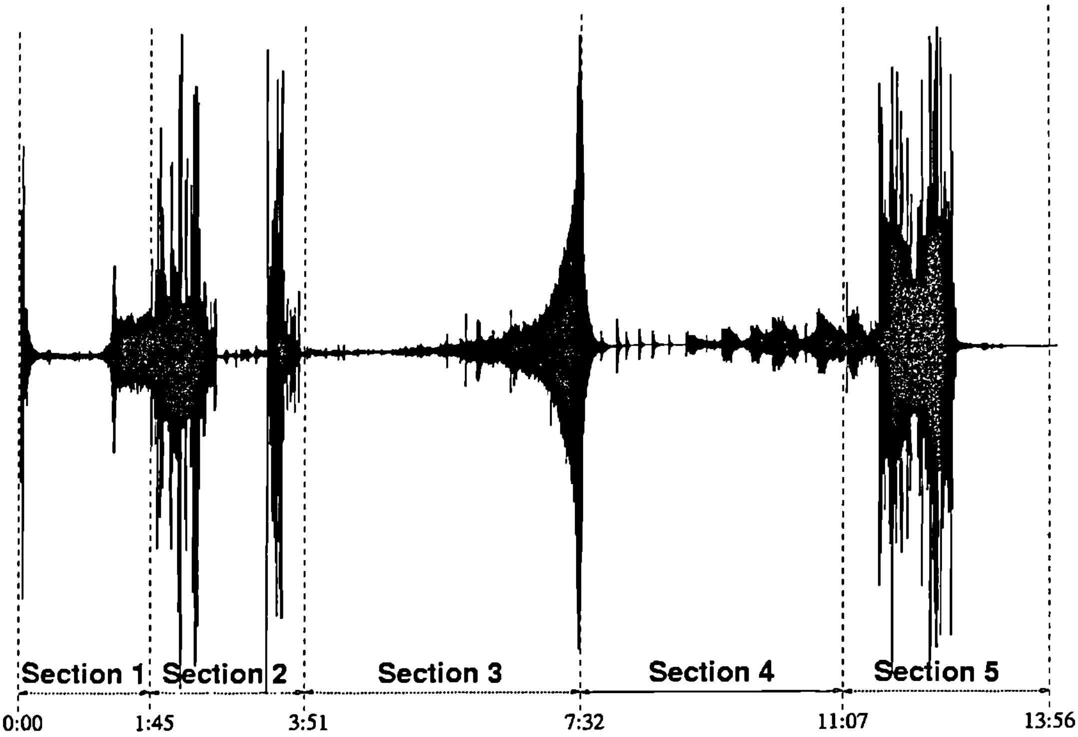


Figure 5.7: Time and amplitude view of *Terra Incognita.*

other times suddenly. The work is inspired by science fiction films, and portrays a journey motivated by curiosity, fear and excitement to an unfamiliar world where the "explorers" encounter a variety of situations and objects which they observe and/or with which they interact. In this piece, as in many of my compositions, I balance on the edge of narrative where I attempt to draw the listener into imagining scenes, only to disrupt the mode of listening and take the listener somewhere else, as if in a dream. I find the possibility of the surround-sound format to create such scenes greater than with two-channel stereo, because the three-dimensional sound sphere of surround-sound composition is closer spatially to our everyday listening environment, and familiar spatial cues can be utilised with greater accuracy and realism. The potential of surround sound to envelope the audience and create a sense of participation in the sound environment was an important motivation for choosing this format for the composition. Being able to move sounds independently around within the sound field makes possible spatialisation which has clear references to real-world scenarios. This was a central part of the compositional idea.

Structurally, *Terra Incognita* can be divided into five sections: the first section spans the first 1'45" of the piece, section 2 goes from 1'45" until 3'51", section 3 from 3'51" until 7'32", section 4 from 7'32" until 11'07" and section 5 from 11'07" until the end at 13'56" (see fig. 5.7). Sections 1 and 3 are of intrinsic spatio-structural orientation while the other three are predominantly of extrinsic orientation. Although the title and the underlying idea of the work might suggest extensive use of environmental space, I chose to explore extrinsic space as a way to suggest an environment. I did this not only because the three-dimensional space of surround sound provides room to play with movement and placement, but mostly because I wanted to abstract the objects of a referential space from the environment in which they might exist, and thereby leave more room for the imagination as to what the sonic environment around these sound objects might be[23]. Thus, I do not expect, and do not want, the inspiration for the work to show through, but rather leave the work more open for the listeners to experience and interpret on their own grounds. There is, of course, a structural progression in the work which may create certain expectations, but I do not necessarily see this as leading different listeners in any common direction in terms of interpretation.

This idea of "reducing" my control and influence was also important on a more detailed level during the composition of the work. Partly due to the meticulous process of defining placement and movement of the sound material in the mixing software I used for the piece, I wrote a small computer programme which would do some of this work for me. In this programme I specified one or several sound files and defined certain parameters for the movements of the sound(s), such as coordinates of a reference point, maximum distance from the reference point, speed of movement and number of movements during the course of the sound or during the course of the event when several sound files were specified. The computer then randomly spatialised the sound material within the given parameters. The output of the programme was a text file formatted as a *VSpace* script which could later be edited, if necessary, or copied directly into the main *VSpace* mix script for the composition. There are several events in *Terra Incognita* where the speed and extent of the movements are important, but where accurate positioning of the sounds are less relevant in the structural context. This can be heard, for example, early in the work, between 1'45" and 2'38", and towards the end, between 11'33" and 12'32".

The sectional division of the composition is mainly based on variations in sound material and spatial behaviour. There is some overlap in both sound and space between sections, which makes the division less straightforward than the case is with *Intra*, for example (see section 4.3). The first section of *Terra Incognita* has an introductory function; it represents the "arrival" in the unfamiliar land. Its function is to establish a foundation and indicate boundaries of the sonic landscape. The boundaries are suggested by a temporally stable sound, circling in a fixed radius from the centre, on top of a more diffuse sound mass lying underneath. The stability of this section in terms of spatiomorphology and spectromorphology points towards an intrinsic spatio-structural orientation. There is an intensification in

[23] I have in mind here the inclination human beings have to try to identify sources of sounds and sound events and relate them to past experience.

the music starting with the peak of a crescendo at 1'15". Here, the overall intrinsic space expands with a low bass sound of great magnitude while simultaneously there is an upwards expansion of spectral space in the diffuse sound mass as some of its frequency components gradually ascend. This leads to a switch in the structural function from stasis to latency in anticipation of a bigger structural change.

That change comes with the beginning of section 2 at 1'45". The spatiomorphology here is of a radically different nature than in the previous section, although a change in extrinsic space was anticipated by the increased speed of the circling sound at 1'15" in section 1. A variation of this circling sound is carried over into section 2, but here it moves faster and in the opposite direction. This section is, however, dominated by another, quite aggressive sound which is randomly and swiftly propelling within the sound field at varying distances from the listener. Although the section quietens down significantly between 2'38" and 3'18", the random, unstable movements in extrinsic space persist and characterise the entire section which thence clearly is of extrinsic spatio-structural orientation.

Section 3, starting at 3'51", is more diverse in terms of spatio-structural orientation, and serves as an example of how orientations can vary and combine within the same structural unit. Again, a variation of the circling sound from the previous two sections appears, but here its movements are less uniform: they alter arbitrarily between circles, arcs and lines. This sound is accompanied by a series of pitched tones which move at random within a limited, but relatively wide, spectral space while maintaining a stable position in extrinsic space. At this point the section is of spectral spatio-structural orientation because the variation in spectral space in the pitched tones are likely to attract the most attention. The pitched tones are eventually swallowed by a crescendoing sound mass which fades in from the distance in front from around 4'40". The crescendo takes place both in terms of gradual movement from the distance towards the centre of the sound field, and in terms of an increase in sound intensity as well as in density. At 4'55" is the first of a sequence of pitched front-to-back moving sounds, which portray some form of flying objects in the sonic landscape. The speed and extent of the movements are underlined by a Doppler shift as the objects move past the centre listening position. Immediately after the peak of the crescendo there is a relatively quick decrescendo as only a few of the components of the sound mass linger in the distance in the back of the sound field while the others quickly fade out. The flying objects contribute to a shift in spatio-structural orientation from spectral to extrinsic, but as the crescendo increasingly dominates the sound field, the overall orientation shifts towards an intrinsic spatio-structural orientation.

A new sound, with a "bouncing" quality, is introduced at 6'38", and although it participates in the intensification towards the peak of the crescendo, its main function is to provide a bridge between section 3 and section 4, which starts at 7'32". This bouncing sound is the central element in most of section 4. It has a sharp, "wooden" quality and is highly directional, and thereby indicates an extrinsic spatio-structural orientation. The pronounced reverberation here clearly points to an indoor environment, something that was hinted at in the beginning of the previous section where a certain reverberant quality could also be heard. The wooden sound bounces quite freely around in the space until 10'15" when another, more abstract sound, reminiscent of the dominant sound of section 2, appears. This sound moves around over a greater area than the wooden sound, and by flying close to, and sometimes through the reference listening position at a high speed may appear somewhat threatening. The sound continues into section 5, which starts at 11'07". Here, a more noise based background indicates a return to an outdoor environment. The most violent part of the piece starts at 11'33" when a highly fluid, storm-like sound mass suddenly enters. This is the most intense part of the whole work, where sounds are randomly thrown around in the sound field in close proximity to the reference listening position. The "storm" sound ends with a massive explosion at 12'32", after which the sonic landscape quietens down significantly and fades away, suggesting a retreat and a "departure" from the sonic landscape which subsequently disappears in the distance. The piece ends with a fade-out at 13'56".

Conclusion

Terra Incognita was an experiment in several ways. Being my first surround-sound composition, a main objective for the project was to investigate the compositional possibilities and implications of working in three spatial dimensions. Part of this was to explore spatial extremes in the three-dimensional sound field. Thus, the piece displays great contrasts in sound density as well as in sound intensity among different parts, and similarly with regard to spatial density and spatial intensity of the sound material. It was particularly extrinsic space as a structural means I wanted to look into with this piece, mostly because it is with regard to placement and movement of sound material the difference between the two-channel stereo space and the multi-dimensional space of surround sound is the most pronounced.

I was also interested in the technical issues regarding surround-sound systems as well as the differences in compositional environment between conventional stereo and surround sound. Working with the real space of surround sound is quite different from the virtual space of two-channel stereo. The main difference, of course, is the multi-dimensionality of the composed space. Unless one is working with source material that is recorded in three dimensions, in which case the desired spatialisation may already be captured in the recording, there is the extra level of compositional decision-making which involves placement and movement of the sound material in this multi-dimensional space. The composition process becomes more complex, not only because there are more possibilities with regard to placement and movement of sounds, but also because enlarging the space affects the density of the sound field—the sound field "opens up" as the sounds are given more room.

With the amount of spatial control available in the ambisonics sound field, it is possible to use extrinsic space as a central structural device in composition, where sophisticated spatial configurations and trajectories can be created and reproduced with great accuracy and realism.

5.5 Sound diffusion

Having considered different basic formats of sound systems as well as general acoustic features of physical spaces, a background is established in order to discuss *sound diffusion*, the standard performance practice for electroacoustic music in concert.

5.5.1 The sound diffusion system

Multi-channel loudspeaker systems have been a feature of electroacoustic music concerts since the early days of the genre. Most electroacoustic works are composed in two-channel stereo and intended to be "spread out" over multiple loudspeaker channels in concert. On a multi-loudspeaker diffusion system the performer, the *sound diffusionist*, has the possibility to articulate and enhance the composition's inherent spatial and dynamic properties and underline the spatio-structural content of the work by physically distributing the sound around in the listening space. By nature, the diffusion of an electroacoustic composition is specific to a particular sound system and concert venue. Sound diffusion is a site-specific performance that has to be prepared for, and rehearsed in, the particular listening space of the concert so that the spatio-structural features of the work can be put across to the listeners in the best possible way in that particular space. A sound system set up properly to match the specific music, room acoustics and audience seating has the potential of making the loudspeakers appear transparent to an extent where they are hardly noticed as the actual sound radiators, and make possible a multi-dimensional sonic experience which is quite unique to the genre of electroacoustic music.

Sound diffusion is an art which requires considerable practice and experience to master. It is my opinion that the combination of concert hall and diffusion system can be considered a musical instrument on which the electroacoustic work is *played*. The aesthetic choices and possibilities of the diffusion itself are therefore inextricably connected to the configuration of the sound system and the architectural layout and acoustics of the hall. Because space and spatial features often are crucial structural elements in electroacoustic works, and—as previously stated—are fragile elements in combination with physical spaces, it is obvious that great care must be taken in the public presentation of electroacoustic music compositions in order to communicate coherently the composed spatio-structural elements of the work in the larger concert space.

The loudspeaker system needs to be "tuned to the room" in order for the room and the loudspeakers to match and appear as one seamless acoustic system. This can be challenging, especially with rooms where there is a considerable difference in acoustic quality with and without the presence of an audience. The placement of the loudspeakers in a concert hall depends on the type and number of loudspeakers available, the shape and size of the hall, the size and placement of the audience and whether the system is primarily for performing stereo or multi-channel compositions. It also depends on the composed space and spatio-structural orientation of the musical works. A diffusion system set up according to the requirements of one type of composition does not necessarily sound as good with another. Thus, already in the choice and placement of loudspeakers lie aesthetic decisions connected to the spatio-structural properties of the specific compositions one intends to perform.

There are two basic approaches to diffusion system configuration. In the purest form these are:

1. *Surround setup.* The loudspeakers are normally arranged in pairs symmetrically around a front to back axis, and set up to surround the audience. The configuration can be in the shape of a circle, a square, a cube, or variations of these.

2. *Loudspeaker orchestra.* Loudspeakers of different sizes and types are placed on and around a stage, symmetrically in pairs and/or asymmetrically in groups or individually.

Most often one finds combinations of these two approaches. A very common configuration is to set up the loudspeakers to surround the audience, but to have more (and the highest quality) loudspeakers in front in order to achieve more locational subtleties in the facing direction. This makes sense perceptually because the spatial (and spectral) resolution of the auditory system is the most accurate in the front, something that can be utilised musically.

Most diffusion systems primarily intended for diffusion of stereo works are made up of loudspeakers of different types and sizes and with different frequency responses in order to spread the sound spectrally as well as spatially around the listening space. Filters are in some cases used to split the spectral response of the system in a more controllable fashion. The spectral space of the work can in this way be divided into different ranges which can be placed in different parts of the room. It is not unusual in a diffusion system to have, for example, multiple tweeters placed above the audience which can emit only the highest frequencies and one or two subwoofers in suitable positions on the floor which only play back frequencies below a certain cut-off frequency. The auditory system tends to locate higher frequencies more easily and rapidly than lower frequencies, and different spectral parts of the sound can therefore be experienced as moving at different speeds and directions depending on the frequency response and directivity of the loudspeakers. Many compositions are layered spectrally in order to diffuse well on this type of system.

Multi-channel works normally require a configuration of a number of the same type of loudspeakers. A common configuration used for eight-channel compositions, for example, is a circular setup of eight full-range loudspeakers. To accommodate such works, a surround configuration of the same type of loudspeaker could form the basis of the diffusion system and more spectrally "coloured" loudspeakers could be added in appropriate locations for traditional stereo diffusion. Matched loudspeakers are also preferable for ambisonics playback, but different types of loudspeakers may work well here as long as they are in strict phase alignment with each other.

5.5.2 Diffusion strategies

What constitutes a successful diffusion is a musical performance in which the combination of composed space and listening space is such that the integrity and meaning of the work are retained and coherently communicated in the specific listening space. This is done by articulating and enhancing the spaces that are composed into the work on the basis of implied spatiality in the sound materials and their relative configuration in composed space. Knowing and understanding the musical content of the work is therefore a key issue. A successful performance requires that the sound diffusionist is intimately familiar with the work in question, and has analysed it in terms of its structurally significant spatial components. One diffusion may be driven by the semantic or referential nature of the sound material, if that exists in the work, while another may be based on locational relationships among the sound material. These are interpretative choices that need to be made on the basis of a spatio-structural analysis of the work. However, not all works are equally suitable for sound diffusion. A good starting point is necessary, where spatio-structural elements are composed into the work in a way which makes further articulation and enhancement in a larger listening space possible. Obviously, knowing the instrument is paramount to any good performance, and sufficient rehearsal time is required on the particular system on which the music is to be played. This is a factor that, sadly, is too often compromised.

What does it mean, then, to articulate and enhance the spaces that are composed into the work? First of all, the spatial essence of the work must be identified to make the basis of the diffusion strategy for the work. A spatio-structural analysis will reveal the spatio-structural orientation as well as more detailed structural functions in the work. For example, a work of extrinsic orientation may ask for a diffusion where the emphasis is on rapid movement of the sound material in easily identifiable locations in the listening space, while a work of intrinsic orientation may require a diffusion where spectral diffusion, envelopment and slow movements between regions in the listening space are necessary. These are high-level decisions which can form the starting point for a performance plan. An examination of lower-level structural functions will suggest diffusion strategies concerning individual spatial features that are central to the work as well as other spatio-structural elements on a detailed level that need to be brought forward in performance. Broad sweeps that mark beginnings or endings of structural units and circular movements that symbolise enclosure are examples of spatio-structural functions that may need to be underlined in diffusion.

I have identified eight categories of spatio-structural elements that are likely to be affected by the translation from composition studio to concert hall, and which can be controlled in sound diffusion. All the categories are not relevant for all types of compositions, but they serve to highlight the dynamic nature of musical space in electroacoustic music. The discussion shows that although the music is composed on a fixed storage medium, a spatial flexibility still exists which ensures that the music itself is far from fixed.

Spatial dimensions

The spatial dimensions of the work are drastically affected by the listening space, but can be effectively controlled in sound diffusion. In the composition studio, the distance from the listening position to the loudspeakers is fixed so the width of the stereo field remains steady for the composer during the production of the work. However, in the concert hall this distance is likely to vary significantly among listening positions. The stereo field becomes narrower the further away from the loudspeakers listeners are seated, with the result that the horizontal width cannot be the same for all listeners. It is fairly standard in diffusion systems, therefore, to have several pairs of loudspeakers in front, and sometimes at the rear, with different widths which can be used dynamically to narrow or widen the horizontal dimension of the work according to how the stereo field is used in the composed space. For example, if the composed space comprises a spatially concentrated sound field, a narrow pair may be used in diffusion to make the sound field appear narrow from most listening positions. On the other hand, if the full spatial resolution of the stereo field is used, a wide pair may be employed to ensure that as many listeners as possible are within the boundaries of the lateral-horizontal sound field (a narrow pair may also be necessary in order to avoid "holes" in the sound field for listeners seated towards the front).

How near the listener sounds can be placed is normally limited to the physical placement of the loudspeakers. Small loudspeakers, such as tweeters, can be arranged in grids or otherwise distributed above the listening area in equal distance to all or most listening positions. This makes possible a physically closer and more uniformly placed sound location than can be achieved with loudspeakers in the horizontal plane. Suggestions of depth and distance in the work can be emphasised by using loudspeakers that are placed at some distance and/or are pointing away from the audience so their output is mostly heard as indirect sound. Natural or electronic attenuation of high frequencies can also be used to enhance or add distance cues to the sound field. Illusions of great distance can be created this way.

The height dimension, which is not physically present in a two-channel stereo field, can be added in diffusion if the listening space permits. Height can thereby be imposed on any sound material where this is appropriate, and elevation in spectral space can be underlined by actively using raised loudspeakers. Most rooms with high ceilings have a large reverberant area above the audience which can be utilised to create a non-directional sound field which effectively can contribute to the envelopment of the audience should the composition call for it. This is effectively done by placing the loudspeakers so that the sound reflects off the ceiling or walls before it reaches the listeners' ears.

Loudness

Expanding the dynamic range of the composition by making loud parts louder and soft parts quieter is an essential technique in sound diffusion. Doing the opposite in order to even out dynamic differences is against the composer's intent and will most likely damage structural cues in the piece. Due to quantisation noise in digital systems and hiss in analogue circuits, the maximum dynamic range of the playback medium is in practice relatively narrow; a realistic estimate is approximately 60 dB. An electroacoustic work prepared primarily for home listening may have an even narrower dynamic range, perhaps as little as 30 dB or less[24]. In comparison, a symphony orchestra can achieve a 120 dB dynamic range, and the range of the human auditory system is approximately 125 dB between the threshold of hearing and the threshold of pain. Extending the dynamic range beyond the limits of the audio storage medium is therefore expected by the ear in a concert situation, where the listener ideally should not be aware of any technological limitations in the performance.

Spatial placement

The translation of sound placement from virtual composed space to placement in the real space of the concert hall is a crucial factor in retaining the spatio-structural coherence of the composition. Greater distances, full horizontal surround and possibly the added dimension of height in the concert hall open up opportunities for more variation in spatial placement of the sound material than is possible in the composition studio. Care must be taken when interpreting spatial interrelations within the composed space and when deciding how these best can be put across in diffusion. For example, which spatial functions depend on close

[24]Pop music is generally compressed in order to achieve as high overall level as possible, often reducing the dynamic range to less than 10 dB.

proximity among the sound materials? Are there spatial functions where a sense of opposition or dialogue needs to be brought out by maximally extending the distance between the sounds? What sound materials may benefit from the social association a location behind the audience may evoke? Do any of the sounds suggest elevation? Are there passages where a full envelopment of the audience is necessary? Are sounds directional and require that only one pair or even just a single loudspeaker is used at a time, or are they non-directional and need to be sent to groups of loudspeakers covering larger areas of the listening space?

The answers to such questions lie in the sound materials themselves, and will only be revealed by studying the work for its spatio-structural properties.

Spatial movement

It is by moving the sound around the listening space that much of the "drama" of sound diffusion is created. Spatial movement of sounds can be brought out in diffusion by exaggerating the extent of the movement. Lateral sweeps, for example, can be widened across the frontal sound stage in diffusion by using one or several pairs of loudspeakers placed at different widths, or extended into semi-circular trajectories by being made to begin or end to the side or in the back. Suggestions of rotations in virtual space can be made real by utilising the full surround possibilities of the diffusion setup. Passages of high spatial activity can often benefit from an active diffusion which plays with the orientation and extent of the spatial dimensions. It is a key point, however, that such movements are implicit in the sound material and the composed space. Adding movement to stationary sounds may contradict spatio-structural functions of the work and be against the composer's intent. It is an important skill of the sound diffusionist, therefore, to know when *not* to move the faders of the mixing console.

Environmental space

Works of environmental orientation may include referential spaces which can be communicated to the audience in different ways depending on structural function. The full surround possibilities of the diffusion system can be used in order to envelop the listeners with sound, and create an experience of being transported to different environments. Placing the environmental material only in the front of the sound field, for example, can be interpreted as transporting the environment to the audience, who in this case may take on an observing role from their position outside of the referential space. Perspective in relation to the sonic activity in the work can be underlined by placing the material in different areas of the listening space with reference to listening position and facing direction.

Distance/presence

The articulation of distance and nearness is an important aspect of spatio-musical structure that needs particular attention in diffusion. This is especially true when dealing with works which play with psychological space and personal distance. It can be problematic to transfer successfully to the concert hall a composition in which spatial intimacy plays an important role. Although sound material such as whispering and other close-up recordings may indicate a certain proximity to the sound source, the size of the concert hall and the physical distance to the loudspeakers may contradict such an indication. Depending on the acoustics of the hall, the distance to the loudspeakers does not have to extend far before the proportion between direct and indirect sound to reach the ear negates the intimacy suggested in the sound material. Furthermore, the concert hall is a public space which normally does not encourage intimate interaction, something which may affect the listener's receptivity to intimate sound material. Compositions relying on physical closeness are therefore often intended to be listened to in solitude in a private space. However, when it is practically possible to place loudspeakers in the listeners' near field, the range of distances at the diffusionist's disposal can be extensive. Just as closeness can provoke strong reactions within listeners, so can vastness and suggestions of great distances. By consciously working with distance and presence as well as envelopment in sound diffusion, the sound diffusionist can play with the audience's experience of being participants near or inside the sonic environment or observers outside the sound field.

Spectrum

Diffusion systems made up of loudspeakers of different makes and sizes makes the task of spreading the sound material in the listening space more flexible. Compositions of spectral orientation, in particular, may diffuse quite well on their own without any active engagement by the sound diffusionist if the sound system is set up for maximum spectral diffusion. Such a system, however, is not necessarily easy to control and, due to the spectral split caused by different frequency response of the loudspeakers, may be restrictive in the type of coherent spatial trajectories that can possibly be made. That can be rectified by employing high quality loudspeakers capable of delivering the (near) full-frequency range and instead control the spectral diffusion by using filters. When this is done at the mixing console, the diffusionist is able to control dynamically the frequency response of the system and actively vary spectrum as well as loudness for the loudspeaker channels. In digitally controlled systems, different filter settings can quickly and easily be put in place to change the characteristics of the diffusion system to accommodate for different types of composition and to utilise the acoustics of the concert hall in different ways for different works.

Pace/energy

Although electroacoustic music is recorded on a fixed medium with a fixed playback speed, the pacing of spatial movements can still to a certain degree be controlled in diffusion. By increasing the extent of spatial movements, a sense of higher speed is created based on the relationship between distance and duration of movement. Longer movements are perceived as being faster than shorter movements of the same duration. Similarly, fast locational shifting between loudspeakers in different areas of the listening space may increase the energy level experienced in the work. In passages with a high degree of spatial activity, the energy level can be underlined or enhanced in diffusion by rapid displacement of the sound material in positions spread out over a (relatively) large area of the listening space.

5.5.3 Case study: Diffusing *(dis)integration*

My composition *(dis)integration* is interesting as an example for diffusion because it is quite varied in terms of spatial composition. With regards to spatio-structural orientation there are sections of clear extrinsic orientation and others of equally clear intrinsic orientation. The dynamic range of the work is wide, and there is extensive spatial movement, which suggests a very active diffusion, particularly in the opening section of the piece. The density and activity level in much of the work suggests a relatively loud performance in order to ensure a physical sensation of the music.

The listening space I assume for this diffusion plan is a rectangular room of sufficient size to house symmetrically placed pairs of loudspeakers at the rear, to the sides and in front (wide, centred and distant) of the room without being too close to the audience (see figure 5.8). This configuration is, as I know from experience and from the report of others (Harrison 1998: 15), a versatile minimum loudspeaker setup for most conventional concert venues. Adjustments to this performance plan, be they minor or major, will have to be made to accommodate for other types of rooms and sound systems and different audience sizes and placements. I describe the diffusion as I imagine it being heard from the optimum listening position in the room, the centre of the loudspeaker array.

Preparation

A high-level spatio-structural analysis of *(dis)integration* reveals that the piece can be divided into three sections. The first section is largely of extrinsic orientation and spans the first 3'21" of the piece. It is highly energetic with an aggressive staccato stream going through most of the section, and with sound material in varying spatial activity of predominantly extrinsic nature interspersed throughout the section. The second section is also of extrinsic orientation, but has a slower pace and is much less dense, and therefore clearly of a

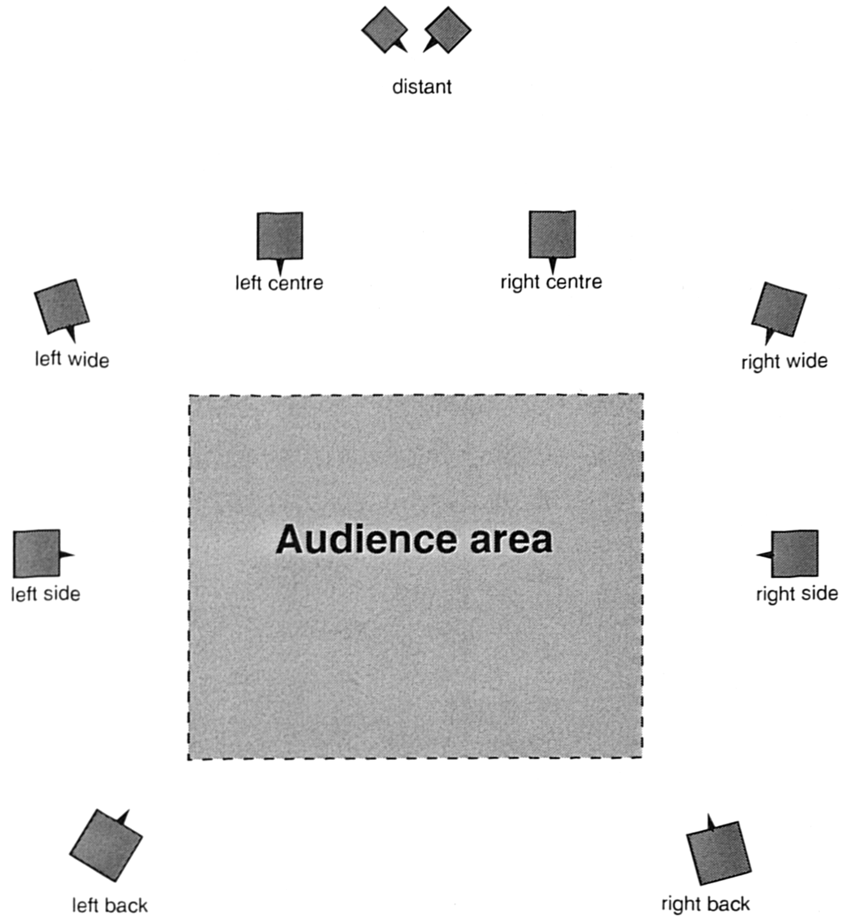


Figure 5.8: Assumed loudspeaker configuration for the diffusion of *(dis)integration.*

contrasting nature to the first section. This section lasts for slightly more than two minutes, until 5'38". There is a gradual transition towards the third section up to this point where a distinct gong-like sound marks an appropriate point to divide the sections. The third section is of intrinsic orientation and consists of a very dense sound mass in the shape of a crescendo-decrescendo. This section lasts from 5'38" until 9'54" where the piece ends with a fade-out. Bear in mind, though, that a sectional division like this is mostly theoretical, and should not necessarily be made explicit in diffusion; transitions between sections need often to be made gradual.

In preparation for diffusing my compositions I normally create a performance score based on a sonogram of the particular work. On this sonogram I indicate key sound events and the intended spatial articulation of them in the concert hall. Appendix IV shows such a performance score for *(dis)integration.*

Section 1

The opening of *(dis)integration* indicates a fast-paced, highly energetic work. This, in turn, suggests a very active diffusion with rapid movements and the use of the full diffusion rig in order to vary locations and trajectories of the sound events as they occur in the composition. Such a diffusion technique will ensure an unambiguous articulation of the pace and energy level of the piece in the listening space.

For the opening of the piece, I choose to position the two channels of the stereo image to the left-side and the right-centre loudspeakers, respectively. During the first two seconds the left channel shifts from the side to the left-centre loudspeaker. This underlines the left-to-right movement in the opening by exaggerating the range of the movement before establishing the stereo image on the centre-front pair of loudspeakers. Already at 0'03", however, there is another, somewhat louder left-to-right moving sound which can be used to expand the sound image further towards the right of the listening space. Keeping the level of the centre-front pair of loudspeakers up, this left-to-right movement is extended by employing the right side speaker. The circling movement which starts at 0'05", entering in the centre of the stereo field and rotating counterclockwise, can then be used to expand the space in the opposite direction, and by following the circular movement all the way around, the full surrounding sound field is established. As the circling sound fades out, the levels of the back, side and front loudspeakers are kept up, with the wide loudspeakers being somewhat louder. The weight is on the front, but the envelopment of the audience is still there, allowing for subtle displacement of sounds in either direction by just slightly adjusting the level for the desired channel when a suitable sound event occurs. Furthermore, in a diffusion rig made up of loudspeakers of different frequency response, employing many different loudspeakers will also ensure a spectral spreading of the sound material which may help clarify a dense sound field.

At 0'11" there is an attack of short percussive sounds which starts in the centre and then bounce back and forth between the horizontal extremes of the stereo field as they fade out shortly after the attack. By increasing the level of the right back loudspeaker from the moment of the attack, the sounds will be perceived as bouncing between the wide left and the right back. The next gesture to catch the ear is a short single-circle movement at 0'21" with its peak on a rapid left-right movement as the rotation ends. Here, a left-back to right-wide movement is implied in the composed space. By going from left rear to the frontal right-wide, the speed of this movement is effectively enhanced because these two loudspeakers are far apart in the listening space and the sound covers a long distance during the course of its short existence.

The sound field is stable for a few seconds until another circular movement starts at 0'30". Compared to earlier movements, this one is somewhat slower. It is timbrally quite sharp and present, and its movement needs only a slight emphasis in the diffusion. An exaggeration of this sound's behaviour would diminish the effect of the sudden disruption at 0'39". Keeping the sound field relatively stable prior to the event will maximise its impact. At this point, the entire sound field is thrown towards the left hand side of the stereo field and quickly fades out. Without compromising the pace, it is soon overlapped by a second onset which quickly fades in in the front of the sound field at 0'41". As the introduction (the first 0'39") suddenly is thrown towards the left and fades away, the rear and side loudspeaker levels should be reduced and the second onset be allowed to fade in on the wide and centre pairs only. It is important at this point that the sound pressure level on the front speakers is high so that the intensity and energy level are not lost when the side and rear loudspeakers are shut off. The overall spectral quality at this re-entry is slightly sharper than in the introduction, something which contributes to a slight increase in perceived intensity level.

With the second onset, the sound field is for a moment settled in the front of the listening space. The left-to-right movement of the percussive staccato-sound fading in from 0'46" needs to be stressed by employing the left-wide loudspeaker before the sound field "collapses" into a narrow, centred sound image at 0'47" where only the centre pair should be used. There are two right-to-left movements at 0'49" and 0'51" which stepwise push the sound image towards the left before a third movement restores the original wide-front stereo field at 0'53". This must be emphasised by a physical widening of the sound field by employing the wide and the side loudspeakers while still keeping the weight at the front.

The activity calms down slightly after 1'00". There are fewer and slower movements about which I choose not to go into further detail as they are not particularly significant structurally. At 1'14", however, there is a rapid right-to-left moving sound which is counterbalanced by a move from left to centre at 1'17" which marks the onset of a more dense sonic texture. These two movements need to be emphasised in diffusion as they comprise an important structural event. There is a sense of instability in the music between 1'17" and 1'22" at which point the onset of a more pronounced return to the dense texture from earlier in the section occurs. The bass drone here has a strong presence, and its ubiquity should be underlined by adding the rear loudspeakers in order to envelop the audience.

From around 1'25" there is a decrease in intensity, the definite turning point being at 1'44" when the staccato stream quickly moves to the left and out of the sound field. This move is particularly important and needs to be emphasised by making the trajectory of the sound long and wide; from right-side to left-wide. The full surround should still be up. As the sound field thins out over the next 80 seconds, a slow shift of the entire sound field towards the distant speakers in the front is appropriate. This forward motion will emphasise the downward movement in spectral space of the drones from 2'00" to 2'40", as well as the overall decrescendo and thinning out of the sound field. A great physical distance to the sound field is achieved and the audience takes on a more observing role than previously, which may lead to a sense of relaxation.

At 3'03", where the sudden attack and recapitulation of the staccato stream occurs, the full surround needs to be employed in order to magnify the impact of this attack and ensure the surprise at its appearance. Some rapid and aggressive physical movements of the sound in the listening space can be made to underline the contrast with the stable sonic texture which preceded it and to contrast what comes after. As section 1 ends at 3'21", the rear, side and distant speakers should immediately be shut off, with the weight being on the wide loudspeakers with the centre pair as underlying support to ensure a relative stability in the sound field.

Section 2

Section 2, which starts at 3'21" and lasts for slightly over two minutes, stands in great contrast to the first section. The virtual space here is sparsely populated, and the reverberation and the distant placement of the sound material in the beginning of the section suggests a vast open space. The most prominent sounds are rocks which are rolling and bouncing in a quite natural but controlled way in the composed space. There are no wide dramatic movements here, but the location of the individual sounds are for the most part easily perceived. Thus, there is an overall focus on extrinsic space; the whole section is of extrinsic orientation. Also in terms of loudness there is great contrast compared to the first section. This section is much quieter than the preceding one, and although there is a gradual increase in loudness over the course of the section, its peak is nowhere near that of the opening section of the piece.

Immediately after the last burst of the first section, the foreground activity in the sound field shifts towards the left hand side. The placement of the sound material is already quite distant in the composed space, and does not necessarily need further enhancement of distance in the diffusion. That could make the contrast to section 1 too great, and the spatial coherence between the two sections might suffer. With the sonic activity starting in the front, the number of loudspeakers employed should slowly increase over the course of the section, and thereby follow the gradual intensification composed into the section in terms of increase in number of sounds and their spreading in the sound field.

As the sonic activity shifts towards the centre at around 3'50" the weight needs to be moved from the wide to the centre pair of loudspeakers. The sound material moves closer in distance in the sound field and the level should be raised slightly to underline this. Over the course of the first minute-and-a-half of the section, the variable placement and movement of the sounds must be underlined by shifting the weight of the sound field between the centre, the wide and the distant loudspeakers. The centre and wide pairs can be employed in more of an *ad lib* fashion than the distant pair, which should only be used at points where distant placement of the sound material is clearly composed into the work, such as at 4'05"-4'10" and at 4'28"-4'30". The side speakers should be added as well, but at a low level in order subtly to add general spaciousness. Their level should be increased slightly when the bass tone fades in from 4'37" in order to make this tone more omnipresent.

A low noise-based drone enters in the background at 4'46", and with the gradual build-up of foreground material since the beginning of the section this point becomes appropriate for adding the rear loudspeakers. This will underline the intensification of the sound material as well as anticipate the final section by enveloping the audience. Simultaneously with the noise-based drone, high-frequency components of similar shape as the rocks are introduced. Together, these two additions expand the spectral space both upwards and downwards. As the number of rocks increases in the sound field, their individual positions become more difficult to perceive, and they gradually fuse into a non-directional granular texture with an increasingly present intrinsic space. The high-frequency components protrude and can still be localised individually. Their position is, however, quite static in extrinsic space. To contribute to the growth, some movement of the entire sound field needs to be added. All the loudspeakers should be sounding and the movement made by varying the level, either individually or in groups to create a sense of the sound mass being a moving organism. The movement should intensify, culminating in a broad articulation of the right-to-left moving sounds at 5'31", 5'34", 5'36" and 5'37". As the gong-like sound appears at 5'38" there should be a relatively quick slow-down of movement in diffusion. The rock-mass moves gradually towards the right from this point until it fades out approximately one minute later, overlapping into section 3.

Section 3

The gong-like sound at 5'38" marks the beginning of the final section. This section is of intrinsic orientation, and although it stands in contrast to the preceding section, there is a relatively smooth transition between the two sections which requires no sudden change in diffusion as was the case with the switch from the first to the second section. The basic layout of this final section is a relatively straight-forward crescendo-decrescendo.

The sound material here is an extremely dense sound mass with minimal locational variation. There is a strong emphasis on intrinsic spatial qualities, and as the sound mass builds up towards its peak, both in spectral span and in magnitude, the onset of spectral components within this sound mass also draws attention to discrete steps in spectral space. The onsets of these spectral components are possible to localise in extrinsic space, but based on the interpretation of the section as a whole, their individual location is not structurally significant and need not be emphasised in diffusion. Only towards the end of the section is there horizontal movement; the right-to-left moving sounds first heard in the beginning of the work reappear and thus create a link to earlier sections.

In accordance with the slow crescendo-decrescendo which characterises the section, the diffusion should also be a slow moving of the sound mass as was done towards the end of the first section. This will further smooth the transition between sections 2 and 3, and underline the transformation from disintegration to integration in the sound world of the piece as is suggested in the title of the composition. A diffusion system which makes possible spectral diffusion is beneficial here. The movement in the diffusion should be intensified in range and pace towards the peak, and similarly follow the decrescendo after the peak by being slower and less extensive. As the sound mass quietens down and becomes less dense, the whole sound field again needs to move slowly forward in the listening space.

Towards the end of the piece there are several right-to-left moving sounds which can be recognised from section 1 and from the end of section 2. These provide important links to earlier parts of the work, and need to be emphasised. The first of these occurs at 8'11", although it is anticipated by a similar sound statically located on the right hand side of the sound field five seconds earlier. The first right-left movements, at 8'11" and 8'47", should be articulated on the wide pair with the side speakers still sounding. The sound field should continuously be moved forward, and when the last drones of the sound mass start sliding downwards in spectral space at 8'53" only the front three pairs should be employed. By the time of the last left-right moving sound at 9'10", only the centre-front and the distant pair should be up with only the distant pair in use from around 9'20" in order to emphasise the extended fade-out. This last diffusion gesture is meant retrospectively to link the decrescendo towards the end of section 1, where a similar move towards frontal distance was made.

6 Perceived space

In any electroacoustic music listening situation, the musical space the listener hears is the combination of composed space and listening space. The musical-aesthetic experience has to do with this resulting *perceived space*, and depends upon a successful interaction between the spaces composed into the work and the space in which the work is listened to. However, these are not the only components on which the spatio-aesthetic experience in music listening is based. The total spatial experience is quite complex and does not confine itself solely to the audible. Visual properties of the listening environment are influential, equally so is the absence or presence of other audience members in the listening situation. Information picked up by other senses such as vision, smell and touch can affect the listener's ability to attend to the music in a concentrated manner, and may underline or contradict spatial cues composed into the musical work.

It is commonly accepted by music practitioners and listeners alike that music conveys meaning. Attempts to define what that musical meaning is have been manifold throughout the history of musical thinking. It is an intricate philosophical problem that I shall not set about discussing here. I simply take for granted that music has meaning. Nonetheless, I shall state that in order for music to convey meaning, the musical elements must be organised *in a meaningful way*. This is where experience and musical training come into play. What is meaningful for one listener is not necessarily so for another, and what means one thing for one listener may mean something different for another. There are more individual differences in the interpretation of artistic meaning than of meaning in everyday communication, but the two are related, and meaning communicated through an artistic medium cannot be studied independently of meaning in general. Hence, space as an element of communication and expression in electroacoustic music cannot be understood independently from space as an element of general interpersonal communication. I therefore propose that the understanding and interpretation of spatio-structural elements in music and the effect of spatial aspects of the listening situation as a whole are largely based on culture-specific learning related to space as a component of interpersonal communication. For example, what is interpreted as uncomfortably intimate for one listener may be an acceptable interactional distance for another. Just as ethnomusical elements are understood differently by listeners native to the particular ethnic group from where the music originated, so is the effect of spatial elements interpreted differently by listeners of different cultural backgrounds. This does not mean that one listener's interpretation is "better" than the other, or that one listener's aesthetic experience is more valid; they are just different.

6.1 Space as communication

Spatial studies in the broader context of interpersonal and environmental relationships come under the socio-psychological heading of *nonverbal communication.* Space communicates and establishes types of relationship between participants in situations of interpersonal interaction, and shapes the individual's relationship to the surrounding natural and cultural environments. The individual can both influence and be influenced by the spatial environment, although not always in an interactive manner. Architecture, for example, is a dominant element in the spatial environment which defines territories and sets up physical boundaries between people and spaces. Particularly in urban areas, where it is inescapable, architecture represents a type of one-way spatial communication which deeply influences and shapes people's lives. Of similar influence on a smaller scale, though highly important in the discussion on musical space and spatial experience in music listening, is the architectural layout and distribution of seats and other objects in the listening space. The proximity of other listeners in a concert situation, the physical proximity to the sound field and the encouragement or discouragement of personal interaction in the listening situation are all factors related to spatial communication that are influential on the musical experience. Of course, most important is the spatial content of the music itself, but this cannot escape the influence of spatial attributes of the listening situation at large which, consciously or unconsciously, affect the state of mind of the listener.

6.1.1 Proxemics

Anthropologist Edward T. Hall focused on intercultural studies of the meaning and use of space in communication. To Hall, culture is defined by communication itself, in which the use of time and the use of space are the fundamental elements of a "silent language" (Weitz 1979: 277; Hall 1990a). It is these silent assumptions about interaction and the other person which separate cultures and give rise to misunderstandings. Hall coined the term *proxemics*, which he defined as "the study of man's perception and use of space [...which...] deals primarily with out-of-awareness distance-setting" (Hall 1979: 293). He described three kinds of spatial organisation:

- *Fixed feature space*, which refers to arrangement of environmental space such as houses, rooms, cities and towns. (In some cultures furniture is also considered fixed);

- *Semi-fixed features*, which are movable objects, such as tables and chairs;
- *Dynamic space*, which has to do with interpersonal distance.

Within these categories, one can distinguish whether space is organised so that it encourages involvement and communication between people (*sociopetal space*) or whether it promotes withdrawal and is organised to keep people apart (*sociofugal space*) (Hall 1979: 302).

Of the three kinds of spatial organisation, dynamic space, or *personal space*, is the most relevant in my discussion since this concerns the distances people unconsciously maintain when they interact. This sense of distance setting, I believe, also comes into play during composition and intentional music listening, and is particularly triggered by referential sound material and familiar spatial configurations.

Hall categorised dynamic space[1] into four distance zones which, in addition to reflecting the relationship between the persons involved, in terms of spoken language also is characterised by type of subject matter (Harper et. al. 1978: 247-48; Hall 1990b: 116-25):

- *Intimate distances* span from body contact and up to a distance of about half a metre. This zone is characterised by maximum olfactory and thermal sensations and great perception of fine visual detail. Visual perception of larger objects (such as the head) is blurred. Speech is generally of little importance, but when it occurs it is kept at a very low level and tends to be on a secret or confidential subject matter.
- *Personal distances* can range from 0.5 m to about 1.25 m. Here, visual perception is no longer distorted and thermal sensation is usually eliminated. Interactants can still touch each other, but only by stretching out their arms or legs. Voice level tends to be moderate and the conversational topic tends to be of a personal nature.
- *Social distances* range from 1.25 m to 3.5 m. Vision and hearing are the two primary senses here. Visual contact is important when interacting at social distances. Speech is normally on a non-personal matter, and may be public information for others to hear.
- *Public distances* range from 3.5 m and up. Eye contact is no longer socially required. The voice needs to be raised or, at greater distances electroacoustically amplified, so that subtleties in meaning expressed through the normal voice are diminished or lost. The spoken messages at such distances are generally addressed to a group of listeners. At these distances most of the nonverbal communication shifts to broader gestures and body stance.

[1] Hall also uses the term *informal space* in relation to dynamic space to underline that this category of spatial communication is unstated and largely out of awareness (Hall 1990b: 111-12).

The distances cited above refer to findings based on studies of middle-class Americans of Northern European descent, and are not universal across cultures. However, from his research on different cultural groups, Hall found that although the distances are culture specific, the four categories of social distances are common. This finding is also backed up by studies on animal behaviour.

Hall suggests that there is no fixed, universal distance-sensing mechanism. All the senses are involved in setting distance, but to what degree each of the senses participates varies from culture to culture. For example, some cultures are visually oriented while others set distance based on olfactory or thermal sensations. Hall concludes that people from different cultures inhabit different sensory worlds. Not only are spaces structured differently, they are also experienced differently. He states, based on findings from ethologists and animal psychologists, that (Hall 1979: 296):

1. each organism inhabits his own subjective world, which is a function of its perceptual apparatus, and the arbitrary separation of the organism from that world alters context and in so doing distorts meaning; and

2. the dividing line between the organism's internal and external environment cannot be pinpointed precisely.

In addition, the circumstances of spatial interaction are influential as they can cause perceived distances to expand or contract according to the situation (Hall 1979: 303).

6.1.2 Territoriality

Personal space is a changeable, psychological space which moves with the individual and varies in size according to situation. Personal *territories* differ from personal space in that they are relatively stationary and do not necessarily follow the individual person. These territories refer to various types of marked-off areas within which intrusion will be reacted to. Common territorial behaviour is, for example, setting up a fence around one's garden or leaving a jacket or a book when temporarily vacating one's place at the library. Equally significant, but much more fragile, is territorial behaviour in situations of personal interaction, such as conversations, where subtle boundary markers are maintained only for the duration of the interaction.

Lyman and Scott (1967: 237-43) categorised human territories into four groups which can be seen in parallel to Hall's four categories of distance zones:

- *Body territory* is similar to the concept of personal space, where entry is only accepted for persons with which the individual has a special relationship. In addition to the (physical) human body itself, Lyman and Scott also include what they term *internal space* and *external space* in their notion of body territory. External space refers to what is most commonly thought of as personal space, namely a non-physical space immediately surrounding the body, which shrinks and expands according to situation. Internal space is of an internal psychological nature, and is the most private and intimate of spaces. Violations of internal space are carried out by various degrees of domination, such as hypnosis, persuasion or involuntary perception of intimate detail.

- *Interactional territory* is a space for social interaction in which there are restrictions as to who is allowed to participate. These territories are mobile and fragile, and their boundaries are invisible and based on (unwritten) social rules commonly understood by the participants. An attempt by a newcomer to join a conversation, for example, may be seen as an intrusion which threatens the ongoing interaction by requiring it to start over again, change subject or come to an abrupt end. The first of these three reactions shows full acceptance and inclusion into the territory, the second shows acceptance, but not on the premises on which the territory was initially established, while the third is a clear rejection demonstrated by a destruction of the interactional territory.

- *Home territory* refers to areas where the individual, or group of individuals, enjoy a relative freedom of behaviour and sense of intimacy and control. There is a certain overlap between public territory and home territory in that the latter does not only concern the home in its most common sense, but also such places as club houses (and genre-specific music venues). A home territory that is publicly accessible is often marked off by its members by use of jargon or specific style of clothing, for example. The boundaries of the home, in the meaning of the individual's private living-quarter, is normally clearly and unambiguously marked off by physical markers and only given access to by invitation.

- *Public territory* is any location which is accessible to all, such as public streets, parks and vehicles for public transport. The behaviour in such territories is usually restricted to what is considered appropriate by those who are the most frequent visitors to these places. Violation of behavioural expectation is reacted to by various forms of exclusion or force, such as a denial of access to a specific area or, in the public territory on the larger scale, imprisonment of lawbreakers.

6.2 Space as an element of expression and communication in electroacoustic music

Spatial content in music can be expressed in terms of spatial references, spatial interrelations among the sound material, extent and intensity of spatial movement and clarity and definition in spatial placement of the sound material. These are all compositional elements that are at the composer's disposal in the expression and communication of musical ideas and artistic meaning. I believe that the composer's organisation of these elements in the musical work and the extension and articulation of them in sound diffusion are based on learned aspects of general interpersonal communication in terms of personal space and territoriality. This is particularly the case when working with referential spaces and when dealing with intimacy and distance as structural elements in the music.

As Hall (1979, 1990a, 1990b) on several occasions points out, many of the communicational aspects that have to do with space and distance setting are so deeply embedded in the individual's personality that they exist out of awareness and are rarely subject to conscious thought. To be aware of all the spatial connotations of the composed space is therefore difficult, if not impossible. However, knowing that such unconscious factors exist is fundamental to integrating space successfully into the structure of musical works, as well as to interpreting spatial elements in music listening.

6.2.1 Spatio-musical expression

The composer's spatial decisions are defined in the composed space and diffusion of the work. Many of the aspects addressed in the discussion on composition and sound diffusion in the two previous chapters, therefore, can be understood in terms of communication and expression of personal space and territoriality.

The dimensions of the sound field itself in terms of size and extent is a substantial factor in spatial experience. There is a significant experiential difference between a small enclosed space and a vast non-reverberant field, for example. The former can be used to create a sense of claustrophobia or crowding, while the latter may be associated with desolation and smallness. The dynamism of electroacoustic space makes possible variations in spatial dimensions, so that different types and sizes of spaces may be used at different points in the composition. This can be done in order to transport the listener to different sound environments and play on associations and reactions to different spaces in the musical structure. Our ability aurally to detect type and layout of spatial environments is quite sophisticated: just by listening we can determine whether a space is empty or furnished and what kind of material the walls are made of, for example.

As pointed out in the discussion on composed space, different directions and extents of movement mean different things. Sounds approaching from the distance, for example, lead to different psychological reactions compared with sounds that are moving into the distance or passing by. Especially when sounds are experienced as violating the body territory, a strong reaction may be provoked. Speed of movement is an influential factor, particularly when sounds are moving towards the listener. Sounds approaching at a high speed can provoke human emergency reactions, such as fear and the mobilisation for fight or flight. A similar reaction can occur when facing sounds of great magnitude or sounds that refer to dangerous, threatening or uncomfortable situations.

Particularly when working in surround sound, spatial cues known from our cultural and natural environments can be utilised in electroacoustic works with a high degree of accuracy and realism. The combination of referential features of the sound material and spatial behaviour in composed space may lead to spatial experiences that are extra vivid due to familiarity of the sound environment gained from real-world experience. Visualisation is key in this regard. Spaces and sounds that are known from personal real-life experience are much easier to visualise than unfamiliar material, and relationships among known spaces are easier to understand. However, relying on such cultural knowledge as spatio-structural functions in a work will almost certainly ensure a great variety of listener interpretations. As someone who grew up in a small coastal community, I associate the sounds of the ocean with immensity and powerlessness, fear and respect, love and dependency. I cannot expect an in-land city dweller, for example, to have the same reactions to the ocean sounds of the opening of my composition *Ebb* as someone from a similar background as myself. Similarly, for somebody who is well travelled, the voice material in my work *Itinera* may be recognised as announcements from airports and train stations and thereby be connected with associations with excitement, expectation, exhaustion, waiting, jet-lag or whatever seems to categorise the individual's travel experiences. Listeners who do not recognise or associate these sounds with travel will most likely interpret the work in a different manner.

The voice is in general a sound source which carries several levels of meaning for human beings. As mentioned earlier, the mere sound of the voice, particularly of somebody familiar, is a very effective cue for judging physical distance with great accuracy. The whisper, for example, is an obvious spectral indication of intimate distance. For this cue to be successful in electroacoustic sound reproduction, however, it is crucial that the perceived distance to the loudspeaker roughly corresponds with the distance indicated by the spectral quality of the voice and thereby does not introduce contradictory spatial cues.

The subject matter of spoken language also carries distance information, as pointed out above in connection with Hall's four distance zones. Obviously, knowledge of the language is a prerequisite for the subject matter to convey spatial meaning, particularly with intimate and personal distances where the conversational topic tends to be confidential and personal in nature. When encountering an unknown language, a similar level of intimacy cannot be achieved because the semantic layer of the speech is not accessible. Even if spectro-morphologically the voice bespeaks intimacy, but the linguistic meaning is not understood and cannot be inferred from context, a psychological distance persists and may give rise to a sense of spatial contradiction. Likewise, intimate distances are characterised by non-linguistic vocal sounds which, similarly, can be contradicted by acoustical cues signalling social or public distances.

Several aspects of spatial information in the voice are taken into account in *kernelMotion*, which forms the second part of my composition *Itinera*. The work deals with the idea of travelling, which is made apparent through the use of sounds of car traffic, trains and aeroplanes. As mentioned above, the voice materials in the work are recordings of various loudspeaker announcements at airports and train stations. The announcements are spoken in German, Russian and English, and for most listeners at least one of these languages will be unfamiliar. For the most part, these recordings have the evident quality of being emitted from loudspeakers, something which results in a depersonalisation of the voice—a clear indication of public distance. This is confirmed by the messages being public information that is communicated in a formal, unemotional way. The placement of the vocal sounds in the background in the mix, therefore, is more an exaggeration of the distance information already defined by the depersonalisation and the public nature of the messages than a primary spatial cue.

6.2.2 Spatio-musical experience

The listening experience depends on how spatial cues in the music are understood in terms of their communicational role in the musical context. Musical training and familiarity with the electroacoustic genre is often necessary in order to connect and comprehend the various structural levels of the work so that an understanding of it can be reached. Some level of understanding, I believe, is a prerequisite for aesthetic experience, which is achieved when there is a satisfactory balance between surprise and fulfilled expectations in listening.

Expectations on a detailed structural level are a result of stylistic familiarity. However, much electroacoustic music contains referential sound material or spatial configurations which convey meaning also for the novice listener. The work on auditory scene analysis and the Gestalt principles of perceptual grouping show that auditory information is perceived as coherent patterns rather than isolated sounds and events. Knowledge of electroacoustic music often determines how and which sounds and events are grouped. Aesthetic experience of electroacoustic compositions is possible for novice and expert listeners alike, but the structural elements on which the understanding of the music and the aesthetic experience are based may differ with training and familiarity.

The listening situation

The influence of spatial aspects of the listening situation at large must be considered in a discussion on spatio-musical experience as it affects the mood and state of mind of the listener as well as the perception of spatial information in the music. There is a considerable contextual difference between solitary listening in a relaxed environment at home and making the effort to go to a concert hall to listen to music in the presence of other people. In addition to differences in sound system and acoustics between home and concert hall, the social situation of listening in a public space affects the listener's receptivity to spatial information and perception of the communicational aspects of space in the music. The architectural layout and seating arrangement of the concert hall and the social rules for the concert situation are important factors here.

In the light of the above discussion on personal space and territoriality, certain observations about some of the extra-musical aspects of the concert situation can be made. In the concert tradition of Western art music, audience participation is generally not desired, and an arrangement of the seats in straight rows all facing forward towards a stage suggests a one-directional flow of communication from performer(s) to audience. The dimming of the lights in the audience area underlines this by making it more difficult to have eye contact with other members of the audience, thus reducing the visual influence of other audience members and minimising the premises for interpersonal communication. The lighting on stage usually remains fixed, signalling a non-changing relationship between performers and audience. With the audience seated, each individual is allocated a certain personal territory which he or she is only allowed to leave on signal (during intermission and when the performance is over). Such an arrangement, where the seats generally are considered as fixed features and physical space is limited, encourages the listener to listen in a concentrated manner and helps suppress any overt physical response to the music. In contrast, the typical rock concert situation encourages a certain audience participation and interaction, which is effectively achieved by not having seats available at all, and is further emphasised by a dynamic use of lighting of varying brightness, colours and directions, both on stage and towards the audience.

The acousmatic concert setting is similar to the traditional Western instrumental concert in many ways, but with the significant difference that there is generally no performer in front of the audience. One consequence of this lies in the lighting arrangement. Often, acousmatic concerts take place in near complete darkness in order to keep visual distractions to an absolute minimum and emphasise the purely auditory nature of the music. This may encourage concentrated listening, but may also decrease the ability to localise sounds, because, as mentioned earlier, localisation is more accurate in the light than in the dark, even for invisible sounds. An alternative is to employ low lighting in the audience area.

This, however, may lead to distractions: there are always listeners who interpret lighting as a permission to read the programme notes during the performance or even communicate with other audience members, verbally or non-verbally. A tried solution to this is to keep the audience area dark and instead illuminate the loudspeakers. This also has negative consequences because it points out the loudspeakers as the actual sound radiators and may direct localisation towards them. However, illuminating the loudspeakers may make it easier to follow with the eyes invisible sound trajectories, and thereby improve localisation. A fourth option, then, is to vary the lighting during the performance in a more or less choreographed way. A visual element is thereby added to the performance, which, unless the work was composed with a specific lighting in mind, involves another level of musical interpretation (in addition to the sound diffusion strategy). Underlining the "atmosphere" in the music with variations in lighting may or may not aid the listener in the understanding and interpretation of the musical work. It can be done subtly with great success, but if the changes in lighting become too abrupt or are too extensive, it can easily become a distraction.

Another consequence of the acousmatic concert setting is the seating arrangement. In concerts involving a traditional sound diffusion system, i.e. where there are more loudspeakers in front of the audience than behind and around, the seating arrangement is normally the same as for instrumental concerts. The audience are all facing the same way, and the diffusionist can treat the directions as the same for everyone. Because there is no performer to watch, however, the audience do not necessarily have to sit in a certain direction, and alternative seating arrangements, such as the audience placed in a circle or divided into groups that are facing each other, are sometimes used, particularly in connection with surround-sound compositions. This has serious implications for the perception of spatio-structural elements in the music. With the audience facing in different directions, what is in front for some listeners becomes behind for others and, with a circular seating arrangement, what is a side-to-side movement for some listeners is a front-to-back movement for others. The social and metaphorical significance of certain directions of placement and movement of sound material can no longer be relied upon as structural means. From the composer's point-of-view this may be intended and called upon for some compositions in order to break free of directional connotations. However, such a democratisation of spatial directions will only be successful if there are no seats in the listening space, so that the listeners are free to move around inside the sound field during the performance, or if the music is consciously composed without any hierarchical organisation of spatial direction. Whether the listeners are willing to move about or prefer to sit down to concentrate on listening is partly a matter of social confidence and security within the territory defined by the acousmatic concert situation. More importantly, though, it has to do with whether the music suggests physical movement or if it demands attentive listening. Most music within the acousmatic genre is of the latter kind, which is also normally expected by the audience.

A sense of *crowding* may come into play in the concert situation. Crowding has to do with the overall level of social stimulation, and the psychological reaction to it is an intensification of feelings which can be good or bad depending on whether the situation is good or bad (Weitz 1979: 284-85). More people at a party, for example, may lead to a good feeling, while more people on the rush-hour train may lead to a bad feeling. Similarly, a full concert hall may positively affect performer(s) and audience alike, while a nearly empty hall, in addition to the acoustical consequences, may have a negative influence on the motivation and concentration of the sound diffusionist and on the listeners' receptivity to spatial cues in the music. It is possible that intimacy in the concert hall, i.e. a full hall where people sit close enough to sense body heat and/or cannot avoid physical contact, may increase listeners' receptivity to intimacy in the music. A nearly empty hall, on the other hand, where people often choose to sit far away from each other may similarly contradict closeness and intimacy in the music[2].

According to Hall (1990b: 115), each individual has a number of *situational personalities* which are associated with responses to the four distance zones. These situational personalities are highly individual and are often not equally developed for each distance zone. For example, some people are particularly uneasy in intimate situations, while others may be afraid of making a presence in a public space. I believe these situational personalities dictate listeners' responses to the listening situation and to certain types of spatial information found in electroacoustic music, particularly in relation to intimacy and distance. Most people are more receptive to closeness and intimacy in private than in public settings. Compositions which play on intimacy are therefore often intended for listening in private spaces. However, when it is practically possible to place loudspeakers in the listeners' near field in the concert hall, works which play on intimacy may be disquieting for some listeners because the intimacy cues are experienced in a public space. Violations of the internal space of the body territory may be especially uncomfortable in the concert hall due to the unstated behavioural restrictions of the concert situation; to turn to escape or other physical defence strategies is generally not accepted.

[2] A nearly empty concert hall is almost always more reverberant than a full hall. More reverberation indicates distance and makes it difficult to create a sense of intimacy and closeness in diffusion unless the loudspeakers are close to the listener. A combination of factors may therefore explain differences in perceived intimacy in listening.

Conclusions

This thesis has shown that *musical space* is a wide term which encompasses aspects of electroacoustic music at all stages of its creation, presentation and appreciation. It has shown that space is an element that *cannot be avoided* when dealing with electroacoustic music. This proposition rhymes with Edward T. Hall's statement that "virtually everything that man is and does is associated with the experience of space" (Hall 1990b: 181). Thus, it is quite likely, as I remarked in the introduction, that when a limited awareness of space is displayed among listeners and musicians, it is due to its ubiquitous nature. Space is always present, and is therefore taken for granted.

Spatial decisions are constantly made in the process of composing electroacoustic music, from choosing individual sounds to creating virtual spaces on the basis of interrelations among the sound materials. As argued in Chapter 6, many of the spatial decisions that are made in composition and performance are based on learned cultural aspects of spatial communication in general, and often take place unconsciously. However, because space is such an omnipresent feature of communication in music, it demands the composer's awareness. In order to become aware of the various aspects of space, then, it is essential to know what these aspects are and how they relate in the musical context. A main objective for this thesis, therefore, was to reveal the various facets of space in relation to composition, performance and perception of electroacoustic works. New aspects of space have been uncovered in each chapter, and have added to the complexity of the concept of musical space. An overview of the various components of space in electroacoustic music can be seen in figure 6.1.

Perception and technology

Throughout this thesis I have placed much emphasis on two forms of perception. The primitive perceptual processes that are basic to audition form the underlying foundation, while a higher-level perception based on musical-aesthetic experience—perceived space, in this context—is the overlying concept for the experience of space in electroacoustic music. The electroacoustic genre is inextricably linked with technology, and aesthetic choices in the composition process and the concert situation are often shaped by the possibilities and limitations of the technology at hand. There are numerous examples in the music literature of compositions that are based on some process or compositional principle worked out on the

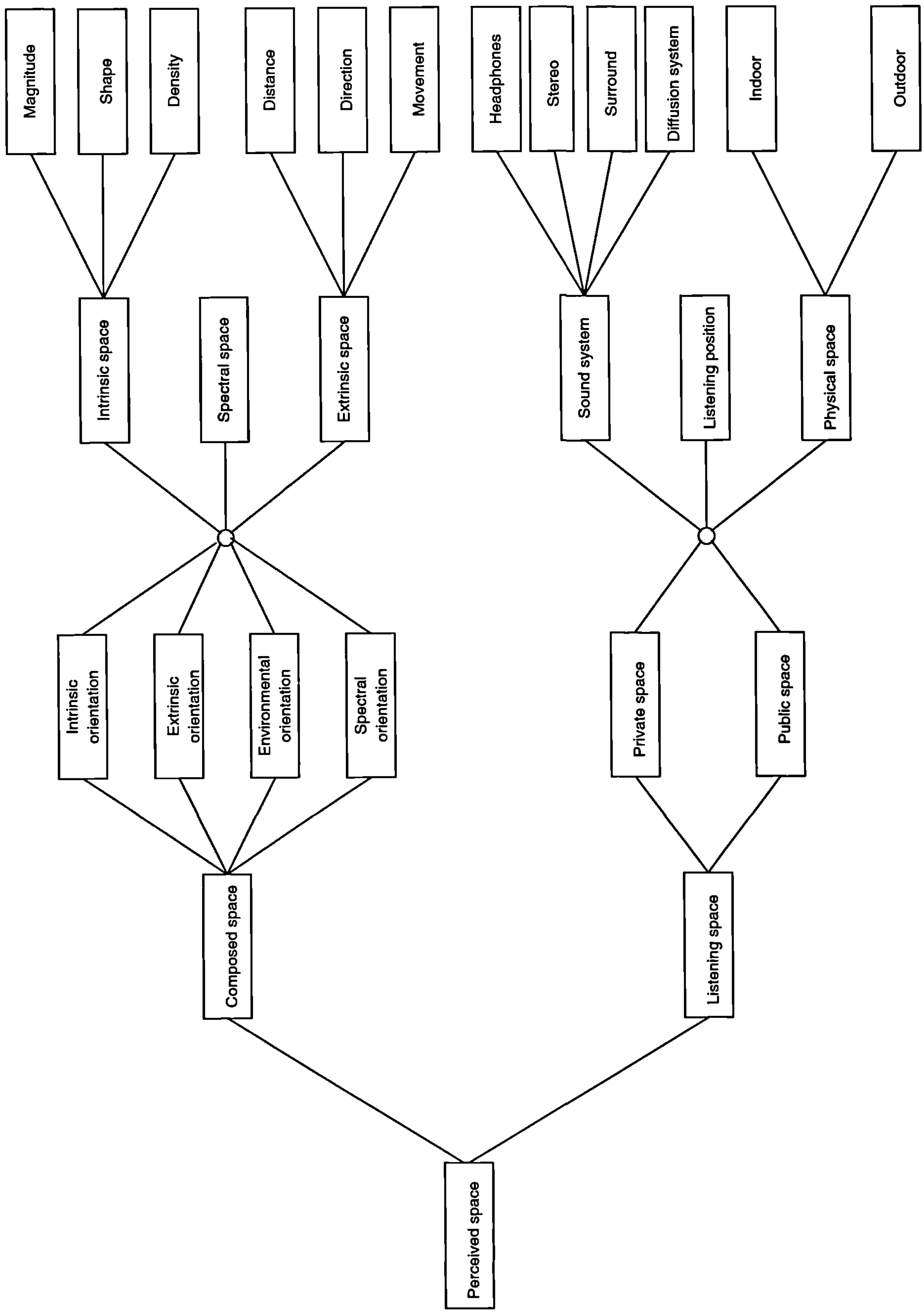


Figure 6.1: Overview of space in electroacoustic music.

computer or with the aid of some other aesthetically "neutral" tool. In many cases, these processes and principles are not shaped by perceptual considerations, but are instead determined by technological potential alone. By examining spatial composition and listening on the basis of perception, but in relation to the technology involved in the creation and realisation of electroacoustic works, I have attempted to show the aesthetic and communicational power of space in music, and how compositional choices must be based on both perceptual and technological considerations in order to utilise this power. Space can then become an essential element of electroacoustic music which can be integrated into the compositional structure as a primary carrier of meaning in the work. Different aspects of space, such as extrinsic space, referential space and spectral space, ensure the potential and versatility of space as a key element in artistic expression and communication. The spatio-structural approach to compositional thinking, as outlined in Chapter 4, is a helpful tool in this regard as it can be used to heighten the awareness of space in the creation of musical work. Spatio-structural theory is not meant to be a recipe for composition, but as a framework for spatial thinking it can aid in the structural organisation of space in the composition process and be of use when preparing for diffusion of electroacoustic works. At this point it is a framework in its infancy, but in it I see a very useful potential. For the spatio-structural approach to develop into a full-fledged analytical system, however, it is essential to expand the framework to reflect the close (structural) relationship between space and sound. Denis Smalley's spectromorphological theory comes to mind as a likely resource in this respect.

With a rapidly growing interest in surround sound in the commercial music industry, the stereo format is likely to be marginalised and eventually phased out as a music distribution format. This has dramatic consequences for electroacoustic music where stereo-based sound diffusion is standard for concert performance, and compositional decisions are often made with such diffusion in mind. The shift towards a multi-channel sound-carrier medium is therefore likely to lead to a change in compositional methods and performance practice in electroacoustic music. This change will no doubt take place gradually, but an awareness of the consequences of the new medium is crucial for electroacoustic composers in order to keep up with the expectations and technological possibilities of the record buying audience. The surround-sound experience is no longer a unique aspect of the electroacoustic music concert. In relation to this, compositional tools in the form of specialised software and hardware for spatial composition need to be developed. Computer applications that make the task of manipulating and organising sound material for surround sound as comfortable as for stereo are needed in order to simplify the composition process so that less time is spent on technical trivialities. I believe that this thesis highlights aspects of space that need to be easily manipulated and controlled in the process of composing and diffusing electroacoustic works.

Conclusions

I have underlined the referential nature of space on several occasions. It is particularly emphasised in Chapter 6 where the communicational aspect of space in electroacoustic music is explained on the basis of everyday spatial interaction. I strongly believe that the experience of music, and of electroacoustic music in particular, is influenced and shaped by associations with knowledge from everyday existence. This is most obvious with regard to works based on recognisable environmental sound material, where the source recognition process links sounds to known objects or situations, but also the experience of a synthesised virtual space is based on reality. Risset (1996: 29) remarks that "our perception has a strong tendency to assimilate the unknown to the familiar." In the case of (abstract) synthesised material we tend to attribute known action or gesture to the sound in order to detect its possible cause in place of an unrecognisable source. The inclination to categorise sounds in the source recognition process (p 38) supports this view, as similarly does the phenomenon of spatial impression (p 39), which is formed based on the real-life experience of enclosed and open physical spaces. Beyond simply linking recognisable sounds to known objects and situations, however, is the deeper symbolic meaning connected to certain sounds and sonic configurations. The symbolic and metaphorical use of space is a highly relevant discussion with regard to the artistic potential of three-dimensional surround sound and of musical space in general.

Bibliography

Adams, G. 1998. "The Room Environment." In J. Borwick, ed., *Loudspeaker and Headphone Handbook.* Oxford: Focal Press.

Ahnert, W., and Steffen, F. 1999. *Sound Reinforcement Engineering.* London: E & FN Spon.

Auerbach, C., and Sperling, P. 1974. "A Common Auditory-Visual Space: Evidence for its Reality." In *Perception and Psychophysics,* 16:1. Cited in Hollander (1994).

Austin, L. 2000. "Sound Diffusion in Composition and Performance: An Interview with Denis Smalley." In *Computer Music Journal,* 24:2. Cambridge: MIT Press.

Barrière, F. 1998. "Diffusion, the final stage of composition." In F. Barrière and G. Bennett, eds., *Composition/Diffusion in Electroacoustic Music.* Bourges: Mnémosyne.

Benade, A. H. 1990. *Fundamentals of Musical Acoustics.* New York: Dover.

Beardsley, M. C. 1979. "Aesthetic Experience Regained." In W. E. Kennick, ed., *Art and Philosophy: Readings in Aesthetics.* New York: St. Martin's Press.

Berry, W. 1987. *Structural Functions in Music.* New York: Dover.

Blauert, J. 1997. *Spatial Hearing: The Psychophysics of Human Sound Localization.* Cambridge, Massachusetts: MIT Press.

Bodin, L. G. 1997. "Kartläggning av den estetiska profilen: ett utkast till en praktiskt-orienterad estetisk analys av elektro-akustisk musik." In *Nutida Musik,* 40:1.

Boeswillwald, P. 1998. "Composition and Diffusion." In F. Barrière and G. Bennett, eds., *Composition/Diffusion in Electroacoustic Music.* Bourges: Mnémosyne.

Boring, E. G. 1926. "Auditory Theory with Special Reference to Intensity, Volume, and Localization." In *The American Journal of Psychology,* 37:2.

Brant, H. 1998. "Space as an Essential Aspect of Musical Composition." In E. Schwartz and B. Childs, eds., *Contemporary Composers on Contemporary Music.* New York: Da Capo Press.

Bregman, A. S. 1994. *Auditory Scene Analysis: the Perceptual Organisation of Sound.* Cambridge, Massachusetts: MIT Press.

Butler, R. A. 1973. "The relative influence of pitch and timbre on the apparent location of sound in the median sagittal plane." In *Perception & Psychophysics*, 14:2.

Camilleri, L. and Smalley, D. 1998. "The Analysis of Electroacoustic Music: Introduction." In *Journal of New Music Research*, 27:1-2.

Chadabe, J. 1997. *Electric Sound: The Past and Promise of Electronic Music.* New Jersey: Prentice-Hall.

Chion, M. 1988. "Les deux espaces de la musique concrète." In *L'espace du son. Ohain:* Musiques et Recherches.

Chion, M. 1994. *Audio-Vision: Sound on Screen.* New York: Columbia University Press.

Chowning, J. 1999. "Perceptual Fusion and Auditory Perspective." In P. R. Cook, ed., *Music, Cognition, and Computerized Sound.* Cambridge, Massachusetts: MIT Press.

Cross, I. 1998. "Music Analysis and Music Perception." In *Music Analysis*, 17:1. Oxford: Blackwell.

Delalande, F. 1998. "Music Analysis and Reception Behaviours: *Sommeil* by Pierre Henry." In *Journal of New Music Research*, 27:1-2.

Deutsch, D. 1982. "Grouping Mechanisms in Music." In D. Deutsch, ed., *The Psychology of Music.* New York: Academic Press.

Dewey, J. 1997. "The Aesthetic in Experience." In S. Feagin and P. Maynard, eds., *Aesthetics.* Oxford: Oxford University Press.

Doherty, D., and Berry, R. 1998. "Sound Diffusion of Stereo Music over a Multi-Loudspeaker Sound System: From First Principles Onwards a Successful Experiment." In *Journal of Electroacoustic Music*, 11. London: Sonic Arts Network.

Emmerson, S. 1999. "Aural Landscape: Musical Space." In *Organised Sound*, 3:2. Cambridge: Cambridge University Press.

Francès, R. 1988. *The Perception of Music.* Hillsdale, NJ: Erlbaum.

Gleitman, H. 1991. *Psychology.* New York: Norton.

Gromko, J. E. 1993. "Perceptual Differences Between Expert and Novice Music Listeners: A Multidimensional Scaling Analysis." In *Psychology of Music*, 21.

Hall, E. T. 1979. "Proxemics." In S. Weitz, ed., *Nonverbal Communication: Readings with Commentary.* New York: Oxford University Press.

Hall, E. T. 1990a. *The Silent Language.* Reprint of original 1959 edition. New York: Anchor Books.

Hall, E. T. 1990b. *The Hidden Dimension.* Reprint of original 1966 edition. New York: Anchor Books.

Handel, S. 1989. *Listening: An Introduction to the Perception of Auditory Events.* Cambridge, Massachusetts: MIT Press.

Harley, M. A. 1993. "From Point to Sphere: Spatial Organization of Sound in Contemporary Music (after 1950)." In *Canadian Music Review*, 13.

Harley, M. A. 1994. *Space and Spatialization in Contemporary Music: History and Analysis, Ideas and Implementations.* PhD dissertation. Montreal: McGill University.

Harley, M. A. 1998. "Spatiality of Sound and Stream Segregation in Twentieth-Century Instrumental Music." In *Organised Sound*, 3:2. Cambridge: Cambridge University Press.

Harper, R. G., Wiens, A. N., and Matarazzo, J. D. 1978. *Nonverbal Communication: The State of the Art.* New York: John Wiley & Sons.

Harrison, J. 1998. "Sound, space, sculpture: some thoughts on the 'what', 'how' and 'why' of sound diffusion." In *Organised Sound*, 3:2. Cambridge: Cambridge University Press.

Harrison, J. 1999. "Imaginary Space: Spaces in the Imagination." In *Proceedings of the Australasian Computer Music Conference 1999.* Wellington: Victoria University of Wellington.

Hollander, A. 1994. *An Exploration of Auditory Shape Perception.* Master's thesis. Seattle: The University of Washington.

Hollander, A., and Furness, T. 1995. *Perception of Virtual Auditory Shapes*. Seattle: The University of Washington. Published on the world wide web: `http://www.hitl.washington.edu/publications/hollander`.

Holman, T. 2000. *5.1 Surround Sound: Up and Running*. Boston: Focal Press.

Keane, D. 1986. "At the Threshold of an Aesthetic." In S. Emmerson, ed., *The Language of Electroacoustic Music*. London: Macmillan.

Kendall, G. S. 1991. "Visualization by Ear: Auditory Imagery for Scientific Visualization and Virtual Reality." In *Computer Music Journal*, 15:4. Cambridge, Massachusetts: MIT Press.

Kendall, G. S. 1995. "A 3-D Sound Primer: Directional Hearing and Stereo Reproduction." In *Computer Music Journal*, 19:4. Cambridge, Massachusetts: MIT Press.

Kendall, G. S., Martens, W. L. and Decker, S. L. 1991. "Spatial Reverberation: Discussion and Demonstration." In M. V. Mathews and J. R. Pierce, eds., *Current Directions in Computer Music Research*. Cambridge, Massachusetts: MIT Press.

Kivy, P. 1979. "Aesthetic Perception." In W. E. Kennick, ed., *Art and Philosophy: Readings in Aesthetics*. New York: St. Martin's Press.

Kupper, L. 1998. "Space Perception in the Computer Age." In *L'Espace du Son*. Ohain: Musiques et Recherches.

Lewis, A. 1998. "Francis Dhomont's *Novars*." In *Journal of New Music Research*, 27:1-2.

Lyman, S. M. and Scott, M. B. 1967. "Territoriality: A Neglected Sociological Dimension." In *Social Problems*, 15:2. The Society for the Study of Social Problems.

MacDonald, A. 1998. "Performance Practice in the Presentation of Electroacoustic Music." In *Journal of Electroacoustic Music*, 11. London: Sonic Arts Network.

Mackenzie, G. W. 1969. *Akustik*. Stockholm: P. A. Norstedt & Söner.

Malham, D. G., and Myatt, A. 1995. "3-D Sound Spatialization Using Ambisonic Techniques." In *Computer Music Journal*, 19:4. Cambridge: MIT Press.

Malham, D. G. 1998. "Approaches to spatialisation." In *Organised Sound*, 3:2. Cambridge: Cambridge University Press.

Manning, P. 1993. *Electronic and Computer Music*. Oxford: Clarendon Press.

Moore, F. R. 1991. "Spatialization of Sounds over Loudspeakers." In M. V. Mathews and J. R. Pierce, eds., *Current Directions in Computer Music Research.* Cambridge, Massachusetts: MIT Press.

O'Connor, N., and Hermelin, B. 1978. *Seeing and Hearing and Space and Time.* London: Academic Press.

Perrott, D., Musicant, A., and Schwethelm, B. 1980. "The expanding image effect: The concept of tonal volume revisited." In *Journal of Auditory Research,* 20. Cited in Hollander (1994).

Pierce, J. 1999. "Hearing in Time and Space." In P. R. Cook, ed., *Music, Cognition, and Computerized Sound.* Cambridge, Massachusetts: MIT Press.

Poldy, C. A. 1998. "Headphones." In J. Borwick, ed., *Loudspeaker and Headphone Handbook.* Oxford: Focal Press.

Rasch, R. A., and Plomp, R. 1982. "The Listener and the Acoustic Environment." In D. Deutsch, ed., *The Psychology of Music.* Academic Press.

Repp, B. H. 1987. "The Sound of Two Hands Clapping: An Exploratory Study." In *Journal of the Acoustical Society of America,* 81:4. Cited in Handel (1989).

Repp, B. H. 1993. "Music as Motion: A Synopsis of Alexander Truslit's (1938) *Gestaltung und Bewegung in der Musik.*" In *Psychology of Music,* 21.

Risset, J.-C. 1991. "Paradoxical Sounds." In M. V. Mathews and J. R. Pierce, eds., *Current Directions in Computer Music Research.* Cambridge, Massachusetts: MIT Press.

Risset, J.-C. 1996. "Real-World Sounds and Simulacra in my Computer Music." In *Contemporary Music Review,* 15:1. Amsterdam: Harwood.

Roads, C. 1996. *The Computer Music Tutorial.* Cambridge, Massachusetts: MIT Press.

Roemer, M. 1997. "The Surfaces of Reality." In S. Feagin and P. Maynard, eds., *Aesthetics.* Oxford: Oxford University Press.

Rogers, M. E., and Butler, R. A. 1992. "The Linkage Between Stimulus Frequency and Covert Peak Areas as it Relates to Monaural Localization." In *Perception & Psychophysics,* 52:5.

Rumsey, F. 2001. *Spatial Audio.* Oxford: Focal Press.

Schafer, R. M. 1991. "Acoustic space." In *L'Espace du Son II*. Ohain: Musique et Recherches.

Schafer, R. M. 1994. *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester: Destiny Books.

Shepard, R. 1999. "Cognitive Psychology and Music." In P. R. Cook, ed., *Music, Cognition and Computerized Sound: An Introduction to Psychoacoustics*. Cambridge: MIT Press.

Smalley, D. 1986. "Spectro-morphology and Structuring Processes." In S. Emmerson, ed., *The Language of Electroacoustic Music*. London: Macmillan.

Smalley, D. 1991. "Spatial Experience in Electro-Acoustic Music." In *L'Espace du Son II*. Ohain: Musiques et Recherches.

Smalley, D. 1996. "The Listening Imagination: Listening in the Electroacoustic Era." In *Contemporary Music Review*, 13:2. Amsterdam: Harwood.

Smalley, D. 1997. "Spectromorphology: Explaining Sound Shapes." In *Organised Sound*, 2:2. Cambridge: Cambridge University Press.

Stockhausen, K. 1961. "Music in Space." In *Die Reihe*, 5. New Jersey: Universal Edition.

Tenney, J. 1980. "Temporal Gestalt Perception in Music." In *Journal of Music Theory*, 24:2.

Tenney, J. 1984. "John Cage and the Theory of Harmony." In *Soundings*, 13. Santa Fe: Soundings Press.

Tenney, J. 1992. *META-HODOS and META Meta-Hodos*. Hanover, NH: Frog Peak Music.

Toole, F. E. 1998. "Subjective Evaluation." In J. Borwick, ed., *Loudspeaker and Headphone Handbook*. Oxford: Focal Press.

Truax, B. 1992. "Electroacoustic Music and the Soundscape: The Inner and Outer World." In Paynter, Howell, Orton, and Seymour, eds., *Companion to Contemporary Musical Thought*. London: Routledge.

Truax, B. 1994. "Discovering Inner Complexity: Time Shifting and Transposition with a Real-time Granulation Technique." In *Computer Music Journal*, 18:2. Cambridge, Massachusetts: MIT Press.

Truax, B. 1996. "Soundscape, Acoustic Communication and Environmental Sound Composition." In *Contemporary Music Review*, 15:1. Amsterdam: Harwood.

Truax, B. 1998. "Composition and Diffusion: Space in Sound in Space." In *Organised Sound*, 3:2. Cambridge: Cambridge University Press.

Warren, D. H. 1970. "Intermodality Interactions in Spatial Localisation." In *Cognitive Psychology*, 1. Cited in O'Connor and Hermelin (1978), p. 75.

Watkinson, J. 1998. *The Art of Sound Reproduction.* Oxford: Focal Press.

Watkinson, J. 2001. *The Art of Digital Audio.* Oxford: Focal Press.

Weitz, S., ed. 1979. *Nonverbal Communication: Readings with Commentary.* New York: Oxford University Press.

White, F. A. 1975. *Our Acoustic Environment.* New York: Wiley & Sons.

White, J. D. 1984. *The Analysis of Music.* Metuchen and London: Scarecrow Press.

Winckel, F. 1967. *Music, Sound and Sensation: A Modern Exposition.* New York: Dover.

Wishart, T. 1986. "Sound Symbols and Landscapes." In S. Emmerson, ed., *The Language of Electroacoustic Music.* London: Macmillan.

Wishart, T. 1988. "The Composition of *Vox-5*." In *Computer Music Journal*, 12:4. Cambridge, Massachusetts: MIT Press.

Wishart, T. 1996. *On Sonic Art.* Amsterdam: Harwood.

Woodgate, J. M. 1998. "International Standards." In J. Borwick, ed., *Loudspeaker and Headphone Handbook.* Oxford: Focal Press.

Worrall, D. 1999. "Space in Sound: Sound of Space." In *Organised Sound*, 3:2. Cambridge: Cambridge University Press.

Young, J. 1996. "Imagining the Source: The Interplay of Realism and Abstraction in Electroacoustic Music." In *Contemporary Music Review*, 15:1. Amsterdam: Harwood.

Appendix I
Programme notes

(dis)integration (1998)

dis'in'te'grate (dɪs'ɪntɪgreɪt) *verb*
dis'in'te'grated, dis'in'te'grat'ing, dis'in'te'grates *verb, intransitive*

1. To become reduced to components, fragments or particles.
2. To decompose, decay, or undergo a nuclear transformation.

in'te'grate ('ɪntɪgreɪt) *verb*
inte'grat'ed, inte'grat'ing, inte'grates *verb, transitive*

1. To make into a whole by bringing all parts together; unify.
2. **a.** To join with something else; unite. **b.** To make part of a larger unit.

(dis)integration was composed in the composer's private studio and in the studios of the City University Music Department in London. It earned a Residence Award in the 1998 Bourges competition.

Ebb (1998)

Ebb was created in London in 1998 during the grey and rainy summer that never became a real summer. Around the same time there were record breaking temperatures in the USA and disastrous tidal waves in South America. The strange weather and the unpredictable seasons in recent years and possible reasons for this were the inspiration for the piece. The title, *Ebb*, does not only refer to the water sounds that the work is based on, but implies other meanings; political, social and psychological. The work deals with regression and decline on several levels, but this is not necessarily obvious on the surface. Ambiguity is created by contrasting beautiful water sounds and drones with distorted sound masses and downward glissandos. There are familiar sounds intermixed with unknown sonic textures. There are natural sounds set up against sounds of human gesture. These are all integrated in a structure that balances on the edge of a narrative. *Ebb* is a journey that does not really end, but continues into the far distance—with or without the consciousness of its passengers.

This composition is dedicated to those who fight against the exploitation and degradation of the world's natural resources.

Ebb was composed in the composer's private studio and in the studios of the City University Music Department in London.

Intra (1999)

The fear of the unfamiliar and the threat of having to challenge one's own values and ideas are nowhere stronger than when confronted with the abnormal mind. The definition and threshold of madness differs from culture to culture and changes with the times. Numerous mythological and religious writings throughout history give proof of this mutable definition of deviant behaviour. The mysterious and incomprehensible sides of the human psyche have been a source of inspiration and fascination for artists in all times—many of whom themselves have been labelled mad.

The underlying model for this composition is based on different stages of schizophrenia, a condition characterised by a fragmentation of mental functions and a blurring of fantasy and reality. Difficulty in distinguishing between internal and external worlds is a key symptom, and hallucinations and exaggerations of sensory input are common. Space and spatial relations among the sounds are central tools in this composition to portray such a state. Representations of real and surreal sonic environments appear both separately and juxtaposed throughout the work. Transitions between different virtual spaces are guided by the development of the sound material, and take place sometimes gradually, other times abruptly. A wide range in dynamics and density of sounds reflect extremes in sensations experienced from overstimulation and desperation for calm and stability. As with the state of the mind, one cannot always anticipate what happens next...

Intra was composed in the composer's private studio and in the studios of the City University Music Department in London. It received an Honorary Mention in the 2000 Métamorphoses competition.

Itinera (1999-2001)

Itinera is, as the title suggests, a piece about travelling. I collected the source sounds while waiting in airports and train stations around Europe and the USA, and the idea for the composition came when dipping in and out of the trance-like state of mind that often accompanies travel. The work seeks to portray such a state by shifting between concrete, unprocessed sounds of trains and traffic, clearly suggesting an outdoor "reality" setting, and more abstract, processed sound material which is less noisy and indicates an indoor, or internal mental, environment. These shifts in perspective, as well as the overall structure of the work, attempt to depict excitement, expectation, exhaustion, waiting, boredom, jet-lag and other attributes of long distance travel. The work is in two parts titled *Trans/Trance* and *kernelMotion*, respectively.

Itinera was composed in the composer's private studio and in the studios of the City University Music Department in London.

Terra Incognita (2001)

The unfamiliar land, *Terra Incognita,* wherever it may be with its promises of riches and happiness, freshness and opportunities, have throughout the ages attracted adventurers, conquerors and exiles alike. In our time, the need to establish a new home due to forced exile is an equally common reason for moving to a new world as the adventurer's dream of starting afresh with a new hope of wealth and happiness. For the most audacious conquerors, other planets and galaxies have become an increasingly appealing target for exploration. In the fantasy literature, outer space and the depths of the ocean have been settings of choice for human expeditions for a long time. Here, imagination has no restraints as to what the journey brings. *Terra Incognita* is a journey to a sonic landscape. It provides a setting which may evoke images of something real or fantastic that borders the narrative. The work is composed in ambisonics in order to create a three-dimensional sound-sphere in which sounds can move about in any direction, and where full surround-sound environments can be set up. The wide dynamics and the variations in range and speed of movement of the different sound materials in this piece indicate ever-changing situations, sometimes expected, other times not.

Terra Incognita was composed in the composer's private studio and in the studios of the City University Music Department in London.

Appendix II
Programme notes for *Lorry Red Lorry Yellow*

Lorry Red Lorry Yellow (2000)

> *There is no question that in our age, in which we are inundated with sound, is unprecedented in history. We need to develop a more caring attitude towards sounds. Presently, the amounts of sound and music in the environment have clearly exceeded man's capacity to assimilate them, and the audio ecosystem is beginning to fall apart. Background music, which is supposed to create atmosphere, is far too excessive. In our present condition we find that within certain areas and spaces aspects of visual design are well attended to, but sound design is completely ignored. It is necessary to treat sound and music with the same respect that we show for architecture, interior design, food, or the air that we breath.*

(Yoshimursa, 1982)

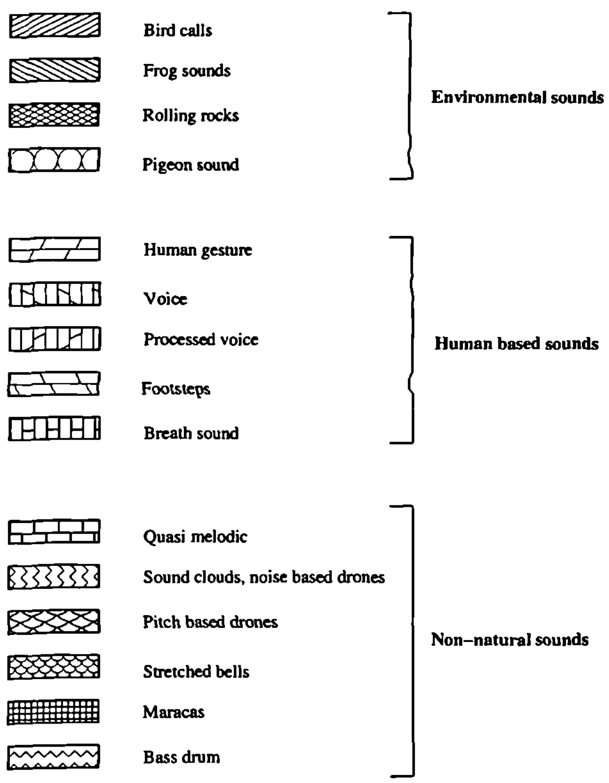
This installation is created in the spirit of Yoshimursa's imperative. Sound is all too often ignored as an essential element in defining space. As silence is increasingly difficult to come by, sound is virtually a constant in our lives. Yet we only really notice sound when it is extremely offensive (a really loud jackhammer) or pumped out of stereos in a familiar and slick package ("oh, this is my favourite song"); everything else is background. But what is background sound? What are we hearing and how does this constant passive semi-conscious listening affect us?

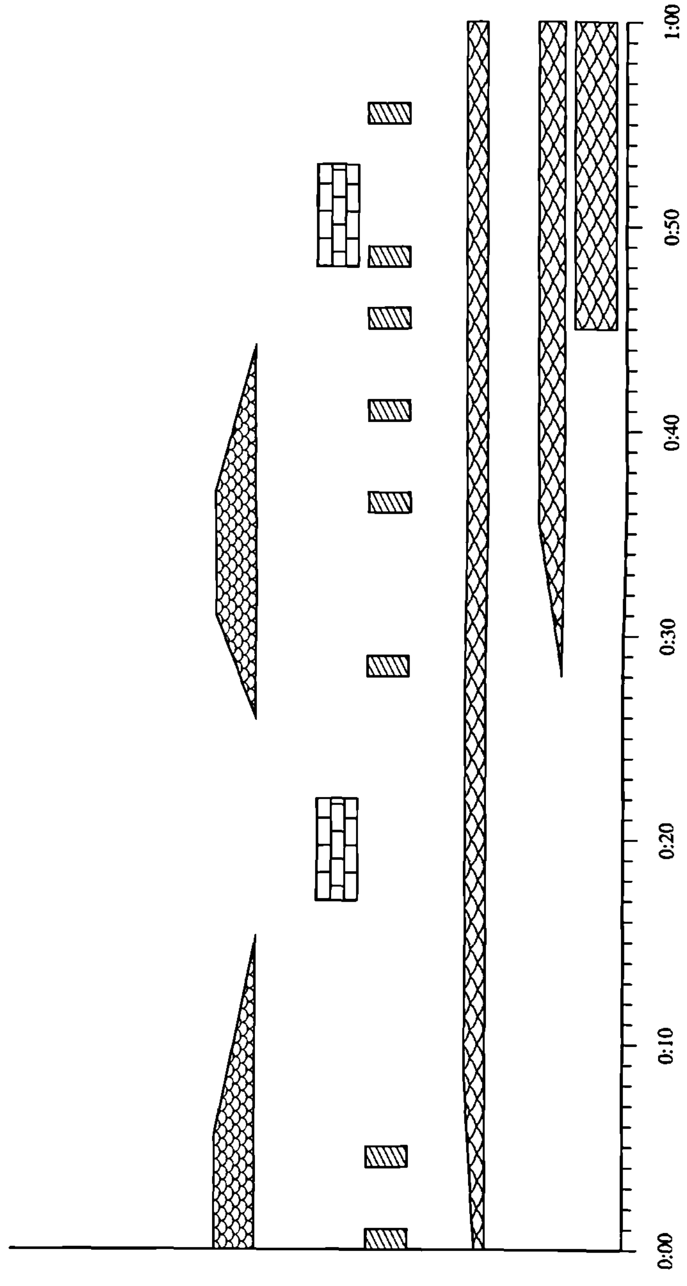
A current Sony advertisement campaign suggests that people use their Walkman to block out the rest of the world. Some of the advertisements assert that playing one's Walkman is a good way to avoid life, while others contradictorily propose that life is listening to pre-recorded commercialised sound—the overriding suggestion of this ad campaign being that one can create a world for oneself by using sound to block out the din of life. This approach is reminiscent of Eighteenth Century personal hygiene where, instead of bathing and showering, people applied thick creams and strong smelling perfumes to mask their body odours.

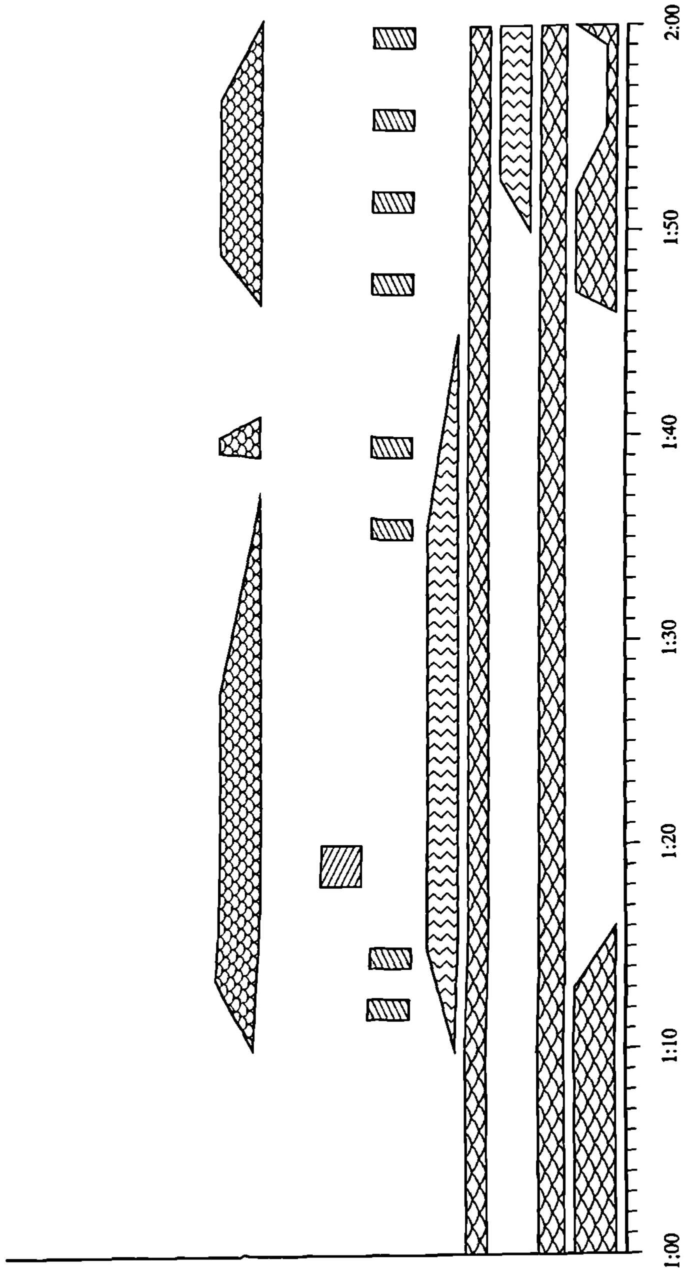
Lorry Red Lorry Yellow consists of a virtually infinite number of permutations of sounds taken from both outside and inside the exhibition space. These ubiquitous yet ignored sounds are taken out of context and put into a gallery where they become highlighted versions of themselves. They are sounds we hear every day, part of the fabric of the background, sounds which we have all learned to block out. Presented in a gallery setting, will they command attention? Out of context, do they become important? In *Lorry Red Lorry Yellow*, we put sound up for consideration. This piece is a nod to sound as an integral part of the formula which defines our quality of life. Above all else, this piece is an invitation to the audience to listen to each sound in the din of life.

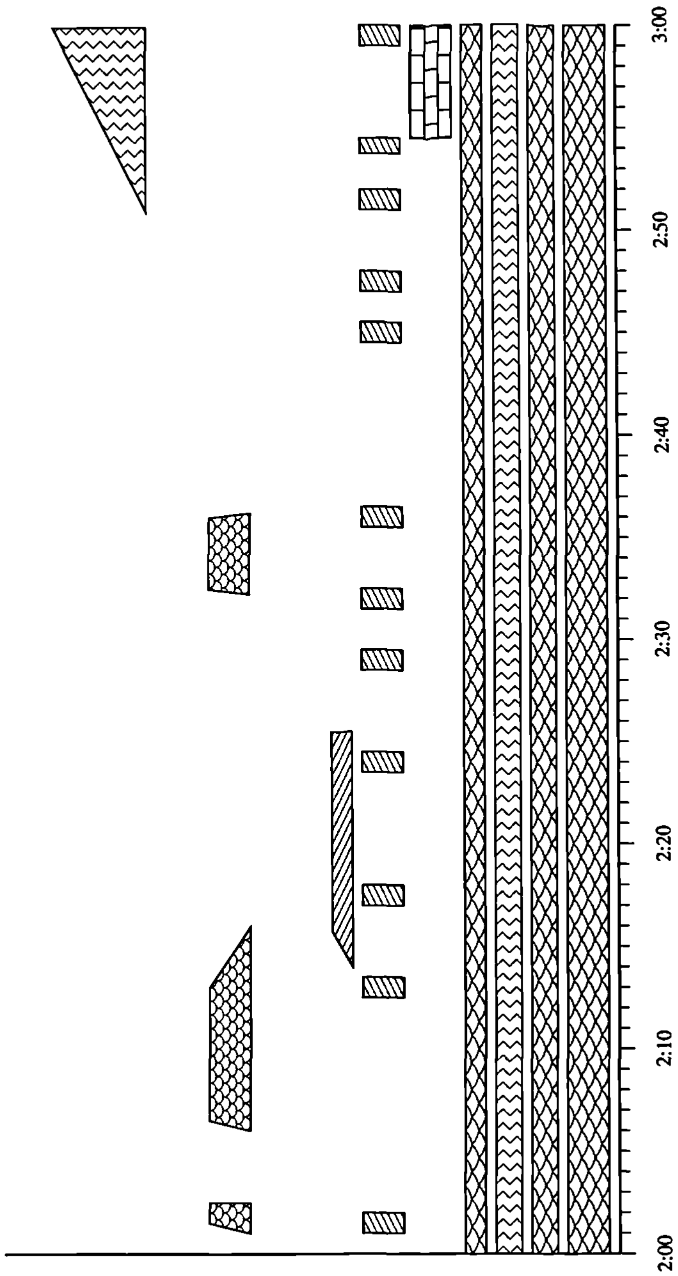
Appendix III
Spatio-structural analysis of *Intra*

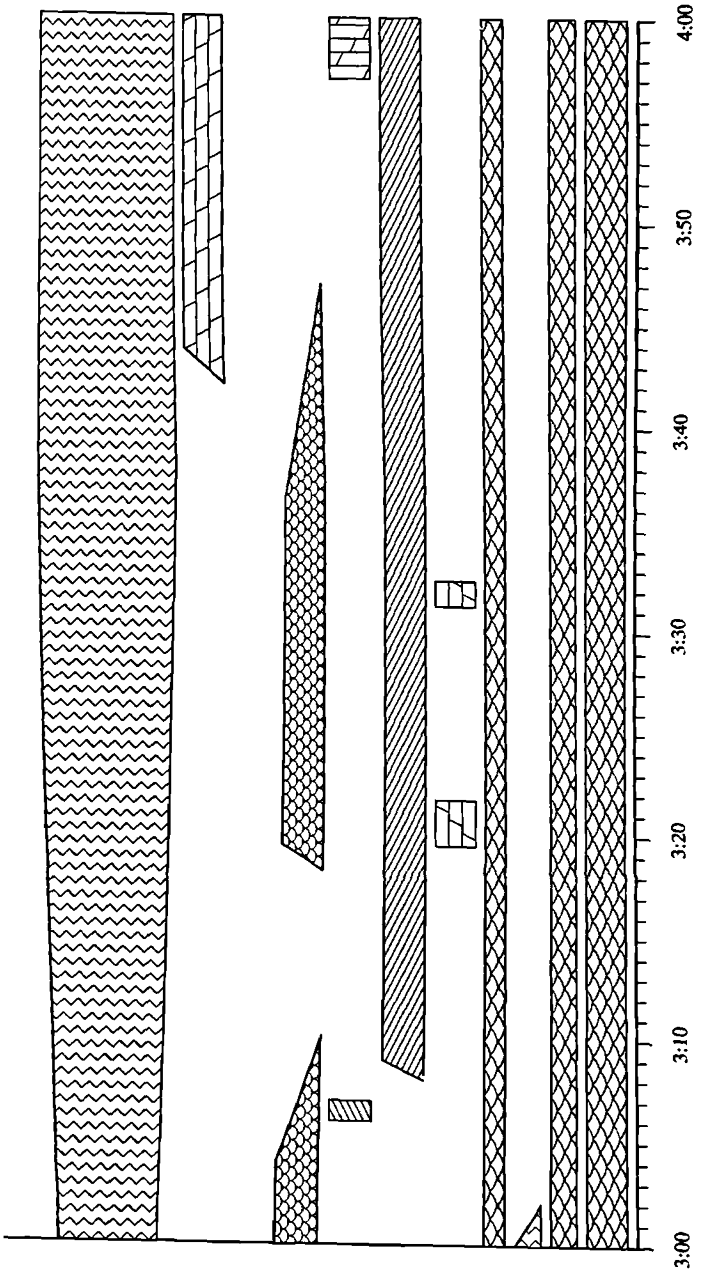
For the analysis diagram I have summarised the sound material used in *Intra* into three main categories: environmental sounds, human based sounds and non-natural sounds. Below is a map of the graphical representation of these sound types as they are displayed in the analysis diagram on the following pages.

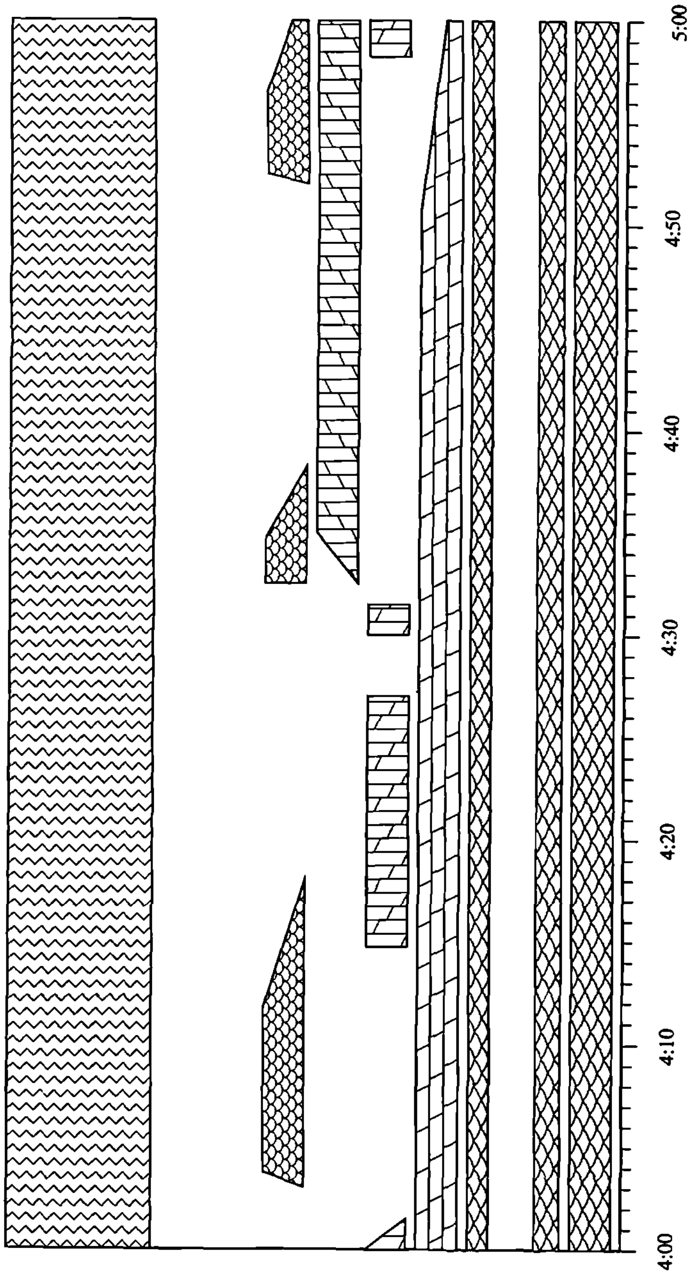


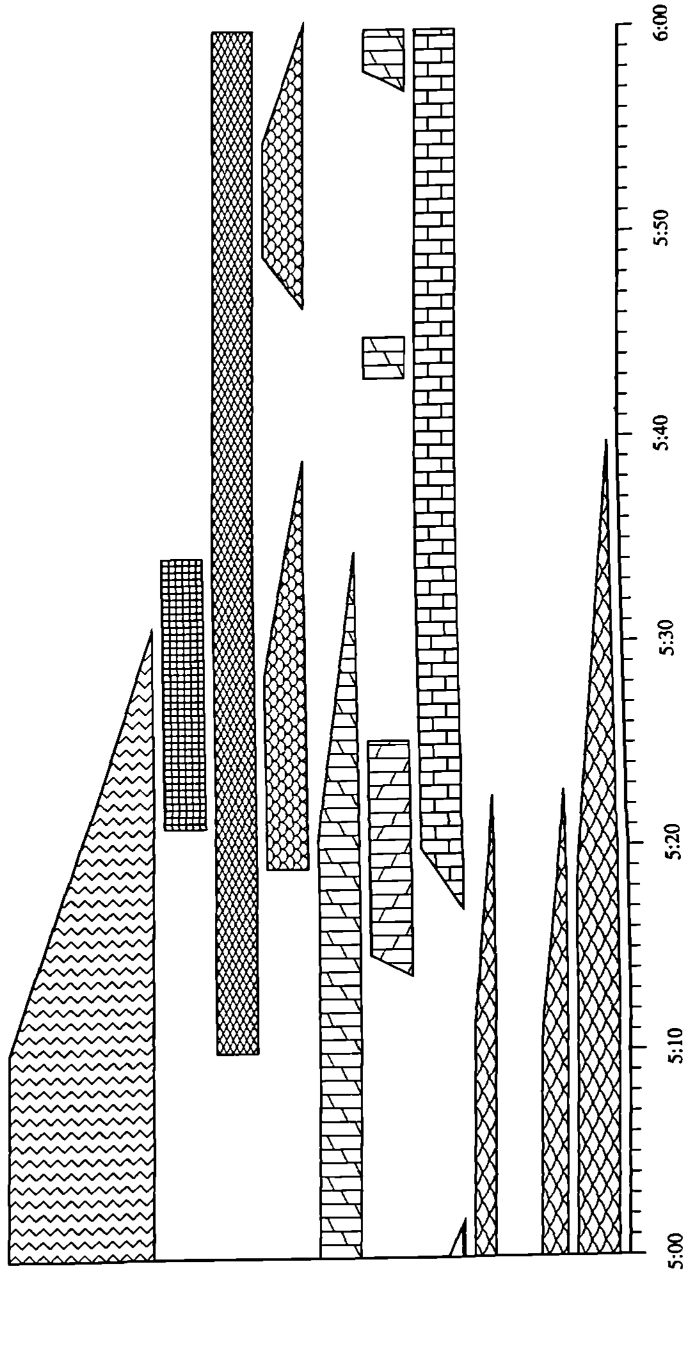

0:00
0:10
0:20
0:30
0:40
0:50
1:00

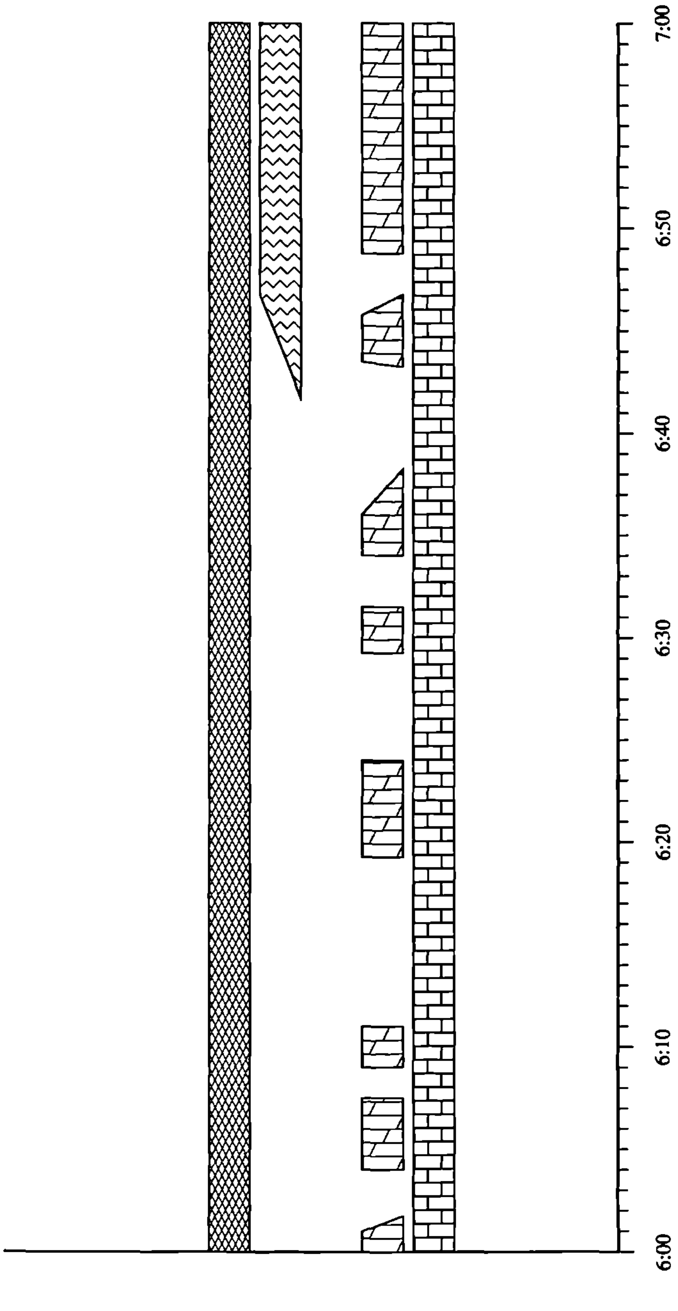

1:00
1:10
1:20
1:30
1:40
1:50
2:00

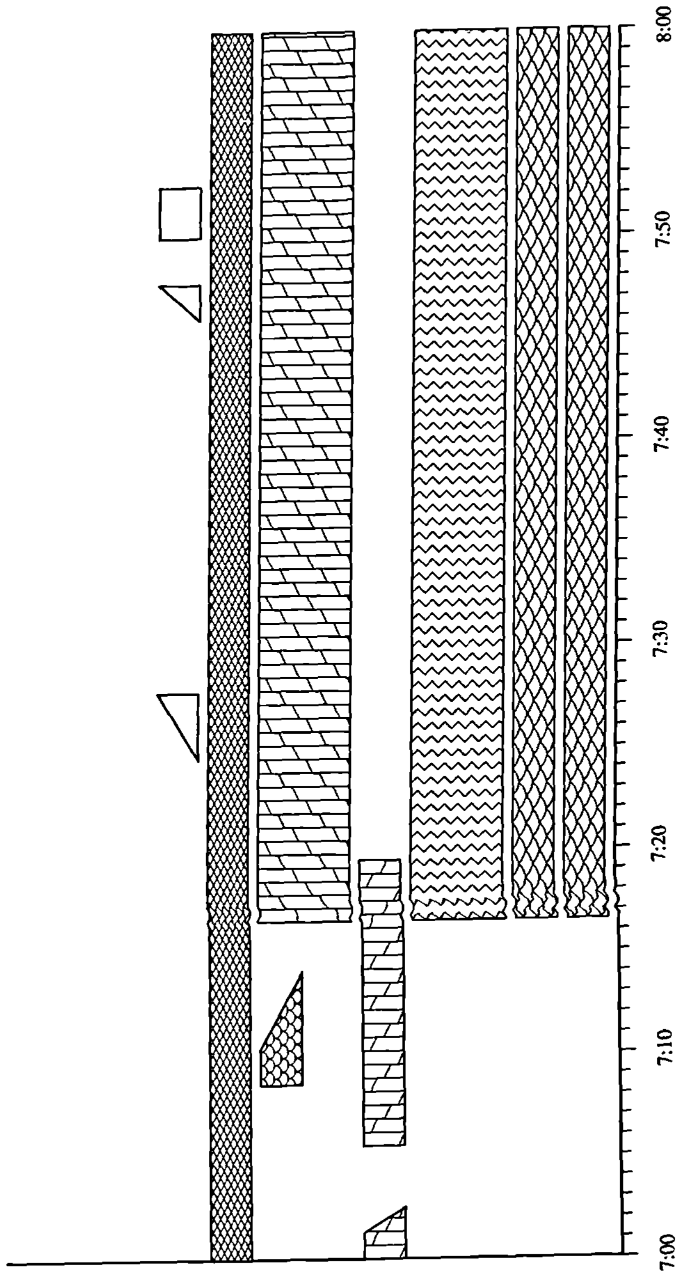

3:00
2:50
2:40
2:30
2:20
2:10
2:00

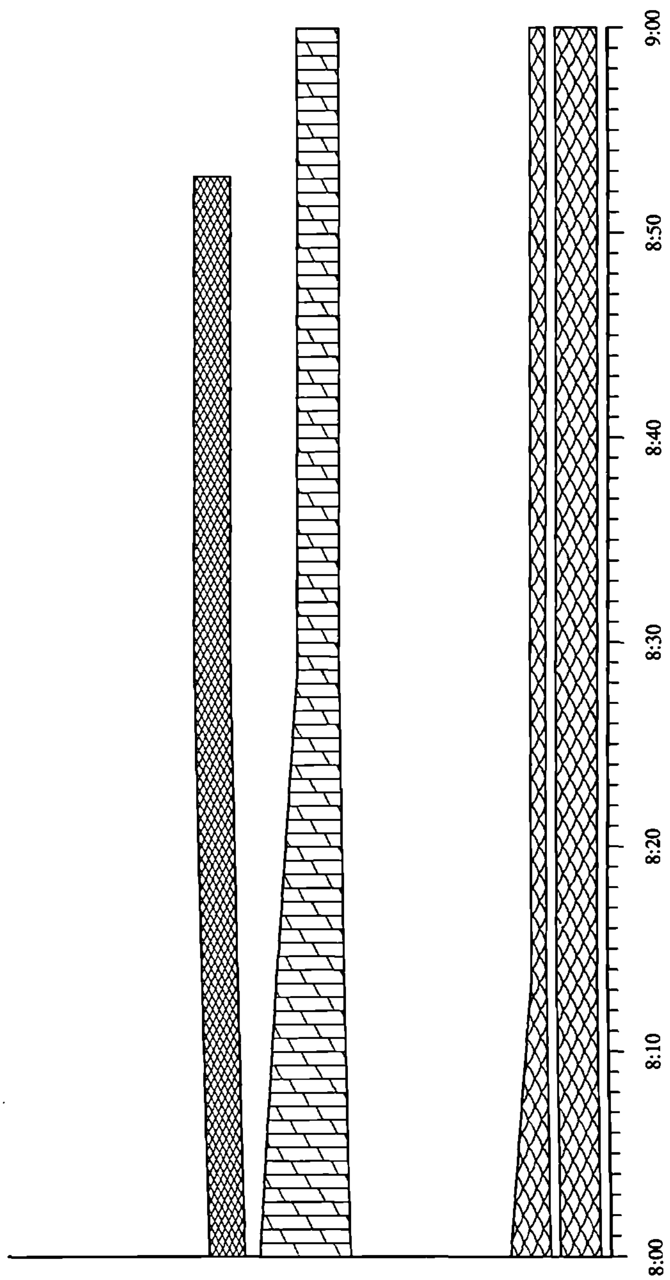

3:00
3:10
3:20
3:30
3:40
3:50
4:00

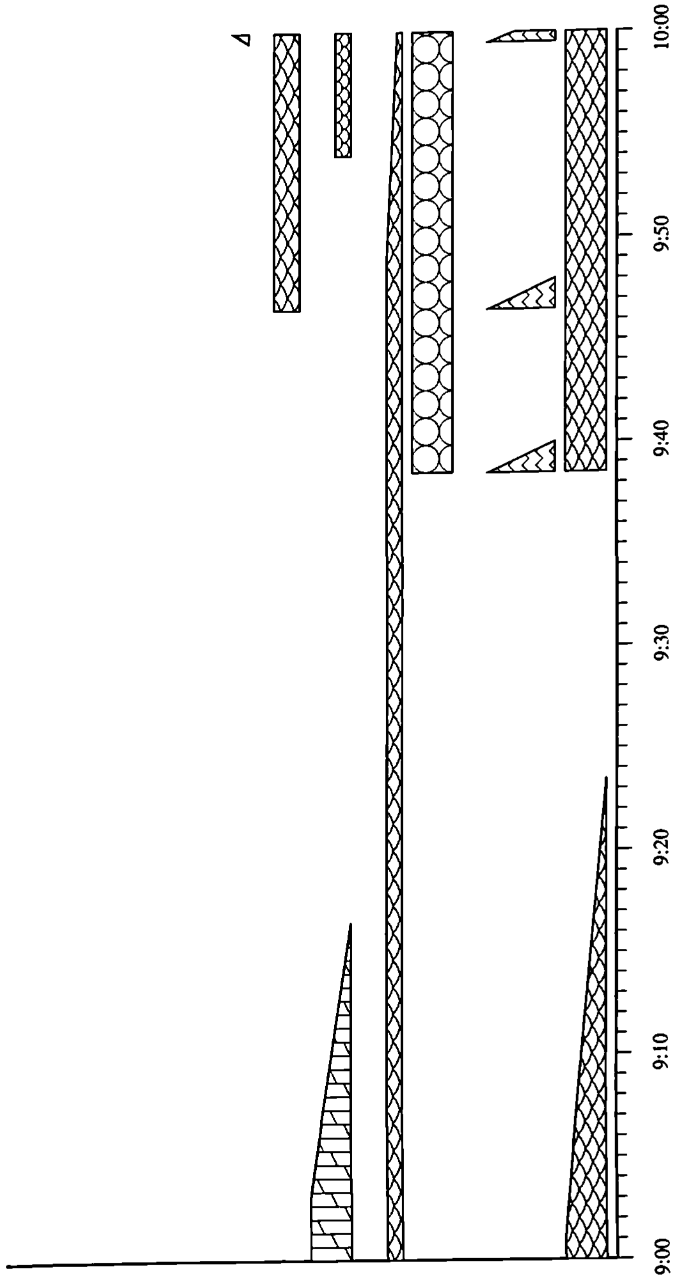

4:00
4:10
4:20
4:30
4:40
4:50
5:00

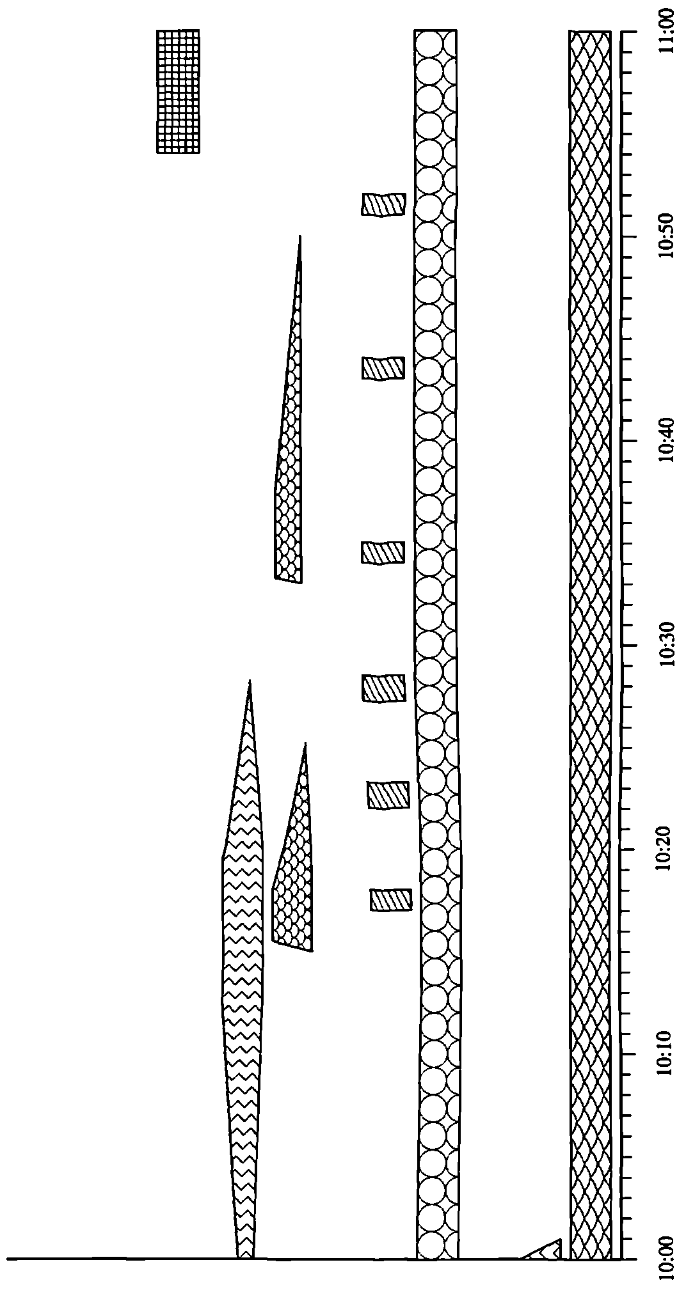

5:00
5:10
5:20
5:30
5:40
5:50
6:00

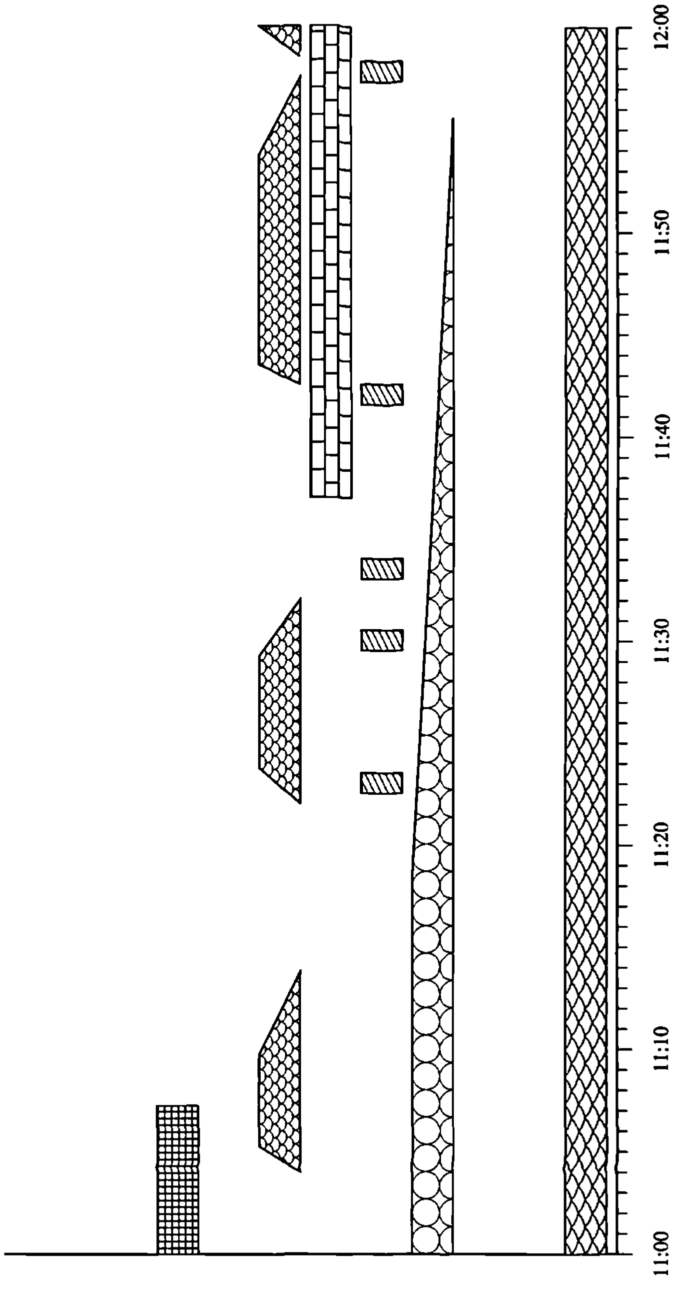

7:00
6:50
6:40
6:30
6:20
6:10
6:00

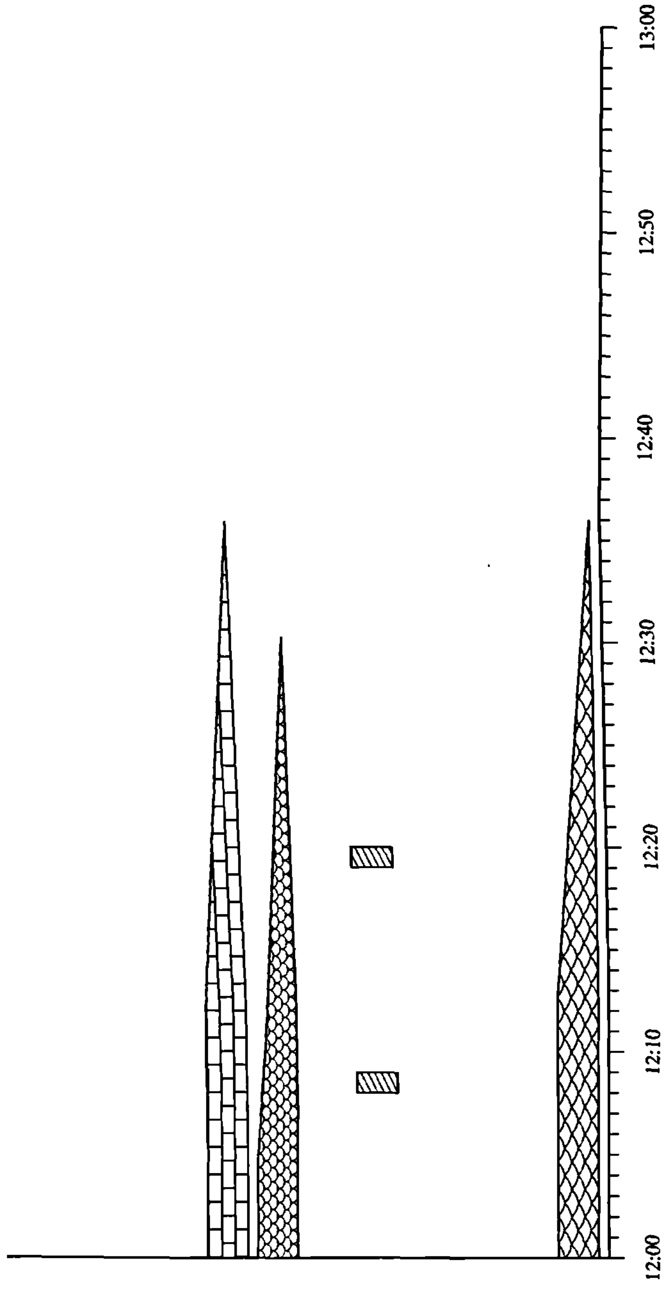

7:00
7:10
7:20
7:30
7:40
7:50
8:00


8:00
8:10
8:20
8:30
8:40
8:50
9:00


9:00
9:10
9:20
9:30
9:40
9:50
10:00


10:00
10:10
10:20
10:30
10:40
10:50
11:00

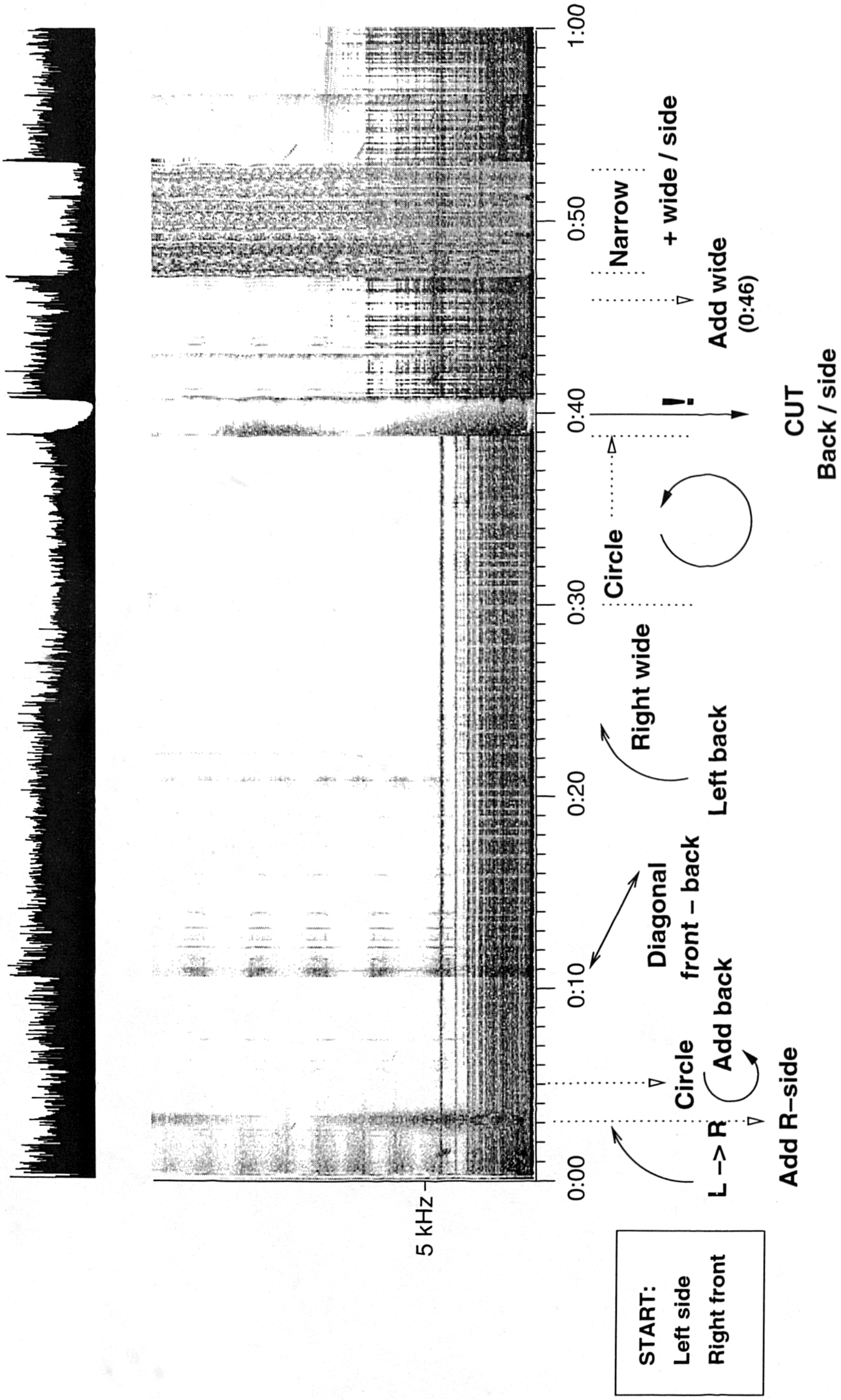

11:00
11:10
11:20
11:30
11:40
11:50
12:00

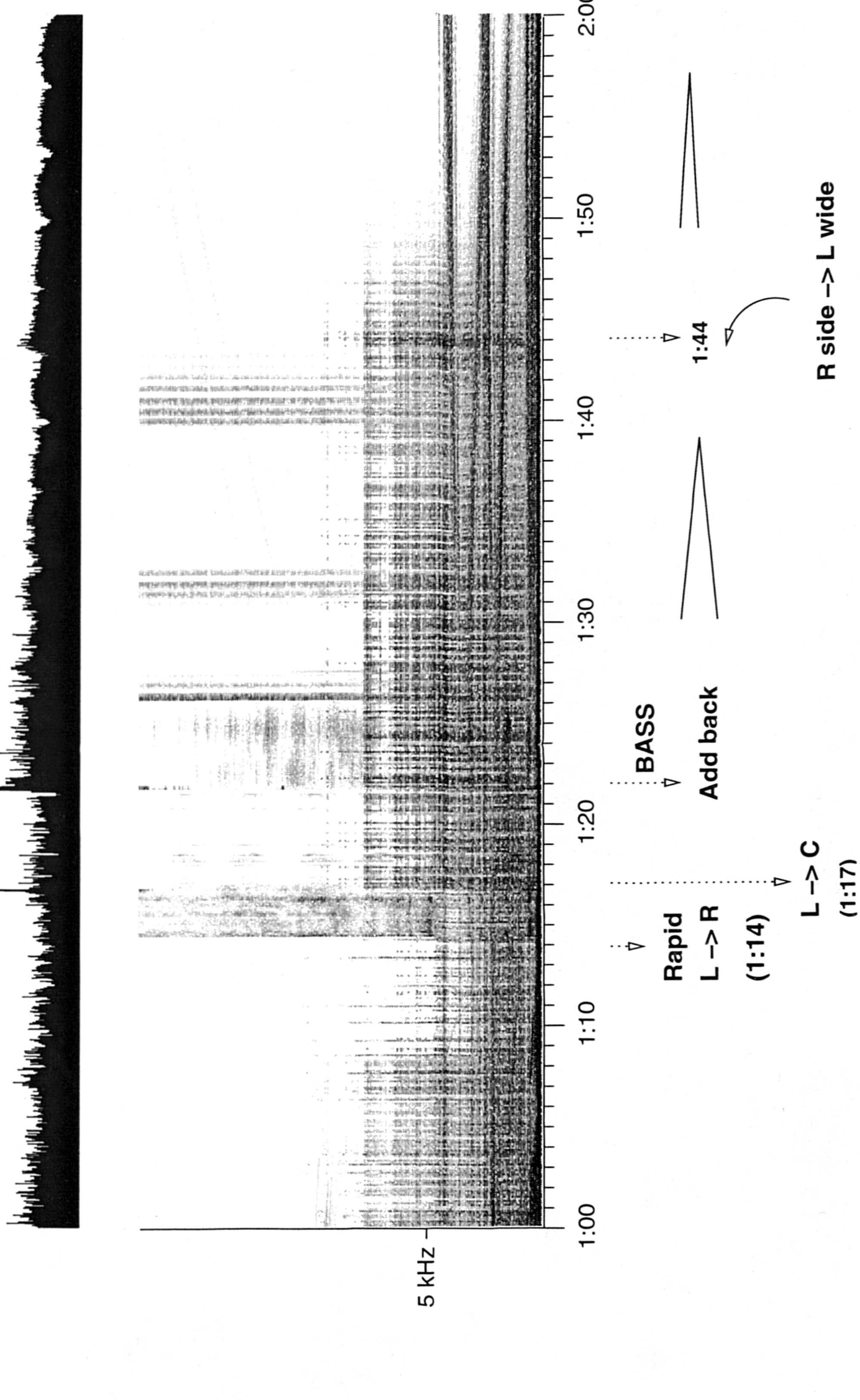

12:00
12:10
12:20
12:30
12:40
12:50
13:00

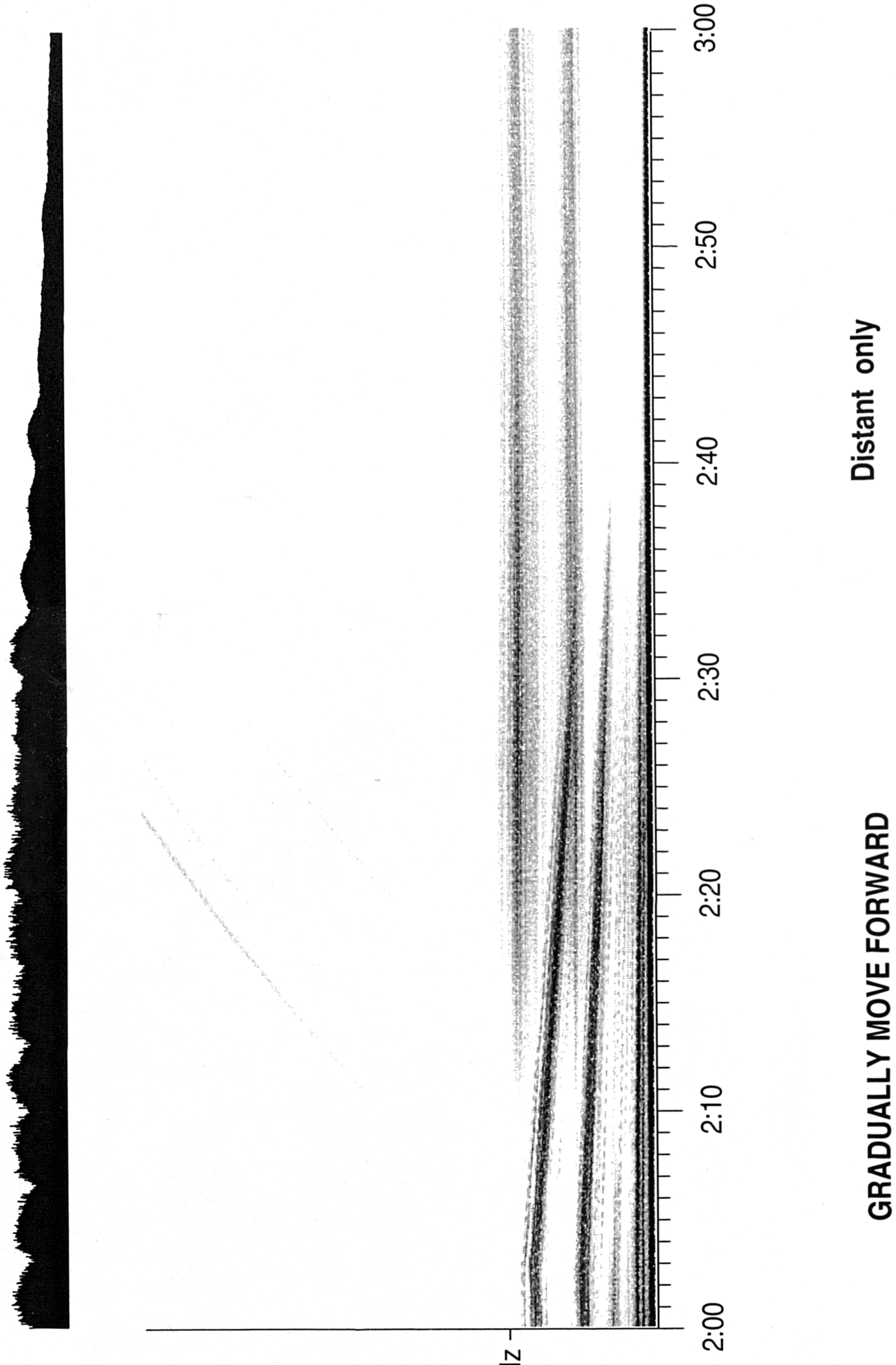
Appendix IV
Diffusion score for *(dis)integration*

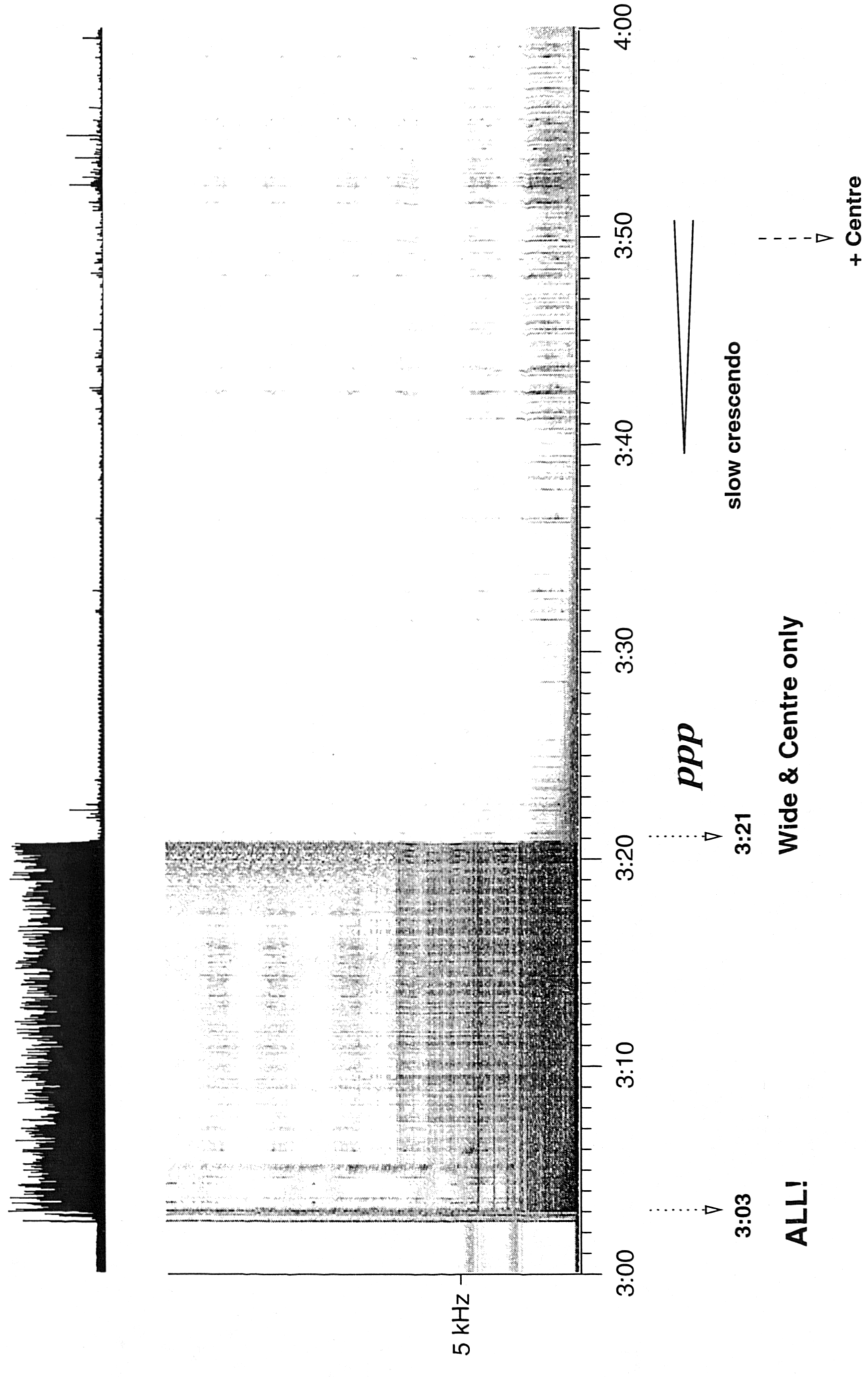
Below is a diffusion score for my work *(dis)integration*. The sonogram is generated by Øyvind Hammer's *Sono*[3], a computer programme running on SGI and Linux computers.

[3]*Sono* is open source software that is freely available from the NoTAM web server at `http://www.notam02.no/notam02/prod-prg-sono.html`..

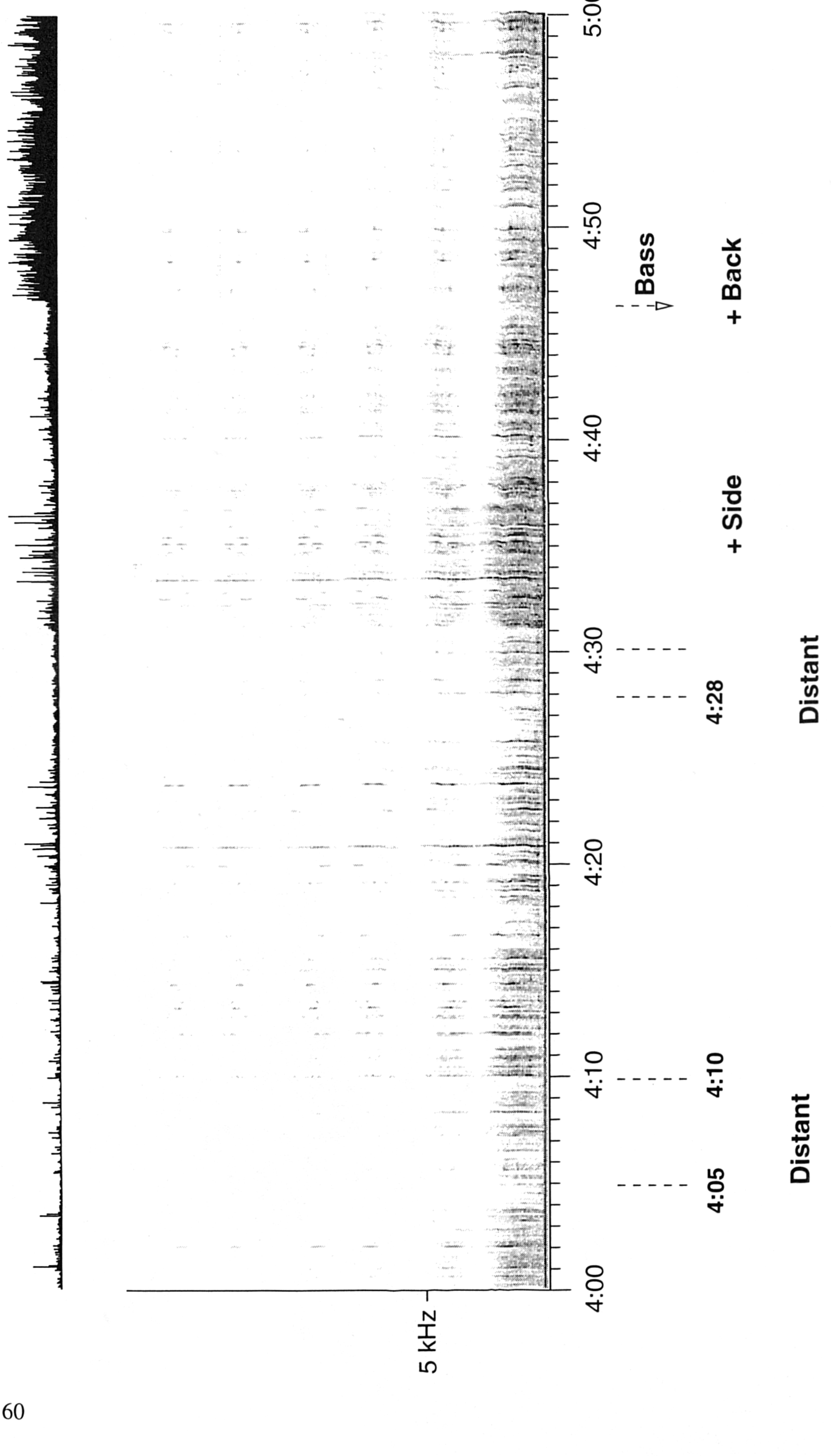

0:00
0:10
0:20
0:30
0:40
0:50
1:00
5 kHz
START:
Left side
Right front
L -> R
Add R–side
Circle
Add back
Diagonal
front – back
Left back
Right wide
Circle
CUT
Back / side
Add wide
(0:46)
Narrow
+ wide / side

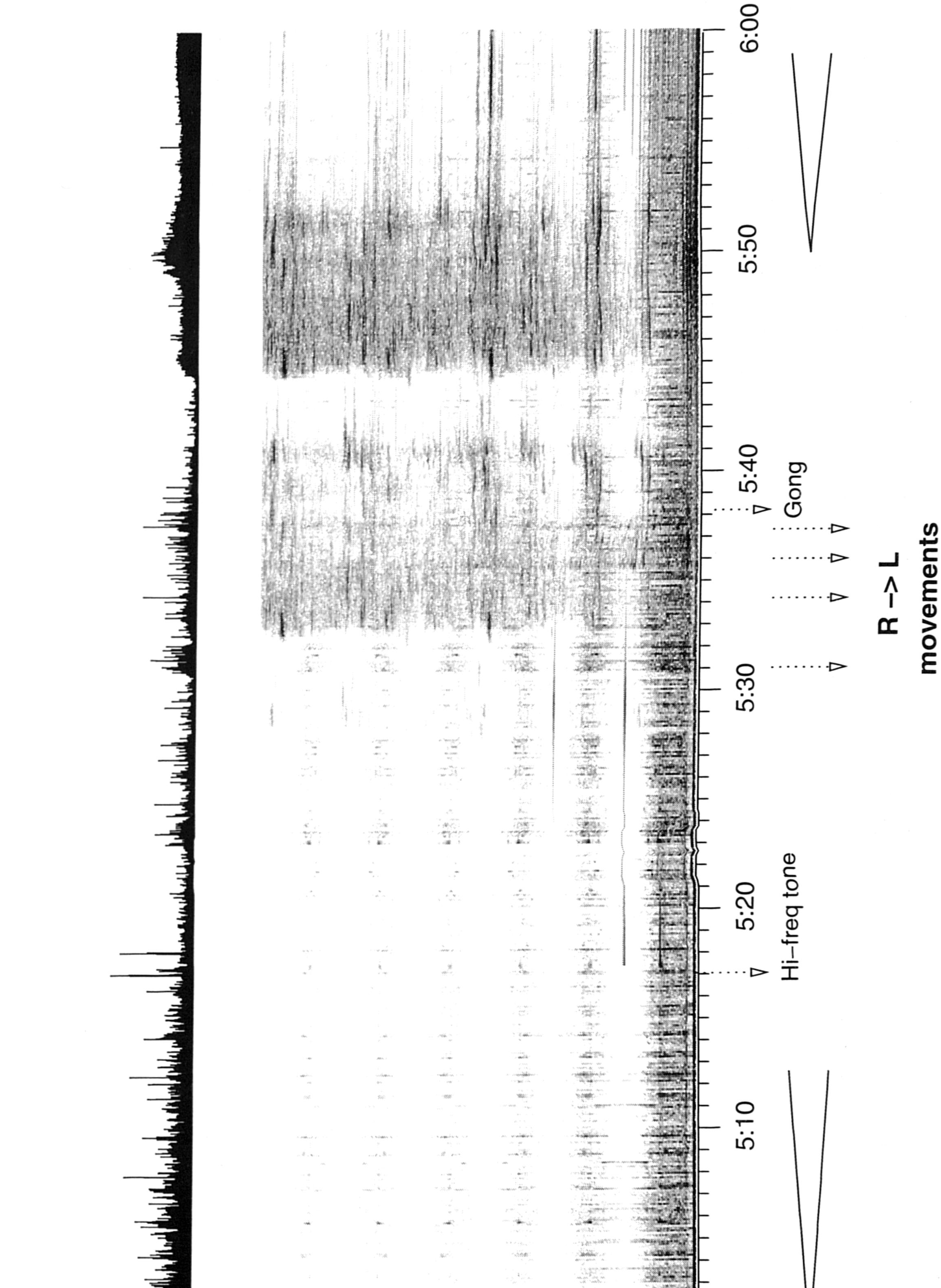

5 kHz
1:00
1:10
1:20
1:30
1:40
1:50
2:00
Rapid
L -> R
(1:14)
L -> C
(1:17)
BASS
Add back
1:44
R side -> L wide

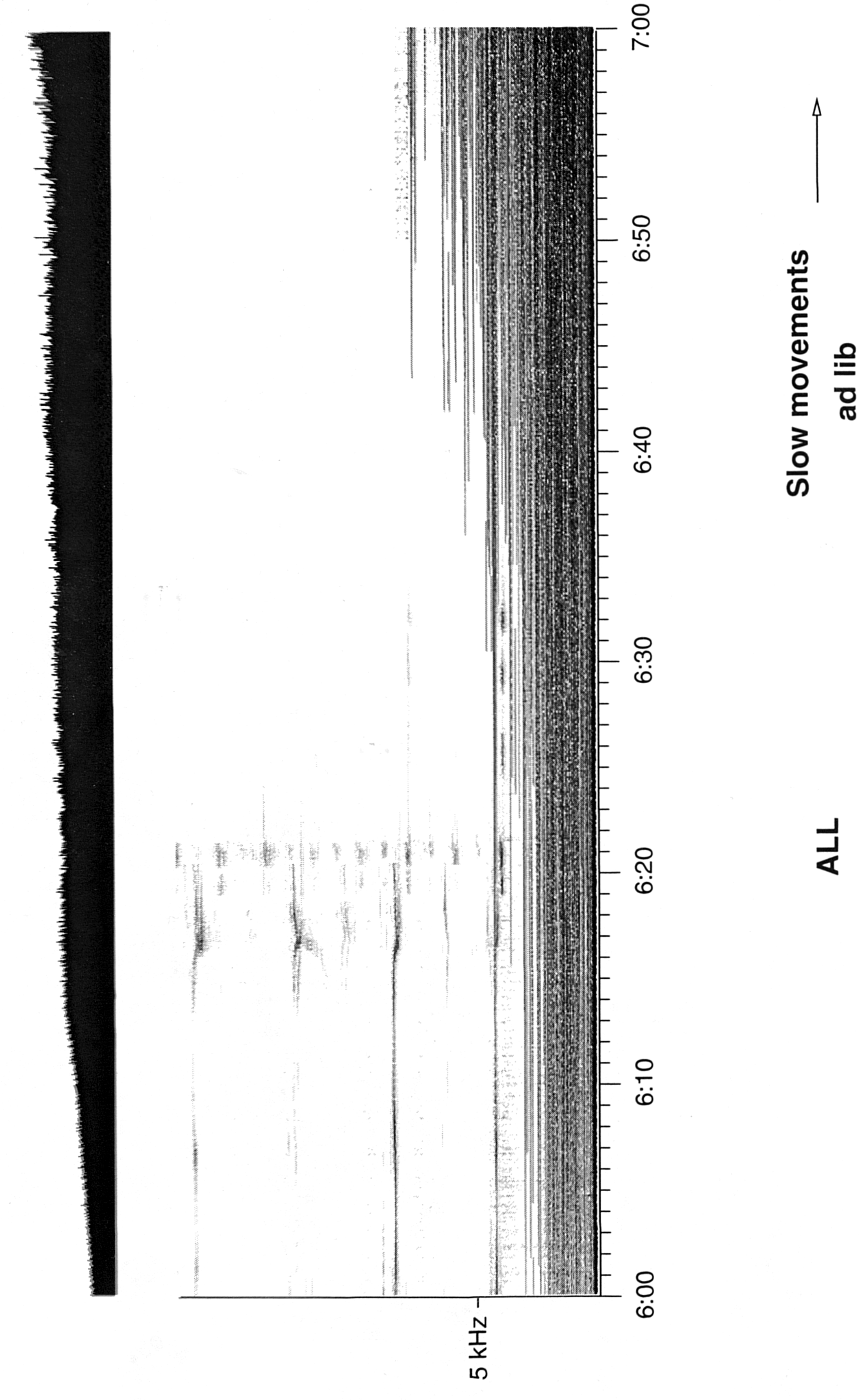

5 kHz
2:00
2:10
2:20
2:30
2:40
2:50
3:00
GRADUALLY MOVE FORWARD
Distant only

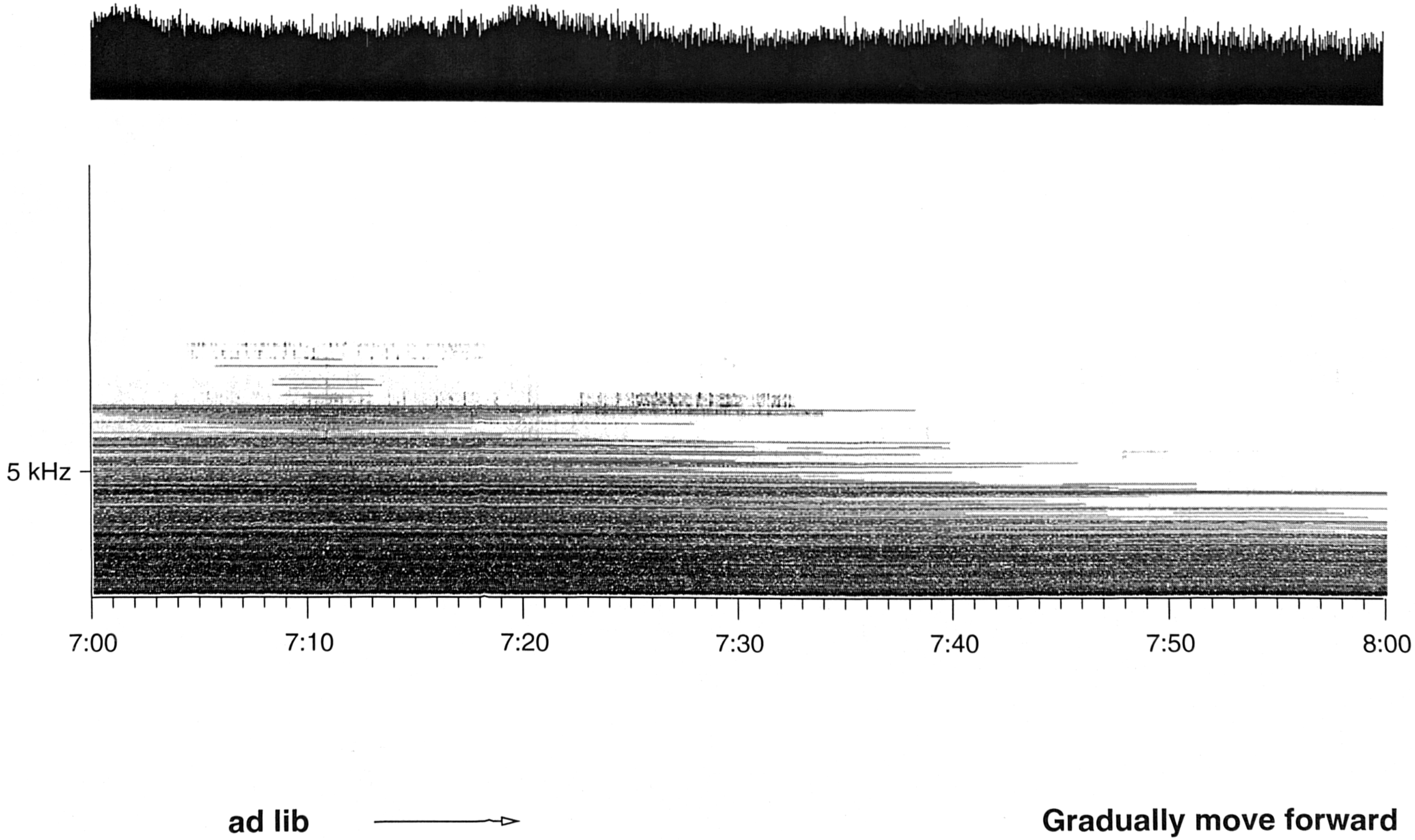


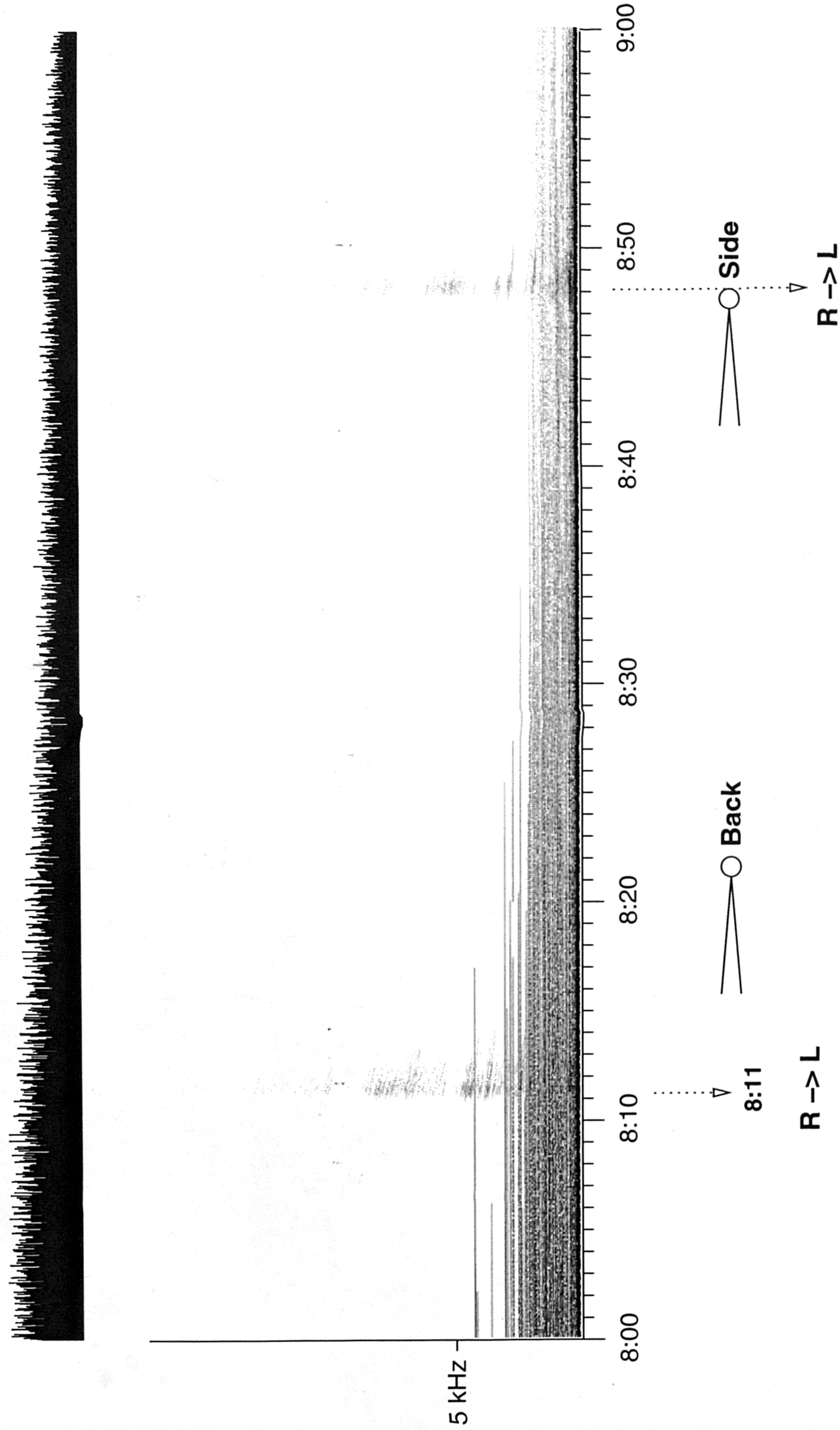
5 kHz
3:00
3:10
3:20
3:30
3:40
3:50
4:00
3:03
ALL!
3:21
ppp
Wide & Centre only
slow crescendo
+ Centre


5:00
4:50
4:40
4:30
4:20
4:10
4:00
5 kHz
Bass
+ Back
+ Side
4:28
Distant
4:10
4:05
Distant

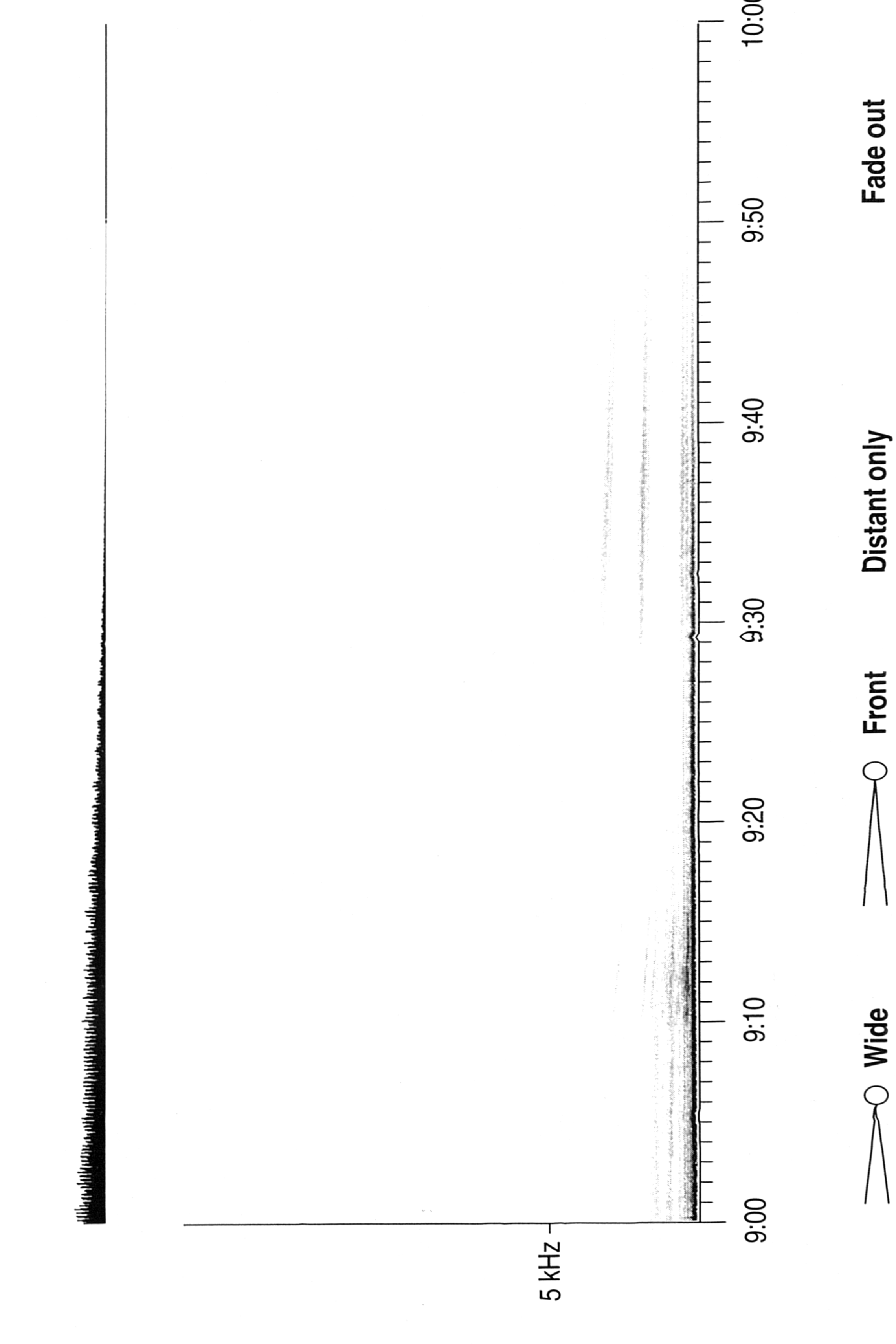

5 kHz
5:00
5:10
5:20
5:30
5:40
5:50
6:00
Hi–freq tone
Gong
R –> L
movements


5 kHz
6:00
6:10
6:20
6:30
6:40
6:50
7:00
ALL
Slow movements
ad lib


5 kHz
7:00
7:10
7:20
7:30
7:40
7:50
8:00
ad lib
Gradually move forward



9:00
8:50
8:40
8:30
8:20
8:10
8:00
5 kHz
Side
R -> L
Back
8:11
R -> L


5 kHz
9:00
9:10
9:20
9:30
9:40
9:50
10:00
Wide
Front
Distant only
Fade out