The University of Sheffield

James R. Mooney

Sound Diffusion Systems for the Live Performance of Electroacoustic Music

An Inclusive Approach led by Technological and Aesthetical Consideration of the Electroacoustic Idiom and an Evaluation of Existing Systems

Volume 1

Introduction and Chapters 1 – 3

Submitted for the Degree of Ph.D. in the Faculty of Arts

Department of Music

August 2005

THIS THESIS CONTAINS A MUSIC CD

-UNABLE TO COPY-

PLEASE CONTACT THE UNIVERSITY IF YOU WISH TO SEE THIS MATERIAL

Abstract

This thesis documents research in the field of sound diffusion for the live performance of electroacoustic music. Broad and inclusive ways of conceptualising electroacoustic music are presented, with the intention of promoting the design of improved sound diffusion systems in the future.

Having defined 'electroacoustic music' in terms of the technologies involved and the unique ways in which these *creative frameworks* are appropriated by practitioners (Chapter 1), a binary interpretation of the electroacoustic idiom, whereby musical philosophies can be regarded as either *top-down* or *bottom-up*, is given (Chapter 2). Discussion of the process of sound diffusion itself reveals two distinct performance praxes, which can also be characterised as top-down and bottom-up (Chapter 3). These differing ideologies, in addition to the technical demands of the electroacoustic idiom and the logistical demands of sound diffusion itself, must be accommodated by the sound diffusion system if live performances are to achieve the desired musical communication. It is argued that this is not presently the case.

A system of criteria for the evaluation of sound diffusion systems is presented (Chapter 4). Two original concepts – the *coherent audio source set* (CASS) and *coherent loudspeaker set* (CLS) – are also presented; these are intended to be practically and theoretically useful in the field of sound diffusion. Several existing diffusion systems are evaluated in terms of these criteria (also Chapter 4). A description and evaluation of the M2 Sound Diffusion System, which was co-developed by the author as part of this research, is also given (Chapter 5).

The final chapter describes ways in which superior future systems can be devised. These range from specific practical suggestions to general methodological recommendations. Overall, the intention is to provide an interpretation of the electroacoustic idiom that can be used as a heuristic tool in the design of new sound diffusion systems.

Acknowledgements

In no specific order, I would like to thank:

My family, for their unwavering support throughout what has, in many respects, been a difficult and challenging four years;

My friend and colleague Dr. David Moore, for collaborative work, guidance, and general advice and inspiration;

Nikos Stavropoulos, Mark Horrocks, Menia Pavlakou, Carla Ribeiro, and Themis Toumpoulidis, who are both friends and fellow Ph.D. students;

Dr. Colin Roth, whose guidance has been extremely significant, and whose writings[1] and theories have been nothing short of inspirational;

My supervisor, Dr. Adrian Moore, for giving me the opportunity to carry out this research, and all of the other staff at the Music Department;

Professor Jonty Harrison, for his enthusiasm, support, and very kindly donated opinions (see Appendix 1);

Tom, Cathy, Rachel, Katie, John, Becky, Simon: sorry to lump you all together, but I had to limit myself to one page!

I would like to dedicate this work to my Mum and Dad, and also to the fond memory of my cat, Beep, who sadly died while I was studying here.

Beep: 1988 – 20th May 2004.

[1] Roth (1997). *Being Happier* (Sheffield; Kingfield Press).

Preface

These are some of the empirical observations and concrete experiences that ultimately led to my writing this thesis.

When I first experienced the live diffusion of electroacoustic music – at the Sonic Arts Network Conference in 1999, which I attended mainly out of curiosity during my undergraduate Music studies at Newcastle University – I was impressed by the results, but confused by the process. 'Isn't it strange,' I thought, 'to play a CD in front of a quietly seated audience?' 'Isn't it strange that people obediently applaud at the end?' Like many newcomers, I also wondered, 'What is the performer actually *doing* with that mixing desk?' It is also true to say that, at the time, I was a newcomer to electroacoustic music in general. As a mainly instrumental performer I was intrigued – but equally confused – by this new musical language, with its strange performance practice conventions.

Later, while working towards an M.Sc. in Music Technology at York University, I was able to experience more electroacoustic music, and began also to compose some myself (as an undergraduate I had specialised in 'music technology,' but my compositional work had always been essentially 'instrumental' in nature). Similar questions crossed my mind in diffusion concerts, only this time they were better informed: 'How many channels are we dealing with?'; 'How is the surround sound effect being achieved?'; 'What is pre-composed, and what is happening live?' and ultimately, 'What *is* the performer doing with that mixing desk?!' I was almost disappointed to learn that, most of the time, we are dealing with only two channels (stereo), and that 'all' the performer is doing, is controlling the levels of various pairs of loudspeakers. 'It really *is* just like playing back a CD in front of an audience,' I thought. I don't recall what I imagined the performer *might* have been doing, but I suppose I hoped it might be more exciting than that!

Also while at York, I was introduced to the Ambisonic surround sound system, and became fascinated with the possibility of controlling sound sources three-dimensionally. As my final M.Sc. project, I wrote two pieces of software for Ambisonic panning. Whilst researching this project, I happened to read a paper by Jonty Harrison entitled 'Sound, Space, Sculpture: Some Thoughts on the What, How, and Why of Sound Diffusion.'[2] This satiated some of my curiosities relating to sound diffusion. It even helped to answer my recurring question, 'What is the performer doing with the mixing desk?' It helped me to gain a clearer idea of the 'what' and 'how' of sound diffusion, but I was still confused about the 'why.' Why would composers restrict themselves to working in stereo when multichannel formats were so easily available? I was also unconvinced of the methods (the 'how,' I suppose) of sound diffusion: how could performers be satisfied with such a crude and apparently imprecise means of controlling the final outcome? Surely, given that the performer is doing so 'little' anyway (this is what I thought at the time), multichannel formats would far more flexible and precise; why bother with this seemingly tokenistic 'performance' stage? I was fairly convinced that my Ambisonic plugins, which would allow composers to accurately control the three-dimensional projection of sound in the *studio*, were a much better solution.

It was not until I arrived at Sheffield University that I had the chance to diffuse electroacoustic music myself. My supervisor, Adrian, asked if I would like to diffuse a piece at the forthcoming concert (if I recall correctly, the concert was entitled 'Xhbt B,' and took place in the Long Gallery of Sheffield's Millennium Galleries in November 2001). The piece was Åke Parmerud's *Les Flûtes en Feu* (1999).[3] I wasn't entirely sure what was expected, but from Harrison's paper I gathered I had to 'make the loud material louder and the quiet material quieter – thus stretching out the dynamic range to be something nearer what the ear expects in a concert situation.'[4] This seemed straightforward enough, and armed with this information, I set about creating a graphical score to aid me in my task.

[2] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 117-127.
[3] Various (2000). *Métamorphoses 2000*. (Compact Disc - Musique & Recherches: MR2000).
[4] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 121.

There were plenty of flute *crescendi* that, I felt, would lend themselves very well to such dynamic exaggeration. I also decided, from listening to the piece at home, that certain sections would sound effective if they disappeared into the distance at the front or rear of the hall.

I enjoyed my first hands-on experience of sound diffusion but, I have to say, I still wasn't completely convinced that it was a strictly necessary procedure, nor that the means for achieving it were especially well-suited to the task. Much of the time, I felt limited by the fact that the material was stereo, and restricted by the at times unrelenting pace of the music in comparison with the cumbersome task of manipulating faders. I still ended up thinking that – enjoyable as the experience was – it would ultimately be better to do a multichannel composition to begin with. Certainly, some of my observations were naïve. On the other hand, some of them were – and I believe still are – valid, and continue to be raised by experienced practitioners and newcomers alike.

One aspect that I found particularly hard to comprehend, both conceptually and aesthetically, was the fact that sound diffusion seemed to involve the spatialisation of material that is itself already spatialised. I had no difficulty, conceptually, with the idea of 'panning' a monophonic source in real time, nor with the idea of spatialising stereophonic sources within a larger sound field: the software I developed at York was capable of both of these things. But something didn't 'add up.' I began to feel that there must be a fundamental difference between 'panning' and 'diffusion.'

During this time, I was also engaged in the preparatory stages of composition. With newly acquired binaural microphones I made field recordings of trams, 'whirring' sounds, traffic, populated environments, sparse environments, and various other materials. I also continued to seek out interesting pieces of free audio processing software, an interest I had acquired whilst in York.[5] In late 2001, two *Studies in VST* resulted, and – in early 2002 – a piece entitled *Coney Street*, named after the road in York where the main source recordings were taken.

In 2002, as an Easter holiday, I spent a week travelling around the costal perimeter of Scotland, gathering many recordings on the way. On my return to Sheffield, I visited the MAXIS Symposium, where I heard and saw many interesting things, among them Shawn Decker's sound sculptures. Two compositions – *Smoo Cave* and *Pointless Exercise No. 911980* – emerged: the former is a 'soundscape' composition comprising recordings made inside Smoo Cave, Durness; the latter uses recordings of two of Decker's sound sculptures as its source material. In the summer of the same year, I paid a twelve-day visit to Bourges to take in the electroacoustic music festival. In the months following my return, two more compositions – *Everything I Do is in Inverted Commas* and *Graffiti 2* – were produced. The latter is in 5.1, reflecting my long-standing interest in surround sound.

In all, between late 2001 and early 2003, I produced just over an hour of electroacoustic music. Overall, I was happy with this, but I must confess to having found the compositional process difficult, and frustrating at times. Regardless of the quality of the results, I somehow didn't feel that I could justify, on a personal level, my work in the studio. I felt as though I was somehow 'missing the point' of what I was doing.

In November 2002 – in an attempt to address this – I recorded five lengthy interviews with composers working at the University of Sheffield Sound Studios: two with M.Phil. students, two with Ph.D. students, and one with my supervisor. I quizzed each interviewee about their views on electroacoustic music, with a particular focus on one of their recent works. Opinions were very divided. I also 'interviewed' myself, recording my thoughts at the time. Seemingly, I was very concerned with the relationship between 'abstract' sounds and 'referential' sounds. In retrospect, I was

[5] ...and one which I have consistently maintained. The 'Freeware VST Plugins Directory' that I started in York has since developed into 'Freeware Audio Resources Online' (FARO), an online database available at http://www.freeware-audio-resources.net.

clearly having difficulty in understanding what the structuring principles of electroacoustic music might be: in the absence of pitch and metrical rhythm (with which I was fully accustomed), how does a structure take the listener from one moment to the next? I could understand that obvious referentiality might be an alternative structuring force, but this seemed to be at odds with a lot of the music I was hearing, which as far as I could tell was completely non-referential. This, I could not understand: it is no wonder I was having difficulties in the studio!

It was also at around this time that I became interested in a piece of software that a fellow Ph.D. student, Dave Moore (whom I had met while we were both teaching at Barnsley College) was working on. This marked the beginning of a long-standing collaboration in the development of the M2 (Mooney and Moore) Diffusion System, incorporating Moore's software and making further software and hardware developments as a means to addressing some of my early (and subsequent) observations.

Further to my formative experiences in the studio and at the diffusion console, I decided – around a year after arriving in Sheffield – that I wanted to focus my attention on gaining a better understanding of the *raisons d'être* of electroacoustic music, the purpose and process of sound diffusion, and the relationship between the two. To my mind, I was convinced that there must be some commonality between the diverse opinions and approaches. What started off as the intention to read more books and papers, take a few notes, and generally organise my thoughts more coherently, rapidly evolved into a much larger enquiry, of which this thesis is the ultimate result.

At an early stage during my research it became apparent to me that there were many different attitudes and working procedures embodied within the electroacoustic idiom. Opinions with respect to what electroacoustic music 'is,' in particular, were divided; comments such as, 'Well, *that's* not electroacoustic music' were fairly common. The trouble is that these opinions were often somewhat at odds with each other: what, to one person, constituted 'electroacoustic music,' was the absolute antithesis to someone else. It also became evident that attitudes towards the process of sound diffusion were similarly fragmented. This stands to reason: if opinions regarding the nature of the music itself are varied, then so too will be attitudes towards the ways in which it should be performed. This thesis is therefore as much about electroacoustic music in general as it is about sound diffusion in particular.

I now personally feel that I have gained an understanding of the theory and practice of electroacoustic music, and as such it is sometimes necessary to remember that newcomers to the idiom are quite possibly experiencing, for the first time, similar difficulties to those I experienced. I now recognise the cause of my original misunderstandings, and hope that this thesis will be useful to anyone experiencing similar problems. Students at De Montfort University, Leicester (where I have been lecturing for the past eighteen months) have certainly responded well. Equally, I hope that practitioners more experienced than myself will gain something from the perspective offered in the following pages. In either case, the intention is to present my ideas about electroacoustic music, sound diffusion, and propose ways in which these can be regarded as homogeneous.

My approach is deliberately broad because I feel that this is an appropriate way to address an idiom that is so diverse. Of course the down-side to this approach is that some of the issues discussed could be taken much further. My argument, however, is this: in sound diffusion, works from across the (constantly shifting and expanding) spectrum of the electroacoustic idiom will need to be performed side-by-side, and the diffusion system used will have to cope with this. If future systems are to be designed with this borne in mind, then a broad understanding of the electroacoustic idiom – even if this is at the expense of certain 'details' – will be highly beneficial. Therefore, if this thesis seeks to answer only one question, then that question (actually three questions!) would be, 'What *is* electroacoustic music, how is it performed, and what can we do to perform it better?'

James Mooney, August 2005.

Table of Contents

Introduction

This research is intended to be of use to anyone seeking to engage in the design, development and implementation of new sound diffusion systems for the public performance of electroacoustic music, or wishing to make modifications and improvements to existing systems. In this capacity, it does not assume a highly detailed knowledge with respect to the nuances of the electroacoustic idiom, although this will, of course, be beneficial. This thesis is also, however, intended to be of interest to practitioners already experienced in the art of sound diffusion, and in this respect it is hoped that new insights will be realised. On the basis of this dual functionality it is the purpose of this thesis, on a very broad level, to deconstruct work in the field of sound diffusion into a practice that at the highest level – like electroacoustic music itself – comprises both technological and aesthetic aspects. The ultimate outcome of this enquiry will consist in the proposition of various means by which new sound diffusion systems that are better suited to the purpose can be developed, designed, and implemented.

From a purely technological perspective, the performance of electroacoustic music can involve the mediation of a wide variety of different technologies. These range from microphones, to compact discs, DAT tapes and other fixed media, to synthesisers, signal processing units, sophisticated real-time control devices, and software running on laptop computers. All of these are, ultimately, subservient to the loudspeaker as the means to making the musical results audible. A sound diffusion system, therefore, must ideally be designed in such a way so as to accommodate all of these technologies (and more), in the context of a live performance. If this objective is to be attained, it will clearly be necessary to know, broadly, what the technological demands are likely to be. This is to be one of the primary concerns of Chapter 1. By considering each technology as a 'creative framework,' an approach will also be made towards an understanding of the unique way in which technology is approached in the composition and performance of electroacoustic music.

The second chapter will follow, in a sense, the opposite path to the first, adopting a line of enquiry that is, in the first instance, aesthetic, and from this standpoint moving towards the more technological considerations. The context of this chapter will also be historical, to an extent. Specifically, it will be observed that the praxes of *musique concrète* and *elektronische Musik* – the cultural phenomena now widely acknowledged to have developed into what we now understand to be 'electroacoustic music' – embodied markedly different aesthetic standpoints. These were, in many respects, diametrically opposed and will be classified, respectively, as 'top-down' and 'bottom-up.' Although *elektronische Musik* and *musique concrète* are often regarded as 'extinct,' it will be demonstrated that top-down and bottom-up approaches are still evidenced in modern electroacoustic music.

The significance of this in the context of sound diffusion rests in the fact that each of these standpoints adopts a fundamentally different approach to the use of technology as a creative framework. It is proposed that the broad division of electroacoustic works into top-down and bottom-up – although in some respects crude – will serve as a useful guide in evaluating the kinds of communication that sound diffusion systems will be required to facilitate.

Chapter 3 will begin to tie together the technological and aesthetic threads exposed in the previous two chapters. It will be noted, importantly, that top-down and bottom-up approaches to the creative process engender not only different kinds of music, but also different diffusion techniques. The basis of this observation will rest in the assertion that top-down and bottom-up works differ fundamentally in terms of what it is they seek to communicate musically, and therefore necessitate entirely different modes of performance. Of course, diffusion systems will ideally be able to facilitate both approaches.

Chapter 4 evaluates several existing diffusion systems in terms of their appropriateness for the task of performing works from across the technological and aesthetic spectrum of the electroacoustic idiom. This evaluation will be realised according to a set of evaluation criteria proposed at the start of the chapter. It is also suggested that these criteria will be useful in the design of future systems. Several observations and conclusions will be made, among them the suggestion that – again, broadly speaking – diffusion systems can be very roughly divided into those that are best suited for the performance of top-down works, and those that are best suited for bottom-up. On this basis there would appear to be a clear need for new systems able to accommodate both types of work.

Chapter 5 will be devoted to a description and evaluation of the M2 Sound Diffusion System, which was co-developed by the author as an integral part of this research. In this respect, Chapter 5 is basically an extension of Chapter 4. It should be noted that the M2 is not being presented as an 'ideal' system, although it will be shown that several of its paradigms and architectures have proved effective.

The final chapter proposes directions in which practical and theoretical research into the development of future sound diffusion systems can be taken. These conclusions will be drawn from observations made in all of the previous chapters. These propositions do not represent a 'design specification,' as such (although several specific suggestions will be made) but are rather less prescriptive in nature. Overall, the conclusions made will seek to engender an approach to the development of new sound diffusion systems that is *inclusive* in terms of both technologies and aesthetic standpoints.

Overall, this thesis should be regarded as a heuristic tool directed toward the development of sound diffusion systems. An interpretation of the electroacoustic idiom, and an evaluation of existing systems, are both undertaken to this end. This enquiry seeks to address the following broad questions:

- What can we learn from the technological demands of the electroacoustic idiom?
- What can we learn from the aesthetic nature of electroacoustic music?
- What can we learn from the logistical nature of sound diffusion itself, and from differing approaches to and aesthetic attitudes towards it?
- What can be learned from existing sound diffusion systems?
- What can we learn from the design and implementation of the M2 Sound Diffusion System and from the feedback obtained from practitioners who have performed with it?
- What might be a constructive way forward?

These purposefully open-ended questions serve as a basic template for the discussion that follows, with each chapter, in turn, addressing one of them. There are no definitive answers, but there are many perfectly valid responses. It is for this reason that this thesis seeks to engender an inclusive attitude toward the development of future sound diffusion systems, which are – after all – the final means by which electroacoustic music is heard.

1. Electroacoustic Music: Defining Characteristics

1.1. Introduction

As a creative and intellectual discipline, electroacoustic music can be somewhat elusive and esoteric, ill-defined and multiplicitous in nature. It is the author's experience that definitions of electroacoustic music tend either to be so generalised as to be of little use, or else they in fact allude to particular 'sub-divisions,' which cannot, in isolation, be regarded as accurate descriptors of the idiom as a whole. The objective of the present chapter is therefore to approach a broad and holistic definition of 'electroacoustic music' that is neither overly embracing nor overly restrictive. The purpose of arriving at such a definition is not specifically with a view to identifying whether or not individual works fall into the category (although some might indeed find this useful or interesting) so much as to present a context within which works from across the technological and aesthetic spectrum of the idiom can be meaningfully interpreted and performed alongside each other without jeopardising their essential nature. In this respect, the proposed definition is intended to be of use specifically in the field of sound diffusion.

1.2. Existing Definitions

The expression 'electroacoustic music' (*musique électroacoustique* in the original French) was first adopted in the late nineteen-fifties, and is described by Chion as 'a term that was intended to be ecumenical and unifying, leaving aside questions of aesthetics and the means used to achieve the end [...] encompassing not only music on a recording medium but also music combining musical instruments and tape.'[6] Taken in its historical context, the term was thus intended to unify the praxes of *musique*

[6] Chion (2002). "Musical Research: GRM's Words". Website available at: http://www.ina.fr/grm/presentation/mots.light.en.html.

concrète and *elektronische Musik*, which differed strongly in both technical means and aesthetics (we will return to these in Chapter 2).

This definition is useful, in one sense, because it engenders an inclusive understanding of the electroacoustic music that embraces a wide variety of technological means and aesthetic standpoints; a similar definition, in this respect, is to be sought in the present chapter. Chion's definition is less useful, however, in as much as it does not positively identify any traits of the electroacoustic idiom: it excludes specific technologies and aesthetic standpoints, but fails to give an indication of what might alternatively be regarded as defining characteristics.

Wishart uses the expression 'sonic art,' which he explains as follows:

> We can begin by saying that sonic art includes music and electroacoustic music. At the same time, however, it will cross over into areas which have been categorised distinctly as text-sound and as sound-effects. [...] I personally feel there is no longer any way to draw a clear distinction between these areas.[7]

Wishart's essential belief that 'all sound is music' is one with which the author happens to agree, but it is not particularly useful to anyone wishing to gain a more in-depth understanding of *specific* musical praxes. Wishart's assertion that 'sonic art includes music *and* electroacoustic music' surely implies that there must be some difference between cultural understandings of these two expressions, yet his explanation of sonic art would seem to suggest that these differences are inconsequential.

Poorly substantiated and/or overly generalised definitions have frequently led to the assertion that 'electroacoustic music' is in fact a meaningless expression:

> The hybrid word 'electroacoustic,' used by this organization and by [Harrison] in founding BEAST, is [...] problematic – rather than reconciling antagonisms, it merely creates confusion because it doesn't really mean anything at all!

---

[7] Wishart (1996). *On Sonic Art* (Amsterdam; Harwood Academic): 4.

'Computer music' tells us very little too, because it describes the tool, not the music – it is hardly any more helpful than a term such as 'piano music.'[8]

However, practitioners of the electroacoustic art must surely have an understanding of the characteristics that define it. If we assume that this is true, then the expression 'electroacoustic music' is certainly *not* meaningless; it just happens not to have been verbally explained in a way that adequately reflects the collective understandings of practitioners. The following sections seek to challenge the assertion that 'electroacoustic music' is meaningless, and re-instate the expression as a useful descriptor through the identification of consistent defining characteristics that are neither superficial nor overly specific.

1.3. Towards a More Useful Definition

Unlike many spheres of musical activity, it is difficult (if not impossible) to define electroacoustic music in terms of its musical, structural, or stylistic attributes: these are simply too numerous and variable. It is therefore erroneous to treat the expression as indicative of a particular 'style' or 'genre' in the same way that one might categorise 'minimalism' or 'heavy metal'; electroacoustic music would seem to be an altogether broader designation than either of these. If electroacoustic music cannot be reliably identified by its musical characteristics alone, then what exactly *is* responsible for its unique identity?

The expression 'electroacoustic music' itself suggests that we are dealing with a practice that utilises electrical and/or electronic ('electro-') equipment to create and present sounds ('-acoustic') that are in some way 'musical.' Accordingly, the following sections will approach a useful definition of the idiom via an examination of the technologies used in its creation and performance.

[8] Harrison (1999). "Diffusion - Theories and Practices, with Particular Reference to the BEAST System". *EContact!* **2**(4). Electronic journal available at: http://cec.concordia.ca/econtact/Diffusion/Beast.htm.

1.4. The Omnipresent Loudspeaker

It is an inescapable fact that electroacoustic music must be reproduced, partially if not totally, via loudspeakers.[9] This condition, in a sense, precedes all aesthetic considerations and is arguably the most basic (but not necessarily the most important) defining characteristic of the idiom. Of course it is not the *only* defining characteristic – if this were the case then any sound emanating from any loudspeaker in any context could justifiably be described as electroacoustic music. Nonetheless, it is a difficult task indeed to find a sensible definition of electroacoustic music, or of any of its fantastically numerous sub-divisions, that does not make reference to this very simple reality:

> It has become common practice to use the term 'electroacoustic music' to describe a wide variety of musical praxes involving the mediation of loudspeakers.[10]

> The performance of electroacoustic music necessarily entails amplifiers and speakers.[11]

It can safely be said that the loudspeaker can be regarded as a defining characteristic of electroacoustic music. In his text entitled 'The Mirror of Ambiguity,' Harvey goes so far as to use the expression 'loudspeaker music.'[12]

1.5. Fixed Media

In addition to presentation over loudspeakers, the presence of a recording medium such as analogue tape or (as is nowadays most common) compact disc is also frequently invoked as a defining characteristic of electroacoustic music. This is mentioned, for instance, by Chion in the quotation given earlier.

[9] This is intended to mean any form of loudspeaker, including headphones or any manner of 'loudspeaker-like' transducer. The term 'audio decoding technology' will later be proposed.
[10] Windsor (2000). "Through and Around the Acousmatic: The Interpretation of Electroacoustic Sounds". In Emmerson [Ed.] *Music, Electronic Media and Culture*: 7.
[11] Rolfe (1999). "A Practical Guide to Diffusion". *EContact!* 2(4). Electronic journal available at: http://cec.concordia.ca/econtact/Diffusion/pracdiff.htm.
[12] Harvey (1986). "The Mirror of Ambiguity". In Emmerson [Ed.] *The Language of Electroacoustic Music*: 188.

It is proposed, however, that the 'fixed medium', as it is sometimes called, is not as indispensable as the loudspeaker in mediating our understanding of what constitutes electroacoustic music. With current technology it is increasingly possible to generate sound in real-time, without the need for an intermediate recording medium, a practice frequently performed in concert with laptop computers (this, of course, would not have been the case in the formative years of the genre). Here, there may be no fixed medium, but the sounds must still make themselves known via loudspeakers. Perhaps some would argue that 'laptronica', as such, is not electroacoustic music. However, it seems valid at this stage to suggest that such live performance could warrant inclusion given that it takes place so frequently alongside more traditional recorded work, and given the undeniable dialogue that exists between 'laptop performers' and 'composers for tape'. It should also be noted that when recorded sounds are processed live in performance (again, perhaps using portable computer technology), the recorded sounds must be stored somewhere – on the computer's hard drive for instance – and in these cases a fixed medium is indeed present. However, take into consideration the fact that *live* sounds could be processed in real time (as in Hans Tutschku's recent work,[13] for example) and once again we find ourselves without a fixed medium.

It is proposed that the fixed medium, as a defining characteristic of electroacoustic music, is nowadays outmoded, and a more generalised approach is therefore required.

1.6. Encoded Audio Streams

An encoded audio stream is a representation of real auditory events in some form analogous to, but immanently different from, their natural existential state. Because electroacoustic music relies on mediation via loudspeakers, it is additionally true to say that it relies on the presence of encoded audio streams to drive the loudspeakers. Encoded audio streams can also,

[13] Tutschku's *Cito* (2002), *SprachSchlag* (2000) and *Das Bleierne Klavier* (2000) are all written for instrument(s) and 'live electronics,' and do not include a 'tape part' as such:
Tutschku (2004). "Hans Tutschku". Website available at: http://www.tutschku.com/.

therefore, be regarded as a defining characteristic of the electroacoustic idiom.

Encoded audio streams can, at present, be analogue or digital, and in either case can exist in a *static* form on some kind of storage medium, or in a *transitory* form in the case of the transmission of encoded audio streams from one device to another. The signals represented in analogue media such as magnetic tape and vinyl are examples of *static-analogue* encoding, whereas media such as compact disc, DAT, and ADAT are carriers of *static-digital* encoding. An electrical signal from a microphone, travelling down a conductive wire is an example of *transitory-analogue* encoding, while a digital signal originating from a CD player and travelling down an optical cable is an example of *transitory-digital* encoding. These examples are summarised in Table 1, below.

	Static (Fixed Media)	**Transitory** (Transmission)
Analogue	Signals stored on: Analogue Tape; Vinyl LP; etc.	Electrical signal travelling down a microphone cable.
Digital	Signals stored on: CD; DAT Tape; Hard Drive; etc.	Optical signal travelling in an digital-optical cable.

Table 1. Summary of currently available audio encoding techniques, with representative examples of each.

When the streams are transitory, the encoded signal exists in a continuously time-varying state and therefore has no immediate 'totality' (i.e. it is not static). Consequently, the encoded stream itself is not fundamentally different to the original auditory phenomenon, insofar as both can only exist over a span of time. Static encodings, on the other hand, *are* fundamentally different because they uniquely offer up the possibility of representing auditory events in extraction from the constant passing of time.

Encoded audio streams can be obtained and manipulated in a number of ways. Microphones are perhaps the most obvious means of obtaining encoded audio streams; such devices therefore fall into the category of 'audio encoding technologies.'

1.6.1. Encoded Audio as an 'Abstraction Layer'

The process of encoding audio results in the abstraction of real auditory events that, in turn, makes the practice of electroacoustic music possible. The existence of audio in an encoded form can be thought of as an 'abstraction layer' (see section 6.2.7, page 289) as it is only ever a means to an end, rather than an end in itself. In other words, the process of encoding audio is only really useful if it is to be *de*coded – that is to say, rendered audible via loudspeakers – at some point in the future. In simple terms the encoded audio thus represents an intermediate stage in the otherwise unitary process of the creation and reception of sound, as illustrated in Figure 1, below.

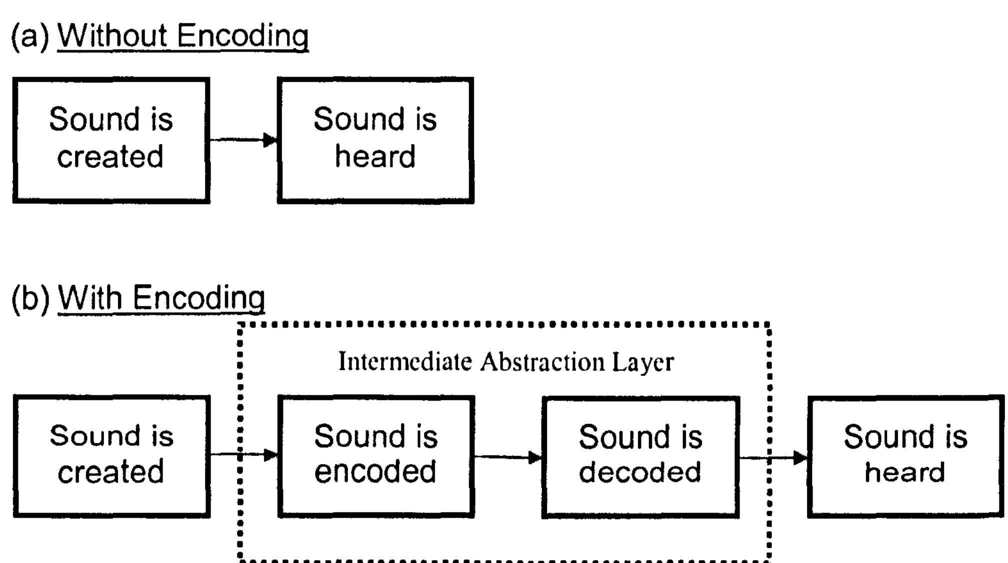

Figure 1. Simplified diagrammatic representation of the process of creation/reception of auditory events both with and without an intermediate encoding/decoding abstraction layer.

The advantage of the abstraction layer is that it creates a 'point of entry' into what would otherwise be a closed process.[14] This has advantages regardless of whether we are dealing with static or transitory encoding. In Figure 1(a), although we may have control over the way in which the sound is created (which objects and how they interact), this is basically the limit of our control – there is nothing else we can do with respect to the sound itself in the interim period between sound creation and sound reception – and in this respect the actuality of the sound is a foregone conclusion once it has been created. In scenario (b), by contrast, *there is the opportunity for various forms of intervention within the abstraction layer itself.*

If we are dealing with a transitory encoding, then the encoded signal can be subjected to various real-time processes before it goes on to be decoded. In the case of a static encoding the possibilities are more radical because we have immediate access to the totality of data within the 'sound object,' as opposed to experiencing it over a period of time. The delightfully-named CDP algorithm 'Sausage' performs a 'granular reconstitution of several soundfiles scrambled together.'[15] This is only possible, of course, if we have simultaneous access to the totality of the encoded audio events, and such processes can therefore only be performed on static encodings.

Another notable feature of static encodings is that there is a *physical* object – the storage medium itself – that indirectly represents the sound and that may itself be open to manipulation. This is particularly true of analogue storage media: tapes can be spliced; records can be 'scratched' by DJs; *et cetera*. In both cases the sound can be accessed and manipulated by various means that take place *within* the abstraction layer.

[14] This is true of any abstraction layer, in any context. The benefits of using abstraction layers as a means to achieving modularity in diffusion system design will be stated in section 6.2.7.
[15] CDP (2005). "CDP Home Page". Website available at: http://www.bath.ac.uk/~masjpf/CDP/CDP.htm.

The abstraction layer has another major advantage in that all it really does is provide a standardised framework (or 'language') for the encoding and decoding of (in this case audio) information. As an example, a system might convert (encode) variations in air pressure ('real' sound) – via a microphone – into a constantly varying voltage; this is a transitory-analogue encoding. Assuming that appropriate equipment exists that understands the 'language' (i.e. 'knows what the constantly varying voltage means') then it should be possible to recover the original sound via the inverse process of decoding. The important point is that the decoder is merely interpreting data of a particular format (in this case, a constantly varying voltage) and, therefore, *where those data originate from is irrelevant*. This means that (to extend the current example) continuously varying voltages can be generated *independently of any real auditory events* and subsequently decoded using appropriate technology. This is, of course, the basis for sound synthesis (the process works in an analogous way in the digital domain) and it is now stated that this possibility arises as a direct consequence of the abstraction layer brought about by the process of encoding and decoding auditory events.

1.6.2. Encoding Spatio-Auditory Attributes in Multiple Streams

Notwithstanding the fundamental difference between static and transitory modes of audio encoding, there is an important respect in which both are similar: multiple independent streams can be utilised in combination to encode certain spatio-auditory attributes. Two-channel stereophony,[16] for example, involves the simultaneous use of two audio streams to encode information regarding spatial positioning along a single axis. This is possible by virtue of 'phantom imaging' between loudspeakers upon decoding, which will be described more fully in section 1.10. The two constituent channels are physically separate

[16] Blumlein (1931). "Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems". International Pat. No. 394325. Alan Blumlein Official Website. Available online at: http://www.doramusic.com/patents/394325.htm.

entities, but in order for the spatial effect to be realised correctly, certain conditions must be observed with respect to their relationship with each other. Conventionally, in the case of two-channel stereophonic recordings, 'channel 1' is 'left' and 'channel 2' is 'right,' with 'left' and 'right' alluding to the opposite extremities of an axis in the horizontal plane. At the most basic level, this convention must be observed during both encoding and decoding stages in order for the phantom imaging to be successful. It should be stressed, however, that this is merely a convention, and there is no strict reason why the two channels of a stereo recording could not be used to project auditory images along any other arbitrarily defined axis.[17] Nonetheless, whatever spatial encoding conventions are observed, it is important that the *same* conventions are assumed during both encoding and decoding stages.

Following the same basic premises it is possible to combine *more* than two encoded audio streams to allow for the storage and transmission of more detailed spatial information. Under present technological constraints the minimum number of channels required to encode fully three-dimensional spatial information is three, with varying degrees of success depending on the particular technique used. Indeed, the multichannel approach is becoming more and more common, with software and hardware applications increasingly supporting various formats. This trend has, as many have predicted, experienced something of an acceleration in recent years with the increasing acceptance of 5.1 surround sound as a standard convention for multichannel audio.[18] The success of 5.1 can, perhaps, be attributed to the fact that, like two-channel stereo, it represents a convenient standard that can easily be

[17] There are in fact reasons, which pertain mainly to the loudspeaker-related criteria for successful phantom imaging (some of these will be discussed more fully in 3.7) and vary depending on the specific phantom imaging technique used. So long as these criteria are met, it is proposed that the present argument is valid.

[18] The '5.1' configuration was devised by Dolby Laboratories and is therefore sometimes also referred to as 'Dolby Digital,' although very similar systems are produced by others (DTS being a good example). Introductory articles are available via the Dolby website: Dolby Laboratories (1999). "Surround Sound: Past, Present and Future". Website available at: http://www.dolby.com/assets/pdf/tech_library/2_Surround_Sound_Past.Present.pdf. Dolby Laboratories (1999). "Some Guidelines for Producing Music in 5.1 Channel Surround". Website available at: http://www.beussery.com/pdf/buessery.dolby.5.1.pdf.

adhered to in the encoding and decoding of audio signals. Accordingly, practitioners can work within these constraints with the relative assurance that appropriate apparatus will exist for the widespread dissemination of their work in both commercial and domestic settings. However, like stereo, it is just a convention, and some composers have criticised the format as being ill-suited in various respects for more creative audio-related applications:

> The specification for 5.1 [...] is very demarcated in terms of what you're supposed to put on each component in the 5.1 array. The LFE is supposed to have *that* and nothing else, and the centre speaker is supposed to have the dialogue and nothing else, and the frontal stereo is supposed to have the soundtrack, the image, the music, *et cetera*, and very little else. The surrounds kind of vaguely interact with that but mostly for effects. [...]. That's about the extent of it. So they're kept fairly distinct in terms of their functions. But most composition usage doesn't want to keep them that distinct. It wants actually to blur those things. It wants to have something that swirls around the space and then it wants something that's very precise and in a very particular location; those kind of things. And I think that's actually quite difficult to do with 5.1. [...] Very often, the surround speakers are a smaller unit altogether than the main stereo. So, actually, its very difficult to get a fully balanced equal surround, for example, in the average 5.1 array.[19]

Consequently, it is fairly usual for electroacoustic composers to wish to supersede the restrictions of 5.1, with eight-channel works at present being fairly common. Spatial encodings encompassing even more than eight channels are not unheard-of,[20] and given the increasing acceptance of multichannel audio in a wider context there is no reason to believe that the demand for more channels should not continue. Regardless of the specific number of channels, these 'spatial encodings' follow the same basic principles as stereo and 5.1 – the same standard must be followed in the encoding and decoding of the audio streams – but unfortunately, in the absence of a more widely accepted convention,

---

[19] Harrison (2004). "An Interview with Professor Jonty Harrison". Personal interview with the author, 8th September 2004. See Appendix 1 for full transcription.

[20] Marco Marinoni's *Del Vuoto Incanto* (2003) is an electroacoustic work for 16-channel tape and was first performed at University of Limerick on 5th April 2003:
Marinoni (2004). "Marco Marinoni's Home Page". Website available at: http://www.geocities.com/marco_marinoni.
David Worrall's *Cords* (1990) also utilises sixteen audio channels and was written specifically for performance in the ACAT Dome, which will be discussed later in section 4.13:
Worrall (2002). "David Worrall Home Page". Website available at: http://www.avatar.com.au/worrall/.

numerous different methods exist (some of these will be described in Chapter 3). It is perhaps even more unfortunate that much commercially available software and hardware is unable to efficiently support these demands, and given that the practice of sound diffusion is confined to a relatively small specialist minority, it seems that this circumstance is unlikely to change in the near future.

All of this indicates a likelihood, in the context of electroacoustic music performance, that (in addition to other varying technical demands yet to be discussed) there will be a high degree of variation inherent in the multichannel demands of works. It is therefore important that arbitrary multichannel configurations should be allowed to co-exist in concert programmes, and that there is a need for sound diffusion systems that are able to facilitate this; it will be suggested, in Chapter 4, that this is not presently the case. This is at least in part attributable to the fact that much hardware and software only allows for encoded audio streams to be efficiently and straightforwardly handled on an *individual* basis, or at best in groups that conform to well-established conventions such as stereo and 5.1 (although the latter case is far less common, particularly in the hardware domain). Ultimately, however, these circumstances are due mainly to consensus (as opposed to any actual physical constraints), and in the vast majority of cases access to individual audio channels is available regardless of any conventional/theoretical 'grouping.' Accordingly, a conceptual framework that allows for an *arbitrary* number of encoded audio streams to be treated as fundamentally related – the 'coherent audio source set' – will be proposed and described in sections 3.5 and 3.6.

It will be noted that the existence of loudspeakers and encoded audio streams is a dialectical necessity: ultimately, neither is useful in the absence of the other. As such, loudspeakers and encoded audio streams can *both* be regarded as defining characteristics of electroacoustic music. The ability to encode auditory events – whether the encoding is static or transitory, digital or analogue – allows us to represent and interact with these transient happenings symbolically and, if necessary, outside of the time domain. This can be regarded as providing an 'abstraction layer' between the normally unitary processes of sound-creation and sound-reception. *Within* this abstraction layer, various forms of technologically mediated intervention can take place before the final decoding of the encoded audio stream(s). *It is therefore within this abstraction layer that electroacoustic music itself exists.*

1.7. Audio Technologies

In surveying the main audio technologies that are currently used in the creation and performance of electroacoustic music, key technologies will be summarised in terms of six categories, namely: audio encoding technology, recording and playback technology, synthesis technology, audio processing technology, software (computer) technology, and audio decoding technology. This summary will be followed by an overview of some of the possible technological combinations that could be implemented in the live performance of an electroacoustic work, and specific examples of each of the primary combinations will be given. The technical demands of electroacoustic music, as presented in this chapter, are not intended to be fully exhaustive, but rather should serve as a basic template for the (purely) technological scope of electroacoustic music. This, in turn, will give an indication of the kinds of technical demands that are likely to be placed on concert sound diffusion systems. It should be noted, however, that these technologies are by no means exclusive to electroacoustic music. Furthermore, electroacoustic music can incorporate many other technologies that will not be listed here. We will return to both of these issues later. The following sections review those key technologies that deal *directly* with encoded audio streams; as such, these can be used to affect various interactions within the abstraction layer of encoded audio. As such, it can be assumed that these technologies will be used in both the composition and performance of electroacoustic music.

1.7.1. Audio Encoding Technologies

These are technologies that mediate between *real* auditory events and encoded audio streams. Microphones and pick-ups of various kinds are the most obvious examples. These convert physical vibrations into (usually) continuously varying voltages, and as such perform a transitory-analogue encoding. It is conceivable that in the near future audio encoding technologies will be developed that encode directly into transitory-digital formats. The abstract principle of converting between real sounds and encoded audio streams, however, would remain the same.

1.7.2. (a) Recording and (b) Playback Technologies

Recording technologies facilitate the conversion of *transitory* encoded audio streams (whether analogue or digital) into *static* encoded audio streams. Playback technologies, on the other hand, carry out the conversion from static back to transitory. We have already alluded to the role of such technologies in electroacoustic music by way of reference to works presented on a 'fixed' recording medium. Recording technology dates back to the late 1870s, when inventors including Edison and Berliner patented 'phonograph' inventions crudely comparable to more modern vinyl record players.[21] Subsequent developments included magnetic recording onto wire (1898), steel tape (1929), analogue-optical recording techniques and, finally, magnetic tape in 1935.[22] Pulse code modulation (PCM), a means of encoding audio signals digitally that is still in widespread use today, was patented by Reeves in 1942; computer implementations of PCM technology were first developed by Mathews in 1957, with the first hardware recorder

[21] Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press): 1.
[22] Ibid.12-13.

developed by NHK Technical Research Institute in 1967 following necessary developments in digital electronics.[23]

1.7.3. Synthesis Technologies

In abstract terms it can be stated that synthesis technologies are used to *generate* encoded audio streams directly, without the need for any audio encoding technology. The earliest example of synthesis technology is generally acknowledged to be the Telharmonium, patented in 1897 by Thaddeus Cahill. Other notable early synthesis tools include the Thérémin (1924), Sphärophon (1927), Dynaphone (1928), Ondes Martenot (1928) and Trautonium (1930).[24] All of these are used to directly generate transitory-analogue streams (in the form of continuously varying electrical voltages) which can then be decoded via loudspeakers. This is in contrast with earlier musical technologies (violins, pianos, *et cetera*) whose sounding characteristics are defined directly by their physical attributes and the means of excitation employed to cause these physical bodies to resonate. More recent developments in synthesis use digital audio technology to implement synthesis algorithms that would be extremely difficult or impossible to achieve by analogue electronic means.

1.7.4. Audio Processing Technologies

An audio processing technology is one that facilitates some kind of *modification* to existing encoded audio streams before the process of decoding takes place. Modifications could be applied 'on the fly' to transitorily encoded audio streams: this is what happens, for example, when an amplifier is used to apply gain to a signal, or when a filter is used to attenuate part of the frequency spectrum. Equally, modifications could be applied to static encodings: tape splicing is an example relevant to the historical development of electroacoustic music, while the process of 'scratching' performed by DJs would represent a more

---

[23] Roland-Mieszkowski (1989). "Introduction to Digital Recording Techniques". Website available at: http://www.digital-recordings.com/publ/pubrec.html.
[24] Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press) 1-3.

contemporary example. Among the earliest examples of audio processing were those performed by Pierre Schaeffer in his early *musique concrète* compositions in 1948. Short recordings were looped, and acetate discs played back at different speeds, thus altering the qualities of the resulting sounds. Physical means of audio processing, such as plate reverbs, were often used in the formative years of electroacoustic music (although it could perhaps be argued that these are actually audio decoding technologies that modify the sound), as well as techniques that exploit the physical means of electrical sound reproduction, such as tape-head delay machines. Electronic means, filters for example, followed presently.

1.7.5. Software Technologies

The impact of software technology on electroacoustic music cannot be underestimated, not least because computer software is effectively able to emulate most of the technologies previously mentioned (with the notable exception of the necessarily physical roles performed by encoding and decoding technologies), and allows the competent user to achieve goals that, for instance, would simply not be possible in the hardware domain. Generally, software used to perform recording/playback, synthesis, or audio processing tasks, will not need to be differentiated from its corresponding category given above. That is, in many cases, implementations of computer and software technology can be regarded as combining elements of the technologies given in items 1.7.2 through 1.7.4. However, certain electroacoustic music praxes (such as algorithmic composition) make use of attributes that are exclusive to software and microprocessor technology, and that do not fall into any of the other categories. Although structuring principles in electroacoustic music will not be discussed in this chapter, they will form a substantial part of the discourse of Chapter 2. Certain live processing techniques also make use of the unique capabilities of software applications. Some examples will be given in section 1.8.4.

1.7.6. Audio Decoding Technologies

In abstract terms these technologies affect the conversion from encoded audio streams back into real sound. Normally this is done by converting the continuously varying voltage of a transitory-analogue stream into physical movements within some kind of material, the resulting vibrations subsequently causing sound waves to travel. In the vast majority of cases, such technologies can be classed as 'loudspeakers,' but alternative means are not unheard-of. Decker has implemented various mechanical systems in sound-sculptures that could be classified as audio decoding technologies.[25] Tarsitani describes 'planephones' – audio decoding technologies that would not perhaps intuitively be regarded as loudspeakers – as follows:

> 'Planephones' [are] special multiphonic sound diffusion systems [in the sense of 'propagators of sound,' as opposed to 'diffusion systems' as they will be described in Chapter 3] developed by Michelangelo Lupone at CRM and presented for the first time at the Musica Scienza festival of 1998. Planephones are vibrating sound systems consisting of panels of different materials (wood, metal, plastic, leather) and shapes installed in artistic venues; with them it is possible to design the acoustic space according to the architecture of the hall and give the sound the timbral quality of the materials employed.[26]

Again, it is feasible that at some point in the future the necessary means to convert from digitally encoded streams into real auditory events will be developed, but the basic premises of the decoding process will remain the same in this eventuality. It will be noted that, because of their dependence on encoded audio streams, each of the audio technologies discussed previously is ultimately indebted to audio decoding technology as the means of generating the final, audible, result. We must therefore continue to consider the loudspeaker as being of primary importance.

---

[25] Decker's *Scratch Studies* (2000), for example, converts encoded streams into mechanical movements of motors attached to pieces of stiff wire, which, in turn, scratch against metal surfaces to create sound. Of course, whether or not this constitutes electroacoustic music is a matter of debate. A piece by the author, *Pointless Exercise No. 911980* (2002), uses recordings of two of Decker's sound sculptures as its source materials.
Decker (2005). "Shawn Decker - Installations". Website available at: http://www.artic.edu/~sdecker/Finstallations.html.

[26] Tarsitani (1999). "Musica Scienza 1999". *Computer Music Journal*, 24(2): 88-89.

1.8. The Technical Scope of Live Electroacoustic Music

It has been proposed that, in terms of technological requirements, the only indispensable prerequisite in the performance of electroacoustic music is the decoding technology (loudspeaker). Invariably – because loudspeakers are effectively useless in the absence of any encoded audio streams to decode – any given piece of electroacoustic music must make use of one or more of the other technologies given under section 1.7, and may of course include all of them.

Of the technologies described in section 1.7, three can be described as 'primary,' because they are able to output a transitorily encoded audio stream with no transitorily encoded input stream. These are: audio encoding technologies; playback technologies; and synthesis technologies. Audio processing technologies must be regarded as 'secondary' because they require an encoded audio stream as input. Such technologies therefore cannot be used on their own, but only in conjunction with one or more of the other technologies. Recording technology can, of course, be used during performance, but if this is the case then it must invariably be used in conjunction with playback technology if the performance is real time and the recorded audio streams are to be part of the live performance itself. Live sampling, for example, would be regarded as a use of audio encoding, processing, and *playback* technologies for the purposes of the present discussion, even though some of the audio streams may have been recorded during the performance. Software technology, as mentioned previously, will not be evaluated separately: if software is used to synthesise sounds, then this will be considered a synthesis technology, and so on. Audio decoding technologies (loudspeakers), as we know, are always present, and therefore are not a variable. Under these criteria, the live performance of a piece of electroacoustic music may exert any combination of the technical demands summarised in Table 2 below.

		...only: no secondary processing	...with (secondary) audio processing technology
Primary Technologies	Playback technology...	**1.8.1.** Streams from a fixed medium	**1.8.4.** Streams from a fixed medium processed prior to decoding
	Synthesis technology...	**1.8.2.** Streams synthesised in real time	**1.8.5.** Streams synthesised in real time and processed prior to decoding
	Audio encoding technology...	**1.8.3.** Acoustic sources (often, but not always, encoded into streams for amplification and/or diffusion)	**1.8.6.** Streams encoded from acoustic sources and processed prior to decoding

Table 2. Summary of the possible technological demands in the live performance of works of electroacoustic music.

The three primary technologies – shown in the left-most column of Table 2 – can be used on their own, or in conjunction with some secondary form of audio processing before the decoding process takes place. This gives a total of six individual technological 'units,' shown in the grey shaded area. A piece of electroacoustic music can incorporate any number (one or more) of these units, in any combination, and could of course feasibly incorporate all six. By factorial calculation, it can therefore be determined that there are a total of sixty-three unique technological combinations, which could perhaps be regarded as electroacoustic 'instrumentations.' Individual explanation of each of these permutations is clearly unnecessary; the following sections will describe each of the six main components, giving examples from the electroacoustic repertoire where possible.

1.8.1. Works presented solely from a fixed medium

This category encompasses the repertoire of works stored on magnetic tape, compact disc, ADAT, hard drive and so on, and intended for playback or diffusion over loudspeakers in performance. Historically this has been the most common category for electroacoustic works, and although this arguably continues to be the case, it seems likely that this primacy will be challenged in the future as technology advances. Of course, the audio statically encoded on the medium can be composed of any combination of acoustic sounds, synthesised sounds, and sounds which have been further processed after initial encoding. Among the first electroacoustic works to combine synthesised and acoustic sounds on a fixed medium was Stockhausen's *Gesang der Jünglinge*, completed in 1956.[27]

1.8.2. Works consisting of sounds synthesised in real-time

This category represents electroacoustic works that are not stored on a recording medium, but rather realised electrically or electronically in real-time using synthesis technology. Again, such performances are invariably made audible by way of amplifiers and loudspeakers. As an example, recent work at University of East Anglia[28] involves the use (or perhaps 'misuse') of synthesisers in unconventional ways for the live real-time creation of sonic material. Here, the internal workings of early (and presumably cheaply available!) digital synthesisers are manipulated directly, using pieces of wire to create short-circuits, with interesting and often unexpected sonic results. Others perform using software-based laptop systems and various other acoustic and electronic sounding devices. The result is an improvised ensemble of electronically generated sounds broadcast over loudspeakers.

As with all of these technological profiles, it is common for the technology in question to be used in conjunction with other forces.

[27] Stockhausen (1991). *Stockhausen 3 - Elektronische Musik 1952-1960*. (Compact Disc - Stockhausen-Verlag: 3).

[28] UEA artists including Shigeto Wada, Phil Archer, Stef Edwards, Adam Green, Jonathon Manton, Nick Melia, Tom Simmons and Bill Vine performed at *Sound Circus* – the Sonic Arts Network Annual Conference – at De Montfort University, Leicester, on Sunday 13th June 2004.

Hindemith's *Concertina for Trautonium and String Orchestra*[29] (1931) might be considered among the first examples of this, this work combining acoustically generated sounds with sounds generated synthetically by the Trautonium. Messiaen's *Turangalîla Symphony*[30] (completed 1948) includes the Ondes Martenot among the traditional orchestral instruments. *Fête des Belles Eaux* (1937), also by Messiaen, is scored solely for six Ondes Martenots. Whether or not such works constitute electroacoustic music, however, is debatable, their having been heavily conditioned by instrumental traditions. One could suggest that Messiaen's *Turangalîla Symphony* makes use of the Ondes Martenot as an orchestral instrument with a novel and interesting timbre, but that in every other respect this particular instrument is treated in the same way as any other. In other words, it could be argued that the piece of music would not be subject to a terminal change in its essential character were the Ondes Martenot to be replaced with an acoustic instrument. The main creative locus of the work has been in composing a piece of orchestral music and not in any aspect specifically brought about by the presence of new audio technologies. We will return to this discussion in sections 1.12 and 1.13, as a means to further refining our definition of electroacoustic music.

1.8.3. Works consisting of acoustic sound sources

Live acoustic means of sound production are frequently used in the performance of electroacoustic music. Such forces typically include the voice and other 'traditional' musical instruments but are not necessarily limited to these. Any piece of electroacoustic music utilising acoustically generated sound, howsoever created, would fall into this category.

---

29 Hindemith and Sala (1998). *Elektronische Impressionen*. (Compact Disc - Erdenklang Musikverlag: EK81032). This CD also includes Hindemith's *7 Triostücke für 3 Trautonien* and works for Trautonium by Oskar Sala.
30 Messiaen (2000). *Turangalîla Symphony / L'Ascension*. (Compact Disc - Naxos: 8.554478/79).

In this context, instrumental sounds may simply be allowed to sound naturally, or equally may be encoded via microphones, pick-ups, and so on. It would be fairly unusual for a piece of electroacoustic music to invoke this criterion as its *only* technological prerequisite, although Nance's *Parables of Control* (2005) could – in a sense – be defined as an electroacoustic work for acoustic sound sources only.[31] It is far more common, however, for acoustically generated sounds to be used alongside other technologies. Smalley's *Clarinet Threads* (for clarinet and tape)[32] and Harrison's *Abstracts* (for large orchestra and tape)[33] are two examples of this, both of these combining playback technologies (fixed media) with acoustically generated sound sources. In the former case the clarinet part is usually encoded via microphones during performance for the purposes of amplification, whereas Harrison's piece does not call for the orchestra to be amplified and therefore instrumentalists do not require microphones.

1.8.4. Works consisting of fixed medium sources processed in real time before decoding

Some work consists in fixed medium recordings that are subjected to analogue or digital signal processing at the time of performance, before amplification and broadcast over loudspeakers. Much work in the field

[31] Nance (2005). Personal email communication with the author. At the time of writing, Rick Nance is completing PhD studies in electroacoustic composition at De Montfort University, Leicester, UK. *Parables of Control* is a series of studies for solo cello in which a pre-composed electroacoustic 'tape' part (itself derived from recordings of cello sounds) – although used in performance – is not actually heard by the audience. The material from compact disc is heard only by the cellist – via headphones – with the instrumental performance being realised by the performer in continual response to it. In this way, the 'tape' part itself acts rather like a performance 'score.' The use of playback and decoding technologies is, of course, still prerequisite to the performance of the work, but as the audience hears only acoustically generated cello sounds, the piece can be regarded as something of a special case. *Analogies of Control* (2005) and *K* (2005) are pieces with tape for cello and solo trumpet, respectively. Here, electroacoustic scores – again, not heard by the audience – are used in the same way, but in this case the sound generated by the performer's instrument is heard alongside a different 'tape' part. The cello pieces were first performed by Thomas Gardner in Liverpool on 29th November 2005, with further performances imminent at the time of writing.

[32] Various (1990). *Computer Music Currents 6*. (Compact Disc - Wergo Schallplatten: WER 20262).

[33] Harrison (2001). In Various (2001). *Cultures Électroniques No. 15: 28e Concours International de Musique et d'Art Sonore Électroacoustiques*. (Compact Disc - Mnémosyne Musique Média: LCD278074/75).

of what has become known colloquially as 'laptronica' involves such practice (here, the fixed medium is a computer hard drive), Joseph Hyde's 'Live Sampling' projects of recent years being a good example. Hyde describes these projects as follows:

> Since 1999, Joseph has been performing as an experimental DJ and VJ using a wide and eclectic range of material (everything from early Musique Concrète to German electronica; from environmental sources to kitsch charity-shop vinyl), mixing across genres and cutting and distorting using a mix of lo-fi techniques such as vinyl vandalism and high-tech methods based on digital sampling [...] and live image manipulation... He performs in concert venues, clubs and unusual sites, both solo and with other performers.[34]

Processing of recorded audio sources is often performed using real-time digital signal processing software such as Cycling 74's Max/MSP[35], Miller Puckette's *Pure Data*[36] (more commonly abbreviated as *PD*), and STEIM's *LiSa*[37], but hardware audio processing technology is also, of course, a possibility.

1.8.5. Works consisting of audio streams synthesised in real time and processed before decoding

In this case, sounds that are synthesised directly in real time are subjected to some form of intermediary processing while the streams are still encoded. This could be achieved by both hardware or software means, or a combination. Again, software applications such as Max/MSP can be used to perform both synthesis and live processing.

---

[34] Hyde (2004). "Projects - Live Sampling". Website available at: http://www.theperiphery.com/projects/LiveSampling.htm.

[35] Zicarelli (2004). "Max/MSP for Mac and Windows". Website available at: http://www.cycling74.com/products/maxmsp.html.

[36] Puckette (2004). "Software by Miller Puckette". Website available at: http://crca.ucsd.edu/~msp/software.html.

[37] STEIM (2004). "STEIM Products - LiSa X". Website available at: http://www.steim.org/steim/lisa.html.

1.8.6. Works consisting of acoustically generated sounds encoded and processed in real time before decoding

Audio streams encoded from acoustic sources via microphones or pick-ups may be manipulated using audio processing technologies, amplified, and broadcast in real-time, where they will be heard in addition to the 'natural' unprocessed sounds. Such works are often described as pieces for acoustic forces 'with live electronics.' Tutschku's *Cito* (2002), *SprachSchlag* (2000) and *Das Bleierne Klavier* (2000), for example, are all written for instrument(s) and 'live electronics,' and do not include a 'tape part' as such.[38] In each case, sound from an acoustic instrument is transduced via microphones, and the encoded audio streams are subjected to various signal processing techniques before being decoded.

It should, of course, be remembered that electroacoustic works may utilise these six technological units in any number and in any combination. In the future it may become necessary to invent further nomenclatures to describe these practices, but at present it can be argued that they all fall under the 'umbrella term' of electroacoustic music. Emmerson states that "mixed' electroacoustic music (instruments and tape), 'live' electronic music (using processing of sound produced by a performer) and [...] 'real-time' computer music' are component parts of a larger 'electroacoustic music field.'[39] Indeed, it is proposed that several of the technological combinations (particularly 'tape only,' 'tape and live instrument(s),' and any of the combinations involving 'live electronics') are – almost of themselves – culturally identifiable as characteristically electroacoustic 'instrumentations.' The technological plurality and eclecticism inherent in electroacoustic music is also observed by Waters.[40] Nevertheless, it seems

[38] Tutschku (2004). "Hans Tutschku". Website available at: http://www.tutschku.com/.
[39] Emmerson (2000). "'Losing Touch?' The Human Performer and Electronics". In Emmerson [Ed.] *Music, Electronic Media and Culture*: 194.
[40] Waters (2000). "Beyond the Acousmatic: Hybrid Tendencies in Electroacoustic Music". In Emmerson [Ed.] *Music, Electronic Media and Culture* AND Waters (2000). "The Musical Process in the Age of Digital Intervention". *ARiADA* 1. Electronic journal available at: http://ariada.uea.ac.uk:16080/ariadatexts/ariada1/content/Musical_Process.pdf.

possible that the inclusion of all of the above within the genre of 'electroacoustic music' may be controversial to some. However, it is undeniably the case that works encompassing all of these categorisations are frequently presented side-by-side in public performance. Several of the examples given previously co-habited the programme for the Sonic Arts Network conference in 2004. This being the case it is clearly necessary that provision be made for the juxtaposition of such varied musical forces; this will become an important item within the criteria for the evaluation of sound diffusion systems, which will be given later in Chapter 4.

1.9. Creative Frameworks

The previous sections have sought to approach a workable definition of electroacoustic music via an examination of the various technological means engaged in its live performance. Although this is a good starting point, the findings are not wholly enlightening because nowadays an enormous volume and variety of musical material is generated and distributed by these means. With the exception of exclusively aural traditions, it is reasonably difficult to imagine any field of musical activity that does not make fairly extensive use of audio technology in some aspect of its production or dissemination. Nonetheless, it seems to be widely accepted that audio technology, in some way, plays a *particularly* significant role in electroacoustic music. Notwithstanding certain technological 'instrumentations' that are characteristic of the idiom (this is somewhat tautologous, and does not constitute an acceptable definition on its own), further examination of the significance of audio technology as a defining characteristic in electroacoustic music, with a particular focus on how the technology is appropriated and how this differs from other applications that utilise the same technology, will be useful.

From this point onwards, the expression 'creative framework' will be used to connote any entity, construct, system, or paradigm – whether physical or conceptual – that is used as a partial or total means to realising artistic output. Western classical notation, for example, is a creative framework frequently used in the composition of instrumental music. Traditional musical instruments are creative frameworks that contribute towards the performance of such musical works. 'Real' sound and synthesised sound are two creative frameworks frequently engaged in the composition of electroacoustic music. The technologies outlined in section 1.7 are all individual creative frameworks within which electroacoustic music is both composed and performed.

The postulation is as follows: every creative framework, due to its very nature, offers up specific affordances[41] to its user and, therefore, engenders particular ways of working. In this respect, creative frameworks are not transparent, neutral mediators of artistic expression but, rather, exhibit certain intrinsic 'biases.' A deliberately crude and exaggerated example may be helpful. As a creative framework, a paint-brush (echoing Windsor) affords painting or, one might say, is biased in favour of – and therefore engenders – this particular mode of operation. A violin affords playing violin music, and is of its nature inclined towards *this* way of working. The violin is no more appropriate as a means to painting pictures than the paint-brush is as a means to playing violin music. We could, of course, *try* to 'misappropriate' the frameworks in this way, but are unlikely to be rewarded with any high degree of success; this is because each framework engenders a particular way of working and is resistant, to an extent, towards different ways of working. This does not mean, of course, that creative frameworks can only ever be used exactly 'as intended.' The example given is deliberately exaggerated as a means of illustrating that creative frameworks each have their own particular affordances and biases, which – in most sensible cases – are not quite so extreme.

---

[41] The expression 'affordances' was devised by psychologist James Gibson, and is explained by Windsor as follows:

> Objects and events are related to a perceiving organism by structured information, and they 'afford' certain possibilities for action relative to an organism. For example, a cup affords drinking, the ground, walking. [...] Gibson's term for this kind of meaning is 'affordance.'

In the present context, the term 'affordance' is used as a convenient way of connoting something that is 'made possible' by a creative framework as a function of its unique nature.

Windsor (2000). "Through and Around the Acousmatic: The Interpretation of Electroacoustic Sounds". In Emmerson [Ed.] *Music, Electronic Media and Culture*: 11.

A creative process can be regarded as one in which multiple creative frameworks are engaged interactively, eventually resulting in the finished work. The expression is deliberately broad because this allows us to examine, at a very abstract and generalised level, the essential characteristics of the means by which electroacoustic music is created and performed.

The categorisations given in Table 2 are merely *technological* requirements, and it is certainly not suggested that *any* practice eliciting these requirements must 'by definition' be a piece of electroacoustic music. We must therefore seek to identify what is missing from our definition at this point. By examining the creative frameworks of electroacoustic music, we can seek to gain an understanding of the specific ways in which these are appropriated by electroacoustic musicians. In this way it is proposed that further defining characteristics of the idiom can be uncovered. The following sections will examine one specific creative framework that is central to all electroacoustic music: the loudspeaker.[42]

1.10. Analysis of the Loudspeaker as a Creative Framework

When an auditory event is *encoded*, it becomes detached, in a sense, from its original causal agent. When a statically encoded (i.e. recorded) sound is made audible by amplification through a single loudspeaker, its placement in space – the locus from which the sound emanates – is determined not by the spatio-temporal location of the original sounding agent, but by that of the loudspeaker (and the source(s) from which its encoded audio streams originate). This may seem obvious but historically the musical possibilities must have been exciting: the spatial location of a sound source was no longer limited by necessity to that physically attainable by a performer. That is to say, loudspeakers can be deployed with relative ease in positions that would be at the very least inconvenient for human performers; on the ceiling

[42] Similar examination of the other creative frameworks may well reveal further defining characteristics of the electroacoustic idiom. For the purposes of the present thesis, this is not necessary.

or up the walls of an auditorium as in (to cite a fairly early example) Varèse and Le Corbusier's installation of *Poème Électronique* in the Philips Pavilion (1958).

Of course, loudspeakers do not have to be used 'monophonically,' as mere surrogate performers. If multiple loudspeakers are used collectively, various 'phantom imaging' techniques can be employed (with varying degrees of success dependent on both the specific technique used and the particular formation of loudspeakers relative to the audience) to create the illusion of sound sources emanating from in-between individual loudspeakers. In other words, the location of a sound source *as perceived* need not even be limited to the physical location of any single loudspeaker. (Surely this must have seemed unbelievable to composers working with new audio technology in its formative years). This process is, of course, greatly aided by the absence of a visual point of reference. Where loudspeakers are used monophonically, as described in the previous paragraph, the loudspeaker will itself tend to be perceived as the sound source. Where multiple loudspeakers are used to create phantom images, however, individual loudspeakers (if the phantom imaging is successful) cease to be perceived visually as spatially rooted sound sources. Technologically, most listeners will realise that in reality the loudspeakers *are* the source of what they are hearing, but perceptually the sound sources will be invisible, because there is no visible entity fixed at the location from which sound seems to be emanating: the perceptual bond between sounding phenomenon and visual referent breaks down. Clozier summarises this situation as follows:

> A single loudspeaker is a sound projector and causes the listener to identify the physical point of emission with the point of origin of the sound source... The sound [...] is located at the point within the general space at which the loudspeaker is placed... We do not hear the sound, we hear the loudspeaker... Two loudspeakers [...] constitute [...] a link between two points situated in general space. They constitute an imaginary space-line on which may be projected singular or particular musical space... Except for extremes at right and left [where loudspeakers would behave 'singly' as visible sources] the sound is created within its [*own*] space.[43]

[43]Clozier (1998). "Composition, Diffusion and Interpretation in Electroacoustic Music". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 246-250.

It can therefore be seen that the ability to (re)create sound over multiple loudspeakers, without the previous necessity of a visible, spatially-rooted sounding agent, leads to the true *emancipation of spatiality* in an auditory context. It is for this reason that electroacoustic music has been described as 'a uniquely spatial medium.'[44] Of course this, in turn, raises the highly pertinent question of how to utilise the phenomenon of 'disembodied space' meaningfully in a compositional context. Unsurprisingly, the need to invoke spatiality on a more-than-superficial level is strikingly abundant among electroacoustic composers, as neatly summarised by Worrall:

> It has become important to find more appropriate ways to explore the nature and use of [spatiality] so as to avoid using it purely decoratively in effects such as sounds 'whizzing' around the auditorium.[45]

The specific ways in which composers seek to attribute meaning to their compositional use of space are invariably tied to the particular aesthetic model to which they subscribe; we shall return to this matter in Chapter 2.

1.11. Methodological Choices in the use of Loudspeakers

Of course all of the above is theoretically applicable to *any* sound reproduced over two or more loudspeakers: voices amplified in real-time in public address systems; live recordings of classical music played back on home stereo equipment; multitrack studio recordings of rock bands projected through loudspeakers in night clubs; sounds broadcast over the airwaves to radios and television sets, and so on. How do we determine whether an implementation of loudspeaker technology is 'electroacoustic' or not? Some might indeed argue that *any* sound broadcast via loudspeakers is electroacoustic. Technically, this is a justifiable assertion: we are using electrical devices for the broadcast of sound. However, recalling section 1.2, if we subscribe to this proposition, we run the risk of delivering a definition of electroacoustic music that is effectively meaningless, and probably at

[44] Austin (2000). "Sound Diffusion in Composition and Performance - An Interview with Denis Smalley". *Computer Music Journal*, 24(2): 20.
[45] Worrall (1998). "Space in Sound - Sound of Space". *Organised Sound*, 3(2): 94.

odds with the intuitive understandings of most practitioners. A greater degree of clarity is therefore to be sought.

In many applications, the role of loudspeakers is to simulate real and familiar sonic events. That is, loudspeakers are used simply as an alternative means to realising auditory scenarios that could ostensibly be facilitated by 'real' sounding agents. For example, the effect of playing a CD of rock music at home is essentially comparable to the sonic experience of a real rock band performing in one's living room. This is a function of the immediately recognisable nature of the sounds themselves (we can identify the sound of a drum kit without the affirmative presence of the visual stimulus itself) and the essentially 'realistic' way in which the spatio-auditory illusion is built (the way instruments have been recorded and deployed within the stereo field). These two factors are mediated by, and dependent upon, the manner in which the listener perceives the stimulus, which will result from a combination of their expectations (sub-conscious) and deliberate listening strategies (conscious). These are all aspects that are fairly specific to the loudspeaker as a functional tool or creative framework, and as such, engender particular 'methodological choices' that must be negotiated by anyone wishing to use loudspeakers for this purpose. It is proposed that the way in which these decisions are negotiated may be helpful in defining electroacoustic music as a practice distinct from other applications of the same technology. Some of the 'framework-specific' negotiations that must be made in the use of loudspeakers are outlined in the following sections.

1.11.1. Choice and Treatment of Sound Material

The reproduction of sound over one or more loudspeakers is a deliberate act, and as such, choices must be made with respect to the sonic material to be presented, and indeed the sonic material to be omitted. Consider the process of recording a performance of an orchestral violin concerto. We might place a stereo pair of microphones in front of the orchestra, a further pair towards the rear of the hall to capture some of the reverberant characteristics of the venue, and a single directional microphone close to the soloist. The resulting recorded signals are mixed together in order to create the notionally 'realistic' illusion of an orchestral performance on reproduction over a pair of loudspeakers. According to this objective, proactive efforts will have been made to capture the 'important' auditory information (the orchestra, the soloist) and, equally, to omit unwanted information (ambient traffic noise, or coughing from the audience for example). Equally, microphones will have been positioned so as to reinforce the 'realistic' aspirations of the recording: a microphone placed too close to the soloist's instrument will result in an unrealistically 'magnified' sound for example. Accordingly, care will have been taken in recording and subsequent mixing to adhere to the aims and objectives of the recording. These represent very specific choices (and omissions) of sound material, and very particular ways in which sound material is obtained and treated. Although recorded sound has been used as an example here, similar considerations would inform the choice and treatment of synthesised sounds. It is here that one must consider the 'purpose' for which one is choosing, creating, or manipulating sound material.

1.11.2. Construction of Auditory and Spatial Illusion

This aspect makes reference to the fact that sound produced over loudspeakers is inherently illusory in nature. When our orchestral recording is reproduced over a pair of loudspeakers, we are not really hearing the sound of an orchestra, but rather we are provided with the illusion of it. When a sound source seems to move around a venue, the *actual* sound sources – multiple loudspeakers – are static in reality, but are again providing the audience with the *illusion* of a moving sound source. Indeed the use of *spatial* illusion in an auditory context is a highly important consideration in the construction of the auditory illusion as a whole. In the orchestral example given, the relationship between orchestra and soloist is essential to the success of the auditory illusion. In all likelihood, the signal from the soloist's microphone will be positioned centrally and statically within the stereo field, while the 'orchestral' and 'ambient' signals will occupy a wider portion of the stereo field, for instance. The soloist is likely, also, to be mixed to a level slightly higher than that of the orchestra, for clarity. The illusion would break down, or at very least be considered bizarre, were the soloist to appear to move around the auditorium, or if the sound of the solo violin was masked by an erroneously loud orchestral accompaniment. These represent very particular decisions regarding the use of auditory and spatial illusion and, again, reinforce the overall intended mode of listening.

1.11.3. Mode of Listening

This final – and arguably most important – aspect refers to the overall *intention* of the broadcast sound with respect to how it should be perceived. It is the mode of listening that completes the auditory illusion, and ultimately determines its relative success. Obviously this is highly dependent on the listener but can also be regarded, to some extent, as a function of the choices made in the previous two areas.

Most listeners, on hearing a recording of an orchestral performance, will know that they are not hearing the live sound of a real orchestra. However, if shrewd decisions have been made regarding the choice of sound materials to be presented and omitted, and in the way that these materials have been recomposed and presented to the listener, then the auditory illusion should be sufficiently persuasive. In this case, the listener will be 'encouraged' to suspend their disbelief, and imagine that they are indeed listening to the sound of a real orchestra. If this happens, then the listener has adopted a particular 'mode of listening' that is congruent with the intentions of the material being presented, as a direct result of the *way* it has been presented. Such modes of listening are often adopted sub-consciously – in playing a CD of the orchestral recording, the 'correct' mode of listening may automatically be assumed by the listener – but can also be conscious acts, whereby the auditory illusion on offer is either wilfully embraced or deliberately subverted. For example, a listener may *expect* (subconsciously) that playing a CD by the heavy metal group Metallica will present them with the auditory illusion of the real band performing, and the results may be more successful if the listener deliberately (consciously) chooses a complementary *listening strategy* in order to 'complete the illusion'. These are two aspects – subconscious and conscious – of the mode of listening. Of course this response is easy for the listener if the nature of the recording has been deliberately engineered to engender it. This, broadly speaking, is a function of the choice of sound materials and the way in which the auditory (and spatial) illusion has been executed.

It is proposed that the successful presentation of sound over loudspeakers, in any context, is dependent on the choices made in areas 1.11.1 and 1.11.2 having been successfully negotiated in accordance with their own objectives as articulated in 1.11.3. Further, it is proposed that any presentation of sound via loudspeakers *necessarily* entails choices in each of these areas, purely because it engages the loudspeaker as one of its frameworks. The extent to which these choices are well-informed or even deliberate may in some cases be debatable, but it is difficult to imagine a scenario in which they do not exist. It can also be seen that the choices made in each of these areas are mutually inter-dependent. Some examples have already been given but a further example may be helpful.

Imagine an announcement broadcast over the public address system in a train station. Some very specific decisions regarding the choice sound material have been made: usually some kind of pitched tone(s), followed by the sound of a human voice (and nowadays the voice may very well be live or pre-recorded). The intended mode of listening is predominantly semantic in nature, and purely functional, and the choice of sound material reflects this: the pitched tones are designed to attract the listening attention of the public and the words 'spoken' are designed to be intelligible (anecdotally this is often not the case). Owing to the nature of the sounding material, a listener will most probably enter the correct mode of listening more-or-less automatically (sub-consciously) but – particularly true in this example – an additional and deliberate (conscious) effort to establish this mode may be necessary in order for the perceived results to be intelligible, and therefore for the intended mode of listening to be successful. The use of spatial illusion is also functional: ideally we should perceive an omnipresent voice, or at least the sound of an announcer who 'appears' to be simultaneously located in several places at once. In other words, several loudspeakers are used, but without any phantom imaging as such: although the sonic illusion of a steward moving around the station may be appealing to some, it would jeopardise the essential aim of the exercise, intelligibility. Thus can it be illustrated that even the most seemingly benign use of loudspeakers necessarily involves explicit decisions in each of the three stated areas: such is the nature of this particular framework. There may, of course, be other areas, but the three examples given are particularly relevant to electroacoustic music, and have served to illustrate the point adequately.

1.12. Towards an Understanding of 'Electroacoustic Music'

It is proposed that the ways in which electroacoustic composers negotiate the methodological choices engendered by the creative frameworks (audio technologies) employed, might reveal some characteristic trends that will help to define the idiom more clearly. Specifically, it is suggested that electroacoustic music is that which *directly and categorically* explores (or seeks to promote an exploration of) the artistic potentials of the particular creative frameworks engaged in its realisation. This sets electroacoustic music apart from applications that make use of audio technologies for purely functional purposes, such as the train station announcement described in the previous section.

Of course, one could argue that *any* artistic practice is bound to entail, at some level, an exploration of the creative potentials of the frameworks it engages, and there is certainly much validity to this argument.[46] In particular, this leaves the relationship between electroacoustic music and

[46] There is a danger here of digressing into an argument about what constitutes 'art.' While this is undoubtedly an area of interest to practitioners of electroacoustic music, it is – realistically – beyond the scope of this thesis.

'popular' music praxes – which would not normally be regarded as purely 'functional' applications and can therefore be considered 'artistic' – open to debate, as these make particularly extensive use of audio technologies.

Let us consider, however, the nature of a deliberately stereotypical 'pop song.' There will normally be a fairly small number of musical instruments: for argument's sake let us assume that this consists of an electric guitar, bass guitar, keyboard/synthesiser, drum-kit, and a vocalist. The musical material will almost always be tonal, rhythmical, and divided into eight-bar sections, often in the standard 'verse-chorus-verse' structure. There will also be lyrics, often of a narrative and/or auto-biographical nature, whose subject matter is dependent to a certain extent on the 'genre' of the song. The list of criteria could go on much further, but it is worth considering, at this point, what the *primary* creative frameworks of our archetypal 'pop' song are. One might suggest: each of the musical instruments employed; (simple) western tonality; metrical rhythm; repetitious (but 'catchy') structuring principles; language (usually English); and so on. *These* are the frameworks within which the majority of the creative process takes place. Audio technologies will certainly be used to record, produce, and disseminate pop songs, but it cannot usually be argued that they represent the primary creative frameworks of this idiom; creative attention is focussed too strongly elsewhere. In short, it can be suggested that most 'popular' music styles do not appropriate audio technologies as primary creative frameworks, as such but, rather, make more functional use of their affordances.

The following sections will describe some musical works that are generally regarded to be 'electroacoustic' in nature. It will be shown that for each of the methodological choices described in the previous section – which, recall, arise from the very nature of the loudspeaker itself as a creative framework – 'electroacoustic' works can be identified that adopt the methodological choice in question as a specific focus for creative exploration. In that case we would expect that a characteristically 'electroacoustic' approach might negotiate the kinds of methodological choices outlined in the previous sections with a creative, experimental, artistic, or otherwise exploratory agenda, as opposed to approaching them as a functional necessity to the realisation of artistic objectives whose creative emphasis lies elsewhere (examples of the latter eventuality will also be given). This particular observation is useful if we wish to differentiate *electroacoustic* music from other musical praxes that make use of the same technologies.

It is also worth remembering that comparable methodological choices could almost certainly be described for any other creative framework imaginable: the loudspeaker is simply being used as an example as it is essential to all electroacoustic music.

1.13. Creative Exploration of Framework-Specific Methodological Choices in Electroacoustic Music

The 'exploratory' nature of electroacoustic music is supported by the broader notion of 'experimental music' (*musique expérimentale*), which is described by Chion as any music 'conceived in a spirit of research'.[47] Much theoretical work is devoted to describing ways in which an experimental approach to these criteria might be adopted: Wishart's explanation of the aural 'landscape,'[48] and Barrett's paper entitled 'Spatio-Musical Compositional Strategies,'[49] are two representative examples from a rich and varied body of literature. The following sections do not seek to imply that *all* electroacoustic music devotes itself to an exploration of the creative potentials of the loudspeaker; merely that the abstract ways in which the technological frameworks are approached *as a primary creative focus* might be indicative of a characteristically 'electroacoustic' methodology.

1.13.1. Exploration of Choice and Treatment of Sound Material

In most applications of audio technology, sound materials recorded or synthesised are generally familiar and/or functional in nature, that is,

[47] Chion (2002). "Musical Research: GRM's Words". Website available at: http://www.ina.fr/grm/presentation/mots.light.en.html.
[48] Wishart (1996). *On Sonic Art* (Amsterdam; Harwood Academic): 127ff.
[49] Barrett (2002). "Spatio-musical Composition Strategies". *Organised Sound*, 7(3).

they exist within cultural norms that are widely understood and accepted. To reiterate an earlier example, the rock music recording exhibits choices of sound material that for the most part fit within a culturally accepted model of 'rock music,' which in turn fits within a culturally accepted model of 'music' as a whole. While these sound sources have undoubtedly been subjected to various studio signal processing techniques, this will have been done in such a way so as not to disrupt the overall cultural congruency of the sound materials. Of course there can be a degree of creative exploration of the sound materials of rock music, but it must normally be subject to restriction if the end result is to be perceived as 'rock music.' The choice and treatment of sound materials in this context, therefore, cannot truly be considered creative, experimental, or exploratory in the fullest sense, insofar as the exploration is, at best, a secondary agenda.

In much (but not by any means all) electroacoustic music, contrastingly, creative exploration of sound materials is a *primary* agenda, representing the very *raison d'être* of the music itself. For example, the juxtaposition of culturally 'recognisable' and 'unrecognisable' sounds as the basis for a creative approach in this area is strikingly common in electroacoustic music. This process is described by McNabb, with reference to his much-cited 1978 tape work *Dreamsong*[50]:

> In computer music, one can take advantage of the unique plasticity of the medium to produce combinations of and transformations between arbitrarily different sounds. Again, it is the interplay of the familiar and the unfamiliar or unexpected which I find to have the most expressive potential. A juxtaposition of two completely unalike sounds need not be instantaneous, but can occur over any time-span, producing a gesture of great musical expressiveness and beauty, poignancy or tension, a concept which was the primary source of inspiration for my work *Dreamsong*. This juxtaposition of the familiar and the unfamiliar seems to me to be of great historical significance. Greater, say, than the introduction of the crescendo to Western orchestral music.[51]

Parmegiani expresses a similar agenda with reference to his fixed medium work *Sonare*:

---

[50] McNabb (1993). *Dreamsong*. (Compact Disc - Wergo Schallplatten: WER 20202).
[51] McNabb (1986). "Computer Music: Some Aesthetic Considerations". In Emmerson [Ed.] *The Language of Electroacoustic Music*: 145-6.

> For each of the 5 movements, I have chosen a pseudo-instrumental or synthesis sound which I sense will allow me to bring out its 'very essence,' to develop it until it is within the deepest levels of the soul... I had to imagine the most suitable interplay to bring out such 'intrinsic resonance.' Now of course, no interplay can be genuine unless certain freedoms are present within or inherent to it, the rule being that such an interplay should remain musical whilst the sounds, 'in a real context' become linked to or opposed to each other. No combat, just interplay for its own sake, for itself alone, and, at the same time, changes in contour, an opening or closing in the tone, range, patterns of rhythm, as if the work were a living being in itself...[52]

The above quote embodies the exploratory approach to sound material that is so characteristic of 'acousmatic' music, a sub-set of electroacoustic music that will be discussed further in the next chapter.

1.13.2. Exploration of the Construction of Auditory and Spatial Illusion

As described earlier, in most non-electroacoustic applications of audio technology, spatial illusion is built in order to affirm the illusion of one or more readily identifiable sounding bodies. A recording of an interview for radio broadcast, for example, may be spatially engineered so that the interviewer and interviewee seem to be located opposite each other; this could be achieved by panning the former towards the left of the stereo field, and the latter toward the right. The same can generally be said in the case of our rock music example: spatialisation will generally be carried out so as to reinforce the illusion of a group of real musicians performing on stage. Here, there is no exploration of spatiality in its own right, a practice that *is* common, however, in electroacoustic music.

Barrett's tape-only piece *The Utility of Space* is described by the composer as follows:

> *The Utility of Space* is an exploration of spatial musical structure. This exploration is done in two ways: with poetic spatial implication, and with

[52] Parmegiani (1996). *Sonare*. (Compact Disc - INA: E5203-275912). Quote from accompanying booklet.

> carefully controlled trajectories and sound magnitudes in real spatial places.[53]

Barrett describes the techniques used to explore the construction of spatial illusion in some depth in the article cited earlier, referring at times specifically to this piece.[54]

Obst's *Solaris*,[55] for chamber ensemble and live electronics, also represents an exploration of the construction of spatio-auditory illusion, although this work is very different, conceptually and aesthetically, from Barrett's *The Utility of Space* (aesthetic standpoints will be discussed more fully in the next chapter). Here, live instrumental sounds are captured by microphones. They are then subjected to various real-time digital signal processing techniques and spatialised – using Ircam's *Spatialisateur* software running on a NeXT computer workstation – among four loudspeakers, where the processed sounds will be heard alongside the unprocessed instrumental sounds. The way in which processed sounds are spatialised is determined algorithmically:

> The control of different parameters during the sound treatment is based on [...] mathematical principles, as is the spatialization of these sounds in the concert hall... The sound-projection uses [...] random numbers [...] by transforming them into distance values between 4 and 8 metres.[56]

1.13.3. Exploration of the Mode of Listening

Much electroacoustic music encourages us to listen in different ways, perhaps attempting to redress the balance of our visually dominated culture. This trait is characterised by an acute awareness of auditory surroundings – whether 'natural' or 'artificial' – and, unsurprisingly, is often reflected in musical output. Much electroacoustic music whose primary focus explores the mode of listening is descended from Pierre Schaeffer's school of *musique concrète*, which will be discussed more

[53] Various (2001). *Cultures Électroniques No. 15: 28e Concours International de Musique et d'Art Sonore Électroacoustiques*. (Compact Disc - Mnémosyne Musique Média: LCD278074/75). Quote from accompanying booklet: 18.
[54] Barrett (2002). "Spatio-musical Composition Strategies". *Organised Sound*, 7(3).
[55] Obst (1998). "Sound Projection as a Compositional Element: The Function of the Live Electronics in the Chamber Opera 'Solaris'". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*.
[56] Ibid.: 321-322.

fully in the next chapter. Examples of this can be found in Ferrari's *Presque Rien* tape pieces. With reference to *Presque Rien avec Filles*, the composer states the following:

> I emit [sonic] images, but more in the form of an empty frame which has to be filled in by listening.[57]

The piece, like its namesakes, comprises environmental recordings, presented in a relatively unaltered form. Thus, the listener is invited to construct the piece, or 'fill in the blanks' by adopting a somewhat intense mode of listening, centred on an appreciation of the evocative characteristics of these 'naturally occurring' sounds. As such, the piece has a lot in common with what is known as 'soundscape' composition, a practice predominantly developed at Simon Fraser University and characterised by the compositions and writings of R. Murray Shafer, Barry Truax, Hildegard Westerkamp and others, and closely linked to the practice of acoustic ecology. A recent composition by the author – *Smoo Cave* (2002) – falls into this category. In *Kits Beach Soundwalk*, Westerkamp 'takes the listener by the hand' and guides them, by way of a spoken-word narrative, through the process of listening to the soundscape (which is presented in a more or less unprocessed form) in great detail:

> It's a calm morning. I'm on Kits Beach in Vancouver... The ocean is flat, just a bit rippled in places... I'm standing among some large rocks, full of barnacles and seaweed. The water moves calmly through crevices. The barnacles put out their fingers to feed on the water. The tiny clicking sounds that you hear are the meeting of the water and the barnacles. It trickles, and clicks, and sucks, and... The city is roaring around these tiny sounds, but it's not masking them... But I'm trying to listen to those tiny sounds in more detail now. Suddenly the background sound of the city seems louder again. It interferes with my listening. It occupies all acoustic space, and I can't hear the barnacles in all their tinyness... Luckily we have band pass filters and equalisers. We can just go into the studio, and get rid of the city: pretend it's not there; pretend we are somewhere far away. [The sound of the city gradually fades into nothing.] These are the tiny, the intimate, voices of nature.[58]

[57] Ferrari (1998). *Electronic Works*. (Compact Disc - BV Haast: CD9009). Quote from accompanying booklet.
[58] Westerkamp (1996). *Transformations*. (Compact Disc - Empreintes Digitales: IMED 9631). Quotation transcribed from 'Kits Beach Soundwalk' (1989).

The above quotation is a condensed transcription of the beginning of the spoken-word narrative that guides the listener through *Kits Beach Soundwalk*. It is evident that, in terms of studio processing, relatively little has been done to the source recordings, and such processing that has taken place has been with the specific intention of encouraging the listener to focus their concentrated attention to the tiny details of the soundscape that might otherwise pass unnoticed. Here, as in Ferrari's piece, experimentation with the mode of listening is of primary compositional importance.

Listening, in the examples given, is not a passive behaviour, but an active one. Implicit here, broadly speaking, are two distinct levels of attention: a focus on the intrinsic characteristics of the sounds regardless of their cultural connotations; and a specific focus *on* the cultural connotations, that is, on the extrinsic or 'extra-auditory' information imparted by sonic events. Of course, the two need not be mutually exclusive, that is, the listener's attention will probably be variously divided between the intrinsic (sound as autonomous phenomenon) and extrinsic (sound as a 'signifier,' of sorts) aspects of the stimulus material. The exception to this can be found in the original incarnation of *musique concrète* (to be discussed more fully in the next chapter), which proposed a deliberate denial of all extra-auditory understandings of sonic material, in favour of a phenomenological appreciation of the sounds themselves, an act now widely regarded as difficult to achieve in practice:

> It proves very difficult to hear sound only in terms of an appreciation of its shape and spectral properties as Schaeffer seemed to advocate.[59]

Accordingly, theories of modes of listening, and compositions reflecting these, have been numerously proposed and documented. For example, some subscribe to the idea that our response to auditory stimulus, familiar or otherwise, is conditioned by innate 'survival instincts' that force us to attempt to identify the cause of everything we hear. Smalley

[59] Emmerson (1998). "Aural Landscape - Musical Space". *Organised Sound*, 3(2): 136.

refers to this process as 'source bonding,' and further refines the theory with the notion of 'gestural surrogacy,' the ability (or inability) to associate a sonic event with an imaginary physical gesture as its causal agent.[60] In brief, 'source bonding' refers to the ability to extrapolate a hypothetical source for a sonic event, and 'gestural surrogacy' to the ability to imagine what kind of gesture that sounding agent may have made to cause that sonic event to happen. The 'ecological' model of response to auditory stimulus is summarised by Windsor as follows:

> A growing body of research is attempting to study the perception of sounds which do not resemble traditional speech or music in a manner that takes account of the perception of sound sources and their potential to us as active (rather than passive) organisms. Within the field of ecological acoustics, sounds are not viewed as being perceived as abstract entities related only one to another, as 'tone colours' or timbres, nor are they perceived as standing for concepts or things, as signs. Instead they are seen as providing unmediated contact between listeners and significant environmental occurrences.[61]

One might regard the notion of 'transcontextuality' as something of an extension to the 'ecological' model of auditory response. While *musique concrète* proposed an interpretation of sonic stimulus that focussed on purely intrinsic characteristics, so transcontextuality can be regarded as something of the opposite: the primary object of attention *is* the context, or cultural understanding, evoked by sounds. Experiments in the relationship between text, context, and super-context are described by Savouret:

> Let us take another example, one that illustrates a text diffused within a foreign context: traditional musicians from Auvergne performing at a public concert geographically far away from the place where their music [...] is [normally] practiced... I discovered with them [...] that producing their text was made easier in the context of the hall (which was simply a volume plus spectators [...] from anywhere but the Auvergne) when I added a virtual sound 'over-context' to which they could relate... One of the musicians [...] said that for him the song of crickets from a valley in Cantal guided him as to how, that particular evening, he had to sing... For another musician, the sound of the passage of a herd of cows, far from hindering him in a work song provided him with indications as to the tempo of the song... Thanks to a better relationship between the performers and the context, brought about by a trick of sound – a super-context, so to speak – listeners who belonged to cultures that had little to

[60] Smalley (1997). "Spectromorphology - Explaining Sound Shapes". *Organised Sound*, 2(2): 110-112.
[61] Windsor (2000). "Through and Around the Acousmatic: The Interpretation of Electroacoustic Sounds". In Emmerson [Ed.] *Music, Electronic Media and Culture*.: 10.

do with the Auvergne were able to receive the entire text/context/super-context in a much more satisfactory way.[62]

Although this quotation does not refer directly to electroacoustic music, it does demonstrate the intriguing and subtle responses that emerge from exposure to auditory 'contexts,' (or, to rephrase, the ways in which one's perception of 'context' can be profoundly affected by auditory stimulus) and hints at ways in which composers can harness this phenomenon as a dimension to their musical discourse. One such composer is Ambrose Field, who exemplifies transcontextuality as follows:

> When applied to electroacoustic music transcontextual working is a method by which the extrinsic meanings of a sound can have a profound impact on their musical surroundings... [It] can be used as a tool to lend old or existing contexts new meanings... In *La Disparition*, [Christian] Calon creates a scene where an aeroplane appears to fly over a tropical jungle. This is initially accepted by the listener as the probability of this event occurring in reality is quite high. However, as the piece progresses, the sound of the aeroplane continues descending in pitch until it eventually forms the bass drone to the subsequent section... The sound of the aeroplane is the main transcontextual agent, as it performs both the function of 'aeroplane' and the musical function of a bass drone. During this transformation from aeroplane to bass drone, this sound has changed the way we perceived the context that surrounds it.[63]

It is clear that the possibility of including recognisable 'every day' sounds, and perhaps contrasting these with unfamiliar sounds (either processed or synthesised) in a musical context, gives rise to the 'mode of listening' as a valid area for diverse musical exploration, as demonstrated in the examples given above.

1.14. Back to Technology

It has been established that an understanding of electroacoustic music goes further than a simple description of the technological frameworks employed in its production and performance. Nonetheless, it can also be seen that the very possibility of a creative exploration in any of the areas discussed is facilitated by the existence of recording, audio processing, and synthesis

[62] Savouret (1998). "The Natures of Diffusion". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 351.
[63] Field (2000). "Simulation and Reality: The New Sonic Objects". In Emmerson [Ed.] *Music, Electronic Media and Culture*: 36, 50, 51.

technologies and the ability (in fact, the *necessity*) to reproduce the resulting encoded audio streams over loudspeakers. In the words of Trevor Wishart:

> The sophisticated control of this dimension of our sonic experience has only become possible with the development of sound-recording and synthesis and the control of virtual acoustic space via sound projection from loudspeakers.[64]

It can be argued that electroacoustic music is one of the few disciplines that attempts to fully explore the creative possibilities uniquely offered by the technologies discussed, and it is therefore something of a paradox that the majority of applications of these technologies cannot be considered 'electroacoustic.'

1.15. Summary

It has been a primary purpose of this chapter to provide a workable definition of 'electroacoustic music' on a mainly technological basis, and with as little discussion of aesthetic principles as possible. From a starting point at which 'loudspeakers' and 'fixed media' were cited as rather ambiguous defining characteristics (see sections 1.4 and 1.5), a clearer definition has been sought.

1.15.1. The Technological Demands of Electroacoustic Music

It can now be restated that electroacoustic music, at present, does indeed depend on performance via loudspeakers. Furthermore, because loudspeakers are effectively useless without feed signals, so electroacoustic music is equally dependent on the existence of encoded audio streams, as described in section 1.6. These two characteristics can be persistently observed in all works of electroacoustic music.[65]

---

[64] Wishart (1986). "Sound Symbols and Landscapes". In Emmerson [Ed.] *The Language of Electroacoustic Music*: 60.

[65] Although, as noted in section 1.8.3, there are certain works that could arguably be described as electroacoustic despite the fact that they do *not* require any loudspeakers or encoded audio streams in order to be performed. However, it can also be argued that loudspeakers and encoded audio streams have played an absolutely central role in the *compositional* process in this case. The works in question are strongly 'acousmatic' in essence and could not realistically have been composed without the use of loudspeakers and encoded audio streams as mediators of acousmatic sound:

Nance (2005). Personal communication with the author.

Encoded audio streams cannot exist without some kind of device to produce them, and therefore electroacoustic music must also be dependent upon other technologies. It is at this point, however, that the technological prerequisites become less unanimous between works. Nonetheless, the following *three primary technologies* have been identified (see sections 1.7 and 1.8):

- Audio encoding technologies
- Synthesis technologies
- Recording and playback technologies

Recording and playback technologies are only useful in the context of audio encoding and/or synthesis technologies, and all three of these technologies – if they are to be useful in a performance context – must be used in conjunction with:

- Audio decoding technologies

It can therefore be concluded, in addition to the prerequisite audio decoding technologies (in most cases, loudspeakers) and mediation via encoded audio streams, that any given piece of electroacoustic music will depend on at least one other of these three primary technologies for its realisation in performance. Two other technological categorisations, which are secondary in the performance context, have also been identified. These are:

- Audio processing technologies
- Software technologies

Audio processing technologies are secondary because they require at least one other of the three primary technologies in order to function usefully within a performance context, and therefore cannot be used on their own. In section 1.7.5 it was observed that software technologies, in many cases, are used as a surrogate for one or more of the three primary technologies in performance. In such cases, for the purposes of the present discussion, it is the primary technology that takes precedence.

Other applications of software technology (as a structuring tool, or as a means to converting physical phenomena into control data, for example) are really beyond the scope of this thesis, but have been noted.

It can now be concluded that, in technological terms, the performance of a piece of electroacoustic music can necessitate any combination (one or more) of the six technological 'units' given in Table 2 (page 23). These consist of the three primary technologies, plus each of these used in conjunction with audio processing technology, resulting in a total of sixty-three possible unique combinations. Of course, all of the technologies communicate via encoded audio streams and must therefore be used in conjunction with audio decoding technologies (loudspeakers).

1.15.2. Particular Approach to Audio Technologies in Electroacoustic Music

It should be recalled that the motivation behind such a system of technological classification is not so much to define electroacoustic music in its totality (indeed, this is not possible) as it is to ensure that all electroacoustic works can be adequately catered for in the performance context. As such, these technical requisites will become an important criterion in the evaluation of sound diffusion systems that will be undertaken in Chapters 4 and 5. In any case, it has been noted that the technologies themselves are not intrinsically 'electroacoustic' in nature: they have a wide variety of applications outside the field of what would normally be considered 'electroacoustic music.'

It can now be concluded that what differentiates electroacoustic music from other praxes that utilise the same technology, is the locus of the creative focus. In electroacoustic music, it is likely that at least one of the objects of creative exploration will be *directly attributable* to the unique characteristics or affordances of the creative frameworks (audio technologies) employed. (This does not, of course, mean that there will not be any other creative avenues explored.) In other applications, the object of creative exploration is more likely to exist somewhere outside the direct sphere of influence of the technology, which is therefore often used simply as a functional tool.

In sections 1.10 to 1.13, a specific framework – the loudspeaker – was used to exemplify this. As a functional tool or creative platform, the loudspeaker has certain peculiarities that need to be addressed in its operation, whatever the purpose. In most cases, the unique nature of the platform is, to a certain extent, disregarded and it is used in an essentially functional or conventional manner, as a means to realising objectives that could equally well be realised in a different way. Where this is the case, the loudspeaker is being used as a secondary, or mainly functional framework. In electroacoustic music, by contrast, there is a tendency to explore the possibilities that are *uniquely* offered by the platform, and therefore could *not* be realised in any other way: examples were given in section 1.13. It is proposed that the same model can be applied to other audio technologies.

It can therefore be stated that a purely technological definition of electroacoustic music that excludes *all* aesthetic considerations is impossible to attain. This is due in part to the fact that the frameworks used are not exclusive to electroacoustic music and therefore have a variety of non-electroacoustic applications. Consequently, part of what defines electroacoustic music is the unique way in which the frameworks are appropriated, and in this sense, aesthetics and technology are intrinsically bound. It remains useful, however, to be able to define the technological scope of electroacoustic music so as to ensure that works exerting various different technical demands can be accommodated side-by-side in live performance.

MISSING PAGE

2. Top-Down and Bottom-Up Approaches to Electroacoustic Music

2.1. Introduction

The previous chapter described electroacoustic music as an artistic practice whose creative driving force derives in part (and, in some cases, wholly) from the unique affordances of the frameworks (audio technologies) employed in its realisation. This is in contrast to many other applications of the same technology whose primary creative interests lie elsewhere, and whose use of the technology is essentially as a functional means to achieving these objectives. Consequently, it was concluded that a purely technological account of electroacoustic music, devoid of any aesthetic discussion whatsoever, would be difficult. Nonetheless, the main focus of the previous chapter was the technology itself. The purpose of the present chapter is to focus more directly on the *aesthetics* of electroacoustic music, with less emphasis on the specific roles of technology.

The context of this chapter is very broad, more concerned with large-scale 'overall philosophies' of electroacoustic music than with specific and individualised views (of which, of course, there are many). With this borne in mind, it will be proposed that electroacoustic music, in terms of how it is composed and performed, falls into two highly generalised categories, which will be referred to as 'top-down' and 'bottom-up,' respectively. On a superficial level it can be stated that these terminologies implicate differing ways in which musical structure and material can be devised but, as will become clear, the full extent and implications of the difference is much deeper. Specifically it will be proposed that these two opposing schools of thought have a fundamental impact on the *performance* of electroacoustic music via the process of sound diffusion: this will be discussed in Chapter 3.

Musique concrète and *elektronische Musik* represent an often-cited historical example of the proposed top-down/bottom-up binarism and are generally regarded as having been two concurrent but highly distinct musical praxes that, between them, eventually developed into what we now understand to be electroacoustic music. Bukvic suggests that these 'have long disappeared in their pure form,'[66] and this view seems to be increasingly accepted. He goes on to note, however, that 'for aesthetic, analytical and historical purposes,' they remain useful.[67] In the spirit of this observation, the present chapter introduces the notions of 'top-down' and 'bottom-up' by briefly accounting the differing attitudes, working procedures, advantages, and short-comings, of *musique concrète* and *elektronische Musik* as representative examples, further proposing that their aesthetic influences are still strongly in evidence.

Two further creative frameworks will be analysed: 'real' recorded sounds and 'artificial' synthesised sounds. These are, of course, the archetypal materials of choice of *musique concrète* and *elektronische Musik*, respectively. It will be observed that these frameworks, by their very nature, inherently lend themselves to the two differing attitudes and methods of working in question, and indeed that any creative framework can exhibit a 'directional bias' in favour of the bottom-up approach or the top-down. It will later be argued (Chapter 4) that sound diffusion systems – as creative frameworks themselves – are no different in this tendency, and can basically be divided into those that are essentially top-down and those that are essentially bottom-up.

The top-down/bottom-up dichotomy is by no means an entirely new concept, and has been described by various writers within and without the field of electroacoustic music. However it is comparatively rare, within the boundaries of electroacoustic music, for the binarism to be invoked in as general a sense as it is proposed here, individual writers tending, rather, to allude to it within fairly particular and specific contexts. An important function of this chapter, therefore, will be in citing a few choice incarnations of this fundamental duality, and presenting these as evidence of a more general underlying trend. The results of this survey will be collated and

---

[66] Bukvic (2002). "RTMix - Towards a Standardised Interactive Electroacoustic Art Performance Interface". *Organised Sound*, 7(3): 277.
[67] Ibid.

presented as a series of 'salient features' (Table 3, page 84). (It will be proposed, as a slight aside, that these abstract criteria also allow for useful applications of the top-down/bottom-up dichotomy, as an analytical tool, to be made outside of the confines of electroacoustic music). Having clearly defined the essential characteristics of both top-down and bottom-up approaches to electroacoustic music, some contemporary examples of each will be given.

In comparison with those aesthetic issues discussed in the previous chapter (concerning the appropriation of creative frameworks *et cetera*), it should be noted that the top-down/bottom-up distinction exists on a different level, hierarchically speaking. Put simply, any creative process can be undertaken in a top-down or bottom-up manner. To continue the example from the previous chapter, a musical exploration of the unique affordances of the loudspeaker, as creative framework, could equally be realised according to top-down or bottom-up criteria. It should also be noted that the terms 'top-down' and 'bottom-up' need not necessarily be mutually exclusive. While it will be proposed that, at a very general level, any given piece of electroacoustic music could broadly be defined as *either* 'essentially top-down' *or* 'essentially bottom-up,' closer inspection may well reveal a combination of both approaches across the various aspects of the composition's creation.

Having established a context within which creative frameworks can be regarded as either top-down or bottom-up, it will be concluded that the creative frameworks of electroacoustic music afford, perhaps for the first time, equal opportunities for both top-down and bottom-up approaches to the creative process.

2.2. Electroacoustic Aesthetics: A Case History

Electroacoustic music, in terms of its heritage, is often regarded as having been binary in nature. The praxes of *musique concrète* and *Elektronische Musik* represent the historical evidence of this. The former expression first appeared in 1948 and is used to describe the musical activity borne out of the music and theoretical writings of Pierre Schaeffer, at what is now known as the *Groupe de Recherches Musicales* (GRM) in Paris.[68] The latter describes its Germanic counterpart, which emerged at roughly the same at *Nordwestdeutscher Rundfunk* (NWDR) in Cologne, guided by the musical and intellectual activity of a group of individuals including composer Karlheinz Stockhausen.

On an extremely superficial level it can be stated that *musique concrète* consisted in music built out of 'real' recorded sounds, while *elektronische Musik* was entirely synthetic, constructed electronically in the studio.[69] In both cases it can be observed that the praxes embraced what were, at the time, entirely new creative frameworks, with their respective emphases on sound recording and sound synthesis. Harrison states that 'this is too simple a distinction, based on a reading of only the most obvious surface features'.[70] This is undoubtedly true but, notwithstanding its over-generalising nature, this particular distinction presents a valuable in-road to the more profound aesthetic differences associated with these two compositional approaches. For the purposes of elaboration, we must momentarily adhere to this simplified model, to which we will return after a brief account of *musique concrète* and *elektronische Musik.*

2.2.1. Musique Concrète

It is widely accepted that the theory and techniques of *musique concrète* were more or less single-handedly pioneered by the French electronic engineer Pierre Schaeffer. In Paris, in 1942, Schaeffer began a period of research into the acoustics of sound, and soon became interested in recording and playback technology (section 1.7.2), which at the time was starting to become more readily available.[71] Schaeffer was

[68] Emmerson (1986). "The Relation of Language to Materials". In Emmerson [Ed.] *The Language of Electroacoustic Music*: 18, 217.
[69] The verbs 'build' and 'construct' are deliberately contrasted here for reasons that will be clarified in section 2.5.1.
[70] Harrison (1999). "Diffusion - Theories and Practices, with Particular Reference to the BEAST System". *EContact!* 2(4). Electronic journal available at: http://cec.concordia.ca/econtact/Diffusion/Beast.htm.
[71] Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press): 19-20.

particularly interested in the use of recording technology to capture 'every day' sounds from the real world. This process is very much taken for granted nowadays, but until the advent of recording technology it simply was not possible to statically encode – and thereby 'fix' – a sonic event in such a way: sound was an exclusively transient phenomenon with no physical object representative outside the incessant passing of time.

When auditory events are recorded, a static encoding that abstractly symbolises the auditory events is created and this, of course, *is* represented by a physical object (an acetate disc, magnetic tape, or CD, for example). Schaeffer later proposed the expression 'sound object' (*objet sonore*[72]), reflecting this state of affairs; prior to the existence of recording and playback technologies it would not have made a great deal of sense to refer to a sound as an 'object.' Physical objects can, of course, be manipulated in ways that transient phenomena cannot. When sound is statically encoded onto analogue tape, for instance, it is possible to cut, splice, and rearrange the tape; it is possible to play the tape at different speeds; it is possible to play the tape backwards; it is possible to play only specific sections of the tape; it is possible to make tape loops by joining opposite ends of a section of tape; and so on. These possibilities are brought about by the physical nature of the recording medium itself, which, as it happens, is an agent for the abstract representation of auditory events. Manipulations to the recording medium will have 'knock-on' effects on the sound when it is eventually reproduced. Such practice is sometimes referred to as working 'directly' with sound, as opposed to working 'indirectly' via notational representations of pitches and durations, or with any other form of prescriptive scheme.[73] The implications of this process may

[72] Schaeffer (1966). *Traite des Objets Musicaux* (Paris; Seuil).

[73] It could, perhaps pedantically but ultimately reasonably, be suggested that this is *not* representative of working 'directly' with sound, but rather of working with an encoding of it, and that working 'hands on' with (e.g.) bows and strings involves more 'direct' contact with sound, as such. However, it cannot really be contested that Schaeffer's methods were *more* direct than notational frameworks, which really have nothing to do with *sound*

nowadays seem incredibly straightforward but, experienced for the first time, they must have been exciting.

In Schaeffer's early experiments (in which sounds were statically encoded onto acetate discs, which – like analogue tape – can be subject to certain physical manipulations), attacks were removed from bell sounds, thus fundamentally altering the perceived characteristics of the sound. This was achieved, very crudely by modern standards, by striking the bell and starting the recording shortly afterwards, when the bell sound had just begun to decay.[74] Experiments in which phonogram recordings were played back at different speeds also yielded changes in the qualities of the reproduced sounds. These procedures represent among the first instances of creative audio processing as an artistic practice in its own right.

Schaeffer also observed that repeated listening to a short recorded fragment (again, *only* possible with the use of recording technology) would reveal intrinsic qualities of the sound that would probably have gone unnoticed on a single hearing. This process might be tangentially likened to the experience of saying one's own name repeatedly until, eventually, the word seems bizarre. It was this observation that led Schaeffer to develop the concept of 'reduced listening' (*écoute réduite*[75]), whereby listening attention is focused on the intrinsic characteristics of sounds (on the particular qualities of sounds themselves), rather than on interpreting the sounds in terms of what might have caused them. An important part of Schaeffer's research was devoted to exploring the ways in which sound, thus defined as an autonomous phenomenon with intrinsically interesting characteristics, could be used as the basis for musical discourse. To this end he set about devising a system by which sound objects could be described in terms of their phenomenological characteristics, correctly observing that

---

whatsoever. It is in this sense that the expression 'working directly with sound' is to be understood.

[74] Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press): 20.

[75] Schaeffer (1966). *Traite des Objets Musicaux* (Paris; Seuil): 270-272.

western classical notation could only accurately describe duration and pitch, and even then only within the confines of metrical rhythm and the chromatic scale. The results were published in 1952 under the title 'Esquisse d'un Solfège Concret,' as part of a book entitled *À la Recherche d'une Musique Concrète*. The salient aspects of this publication are summarised, in English, by Manning.[76]

What is regarded by many as the first piece of electroacoustic music was composed during this period of research. Schaeffer's *Étude aux Chemins de Fer*,[77] completed in 1948, was fashioned from recordings of steam trains made at the *Gare des Batignolles* in Paris.[78] As a piece of *musique concrète* of the purest calibre, the intention was that the listener focus solely on the phenomenological characteristics of the sounds themselves, and the musical possibilities inherent in them. In practice, Schaeffer observed, it was difficult to deny a recognition of the 'train-ness' of this material; an important observation that foreshadowed years of subsequent research into empirical sound perception and psychoacoustics. Schaeffer's groundbreaking research has influenced (among many others) Denis Smalley, whose principles of 'spectromorphology' represent the aesthetic foundations on which much electroacoustic music has been, and continues to be, built.[79]

That Schaeffer's new musical practice was referred to as *musique concrète* is to do with the nature of its source material. In one sense, as previously stated, transient sounds could be 'halted in their tracks,' statically encoded onto a physical recording medium, and in this respect 'concretised.' Another (more important) sense in which 'concrete' is invoked is with reference to the *actuality* of the material. The

[76] Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press): 30-40.
[77] Various (2000). *OHM: The Early Gurus of Electronic Music.* (Compact Disc - Ellipsis Arts: CD3670).
[78] Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press): 20.
[79] Smalley (1986). "Spectro-Morphology and Structuring Processes". In Emmerson [Ed.] *The Language of Electroacoustic Music*: 61-93.
Smalley (1997). "Spectromorphology - Explaining Sound Shapes". *Organised Sound*, 2(2): 107-126.

characteristics of the sound have not, themselves, been 'composed'; they are 'already there,' and in that sense they are 'concrete'; they are 'given.' Accordingly, and as Harrison observes, this is reflected in a 'concrete' method of working:

> A further dimension of what was 'concrete' about *musique concrète* was the *method* of working and, by extension, the relationship between composer and material: as in sculpture or painting where the artist produces the finished product on or in a fixed medium by manipulating the materials (paint, wood, stone) directly, so in *musique concrète* the composer is working *directly with sound.* [Harrison's italics].[80]

An important implication here is that the composer has *no choice* but to 'work with' the particular qualities of the chosen material, just as a sculptor must work with the particular characteristics of, say, granite. It is not normally possible to alter the fundamental characteristics of granite. One can sculpt it into a variety of different shapes and textures, but the essential character of the source material remains the same. Of course, the artist is at liberty to choose another material if necessary (sandstone offers different possibilities to granite) and combine materials as appropriate, but ultimately the particular qualities of each material must be considered of paramount importance. Such is the doctrine of *musique concrète*, and of much subsequent and contemporary electroacoustic music derived from its principles.

It will be noted, of course, that Schaeffer's manipulations of sound objects took place *within the abstraction layer* that is brought about by audio encoding (see section 1.6.1 and Figure 1 on page 11), that is, at a point in between the encoding of audio streams and their subsequent decoding. As such, these procedures demonstrate a specific exploration of the creative possibilities *uniquely* afforded by new audio technologies, and it can be demonstrated that the aesthetic of *musique concrète* arose as a direct consequence of these explorations. This is, of course, completely in keeping with the definition of 'electroacoustic music' given in the previous chapter.

---

[80] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 117.

2.2.2. Elektronische Musik

While *musique concrète* can be regarded as having been an entirely new model of musical thinking, a more or less complete departure from what was previously considered (perceptually and syntactically) to be 'musical,' arising from the innovation of one pioneering individual, *elektronische Musik* is best described as a continuation of certain aspects of established musical practice, brought about by a group of like-minded individuals.[81] Nonetheless, it will be observed in due course that *elektronische Musik* is similar to *musique concrète* insofar as both pursue creative interests brought about specifically by the new audio technologies.

The process of setting up a dedicated studio for the production of *elektronische Musik* began in October 1951 after a meeting between Fritz Enkel, technical director of North-West German Radio, and a group of interested parties including Robert Beyer (also of NWDR), Werner Meyer-Eppler (head of the Phonetics Department at Bonn University), and composer Herbert Eimert, with the first compositions appearing while the studio was still under construction.[82] Work on the studio was finally completed towards the end of 1953, by which time several other compositions had been completed, and Stockhausen had begun work on his two *Studies*.

The compositional aesthetic of *elektronische Musik* centred on the structured organisation of, and *ergo* complete control over, all aspects of sonic material. In this respect it can be regarded as a development of serialism. The process of serial composition can essentially be regarded as one of compartmentalisation, or parameterisation. Firstly, compositional material must be abstractly described in terms of a number of independent constituents which, in combination, express its totality. For instance, sound might thus be described in terms of pitch,

[81] Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press): 43ff.
[82] Ibid.43-45.

dynamic level, duration, and timbre, although many parameters, some more abstract than these, are possible. In this way, individual constituent attributes of sound are defined and therefore compartmentalised. Each of these independent parameters is then further defined in terms of a number of discrete, perceptually different, states. This procedure is described by Stockhausen with reference to one of the earliest pieces of *elektronische Musik*, *Studie I*, which was completed in 1953:

> A 'serial system' for sensorially evaluated frequency differences will begin in the middle of the auditory range and extent to the limits of pitch audibility.[83]

Here, Stockhausen is isolating (compartmentalising) the parameter of 'pitch' (frequency) by specifying a point of reference from which differences can be measured. A less esoteric example is the chromatic scale which, conveniently for the early serialists, expresses the infinitely variable parameter of 'pitch' in twelve differentiated sub-divisions. Of course we know that between 'C' and 'C-sharp' – notionally the 'smallest possible' subdivision of the chromatic scale – there is in reality an infinite number of possible intermediate pitches, but this presents organisational and representational difficulties, for how is one to positively identify one member of such (to use a mathematical expression) an nondenumerably infinite group?[84]

Denumerably infinite (or 'countably infinite') groups, such as the chromatic scale, are more easily manageable from a serialist perspective: we know that pitches theoretically carry on ascending and

[83] Stockhausen (1991). *Stockhausen 3 - Elektronische Musik 1952-1960.* (Compact Disc - Stockhausen-Verlag: 3). Quotation in accompanying booklet: 102.
[84] The concept of infinity implies an unlimited quantity, or a lack of definable boundaries. Within this context, however, it is possible for a measurement to be either 'denumerably infinite' or 'nondenumerably infinite.' For the former case, integer numbers represent a good example: we know that there is an infinite quantity of them, but it is hypothetically possible to positively identify every single one of them by counting (even though this would take an infinitely long time!). Floating point (or fractional) numbers, on the other hand, represent members of a nondenumerably infinite group. While we are able to enumerate the quantity of integer numbers that exist, for instance, between zero and ten, we are unable to evaluate the quantity of floating point numbers that exists between these boundaries.

descending *ad infinitum*, but at least we could hypothetically give a name (or number) to any member of that infinite group. Abstractly, in order for parameters to be sub-divided in this way, a constant unit of measurement that arbitrarily describes the smallest interval between two states must be defined for each parameter. This results in what Wishart describes as a 'lattice' of discrete values[85] and therefore renders the number of possible values for that parameter denumerably infinite. The composer may wish also to define upper and lower thresholds – for example the extents of the human hearing range, to reiterate the Stockhausen example given above.

Having defined a parametric lattice, the parameter in question can become subject to mathematical operations: relationships *between* latticed parameters can be easily defined; a succession of pitches can be directly mapped on to a succession of, say, amplitudes. Such a procedure would be absurd – for what exactly *is* the concrete relationship between pitches and amplitudes? – if not for the fact that each parameter is, effectively, represented numerically. In other words, once compartmentalised parameters have been divided into ordered sets of differentiated values, the development of each parameter throughout the course of the music can be *deterministically* calculated according to predefined rules. Stockhausen continues:

> The duration of each note will be inversely proportional to its thus-defined frequency difference, so that as the distance from the middle frequency range increases, the duration decreases. The amplitude series is to decrease proportionally to duration, as the frequency difference increases. Thus the tendencies away from the middle register towards the lower and upper limits of audibility will be perceptible from the correspondingly decreasing duration and amplitude of the notes.[86]

Here, Stockhausen has defined a mathematical relationship (specifically, inverse proportionality) between frequency interval and duration, and, in turn, between duration and amplitude. In this context, this is a *compositional* act. Such a relationship would be difficult to

[85] Wishart (1996). *On Sonic Art* (Amsterdam; Harwood Academic): 23-30.
[86] Stockhausen (1991). *Stockhausen 3 - Elektronische Musik 1952-1960.* (Compact Disc - Stockhausen-Verlag: 3). Quotation in accompanying booklet: 102.

justify, however, in the context of *musique concrète*, because there would be no such system of enumerated frequencies, durations, and amplitudes. The *elektronische* composer is not, therefore, composing directly with the sound itself (as is the case in *musique concrète*), so much as with abstract mathematical constructs that happen to then be applied to compartmentalised sonic or structural parameters. Similar techniques have been documented with respect to other works composed in the formative years of *elektronische Musik*:

> During the first half of 1953 Beyer and Eimert composed *Ostinate Figuren und Rhythmen*, and Eimert alone composed *Struktur 8*. These pieces are characterised by the strict application of serial procedures to the processes of tone selection and processing. *Struktur 8*, for example, is derived entirely from a restricted set of eight intervallically related tones.[87]

It was previously suggested that the composer of *musique concrète* 'works with' the pre-existing characteristics of 'real' sound materials, and that musical structure is therefore born out of these intrinsic characteristics. In the case of *elektronische Musik*, the composer designs the *musical structure*, in this respect, directly, and the intrinsic characteristics of the resulting sounds are born out of the predetermined relationships between their constituent parameters. The 'directionally opposite' nature of these compositional approaches is neatly encapsulated by Francis Dhomont:

> [In *musique concrète* the] compositional method begins with the concrete (pure sound matter) and proceeds towards the abstract (musical structures) – hence the name *musique concrète* – in reverse of what takes place in instrumental writing, where one starts with concepts (abstract) and ends with a performance (concrete).[88]

The strong relationship between *elektronische Musik* and established traditional compositional practice (most obviously, serialism) has already been observed, and it can therefore be assumed that within 'instrumental writing' Dhomont would also include the compositional procedures of *elektronische Musik*.

---

[87] Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press): 45.
[88] Dhomont (1995). "Acousmatic Update". *EContact!* 8(2). Electronic journal available at: http://cec.concordia.ca/contact/contact82Dhom.html.

2.3. The Implications of 'Real' and 'Artificial' Sounds

As previously observed, it can essentially be stated that *musique concrète* was concerned with 'real' recorded sounds, while *elektronische Musik* was concerned with 'artificially' synthesised sounds. The reasons behind this generalisation may already be clear, but as a step towards further clarifying some of the essential aesthetic differences that co-exist within the broad practice of electroacoustic music, it is helpful to examine the implications of 'real' sounds and 'artificial' sounds – as two distinct creative frameworks – more closely.

'Real' sound occurs naturally, in the real world: it develops according to physical laws governing the interactions between the various sounding bodies that give rise to it. Here, there is no 'abstraction layer,' and the processes of sound creation and sound reception are (to all intents and purposes) intrinsically bound, as illustrated in Figure 1(a) on page 11. Therefore, there are certain constraints with regard to the kinds of control a human agent may exert on it. The nature of the physical interactions themselves – when a violinist tilts the bow away from the bridge in order to play more quietly, for example – may be contrived, but ultimately the particular character of the resulting sound is the product of natural, physical, interactions. In this respect, the way in which a 'real' sound develops through time can be described as organic process.

With synthesised, or 'artificial' sound, exactly the opposite is true: a high degree of precision *can* be directly exerted on the intrinsic characteristics of a sound. Assuming that the composer is familiar with the techniques of the various synthesis algorithms, then he or she notionally has absolute control over the spectral profile of a sound at any given moment throughout the course of its development. However, there are no physical interactions that, directly and by virtue of the laws of physics, result in the sounding phenomenon.[89] In this respect, sound synthesis presents a direct point of

[89] Even in the case of real-time synthesis (which of course was simply not available during the historical period in question), we are not dealing with 'physical interactions' so much as 'data entry'. Although the ultimate goal is, perhaps, to give the *impression* of direct control

entry into the abstract domain of encoded audio streams, offering up the possibility of creating sound – literally – from nothing. The way in which an 'artificial' sound develops through time is governed by a concept – an algorithm – which, as Wishart puts it, has 'been generated by an electrical procedure set up in an entirely cerebral manner.'[90] The process that gives rise to synthetic sound, as the result of predetermined abstract construction, cannot therefore be described as 'organic' and would be better described as 'architectonic.'[91]

It can therefore be said that 'real' and 'artificial' sounds differ primarily – or, more precisely, are *opposite* – in terms of the ways in which their intrinsic sounding characteristics can be accessed. In the case of 'real' sounds, the nature of the sound arises from a combination of physical gesture and the physical attributes of the sounding bodies in question. With 'artificial' sounds, however, there is no physical gesture as such (although many would argue that there is an *implied* gesture, or gestural 'surrogacy'[92]), nor sounding bodies in the same sense, and the sounding characteristics (or a higher-level algorithm that will, in turn, generate the sounding characteristics) are determined directly by the composer. This reading is directly analogous to the respective working methods of *musique concrète* and *elektronische Musik.*

It can be seen that the process of synthesising an 'artificial' sound, generally speaking, requires work at a 'low' level: it is necessary to define the ways in which basic fundamental parameters interact, and in this way work *up* to a

(think of a synthesiser controlled by a MIDI keyboard or other such gestural controller) we are effectively dealing with specific ways of providing an intermediate synthesis algorithm with the correct variables. If the algorithm and data entry method are good, we may well be convinced that we are interacting with physical sounding bodies, but ultimately changes to the synthesis algorithm will have a far more radical effect on the resulting sound than changes to the input data, and there is no way we can provide the algorithm with any information that it has not been designed to receive. It is therefore true to say that the resulting sound is essentially governed by the synthesis algorithm, and not by the physical gestures.

[90] Wishart (1986). "Sound Symbols and Landscapes". In Emmerson [Ed.] *The Language of Electroacoustic Music*: 58.

[91] We will return to the expressions 'organic' and 'architectonic' – originally proposed by Harrison – in section 2.5.3.

[92] Smalley (1996). "The Listening Imagination - Listening in the Electroacoustic Era". *Contemporary Music Review*, **13**(2): 77-107.

'finished' sound. In generating 'real' sounds, the laws that govern the evolution of the sound's characteristics are already *physically* in place, and the performer essentially provides these sophisticated natural 'algorithms' with *high-level* parameters. This reading can, again, be regarded as a microcosm of the proposed binary model of composition as a process whose various aspects can be located at any point on a continuum that extends from 'top-down' to 'bottom-up.' It can therefore also be said that 'real' sound – as a creative framework – is inherently biased towards top-down methodologies, whereas 'artificial' sound is biased in favour of a bottom-up approach. Such attributes can be regarded as intrinsic characteristics of the frameworks themselves, and it is proposed that similar biases are potentially to be found in all creative frameworks.

2.4. Musique Concrète versus Elektronische Musik

Having briefly examined the materials and working practices of *musique concrète* and *elektronische Musik* it may be restated that the two are, in many respects, opposite. It is worth examining these differences further, as it will later become clear that, although *musique concrète* and *elektronische Musik* are often regarded as 'extinct,' the underlying aesthetic discrepancies remain in many instances central to contemporary electroacoustic musical discourse and performance practice.

As already noted:

> The Germans held the work of the Second Viennese School in high esteem, and many became avowed disciples of the cause of furthering the principles of serialism. An increasing desire to exercise control over every aspect of musical composition led to a keen interest in the possibilities of electronic synthesis, for such a domain *eliminated not only the intermediate processes of performance but also the need to accept the innate characteristics of natural sound sources...* This [was a] movement towards total determinism. [My italics.][93]

It is therefore clear that the aesthetic standpoint of the Cologne school not only represented a *different* approach to that of its Parisian counterpart, but, by definition, sought actively to exclude it. It is logical to deduce that, in an art-form that relies so heavily on 'total determinism,' not only is it necessary

[93] Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press): 46.

to exercise a high degree of control over one's source materials – hence the preference for 'artificial' synthesis over recordings of 'real' sounds – but it is also important that *the precision of the compositional process be accurately recreated in performance.* After all, there is little point in defining intricate relationships between musical parameters if these relationships will be lost or inadequately represented in performance. It is for this reason that the Cologne school found human performers, in some cases, to be inadequate for the performance of their music. The practice of *elektronische Musik*, therefore, only becomes possible as a direct consequence of the new creative opportunities offered up by synthesis technology, and can therefore be regarded as an 'electroacoustic music' practice according to the definition proposed in Chapter 1. Harrison makes the following observation, with reference to certain aspects of Stockhausen's instrumental work:

> One of the primary reasons for the emergence of *elektronische Musik* was the need to be able to realise with absolute precision in the studio the kind of serialised dynamics presumably vital to the structure of works like Stockhausen's *Klavierstuck I.* This piece, famously, features a simultaneously struck nine-note chord containing five different dynamic levels – a fairly unrealistic demand on any pianist; but if it cannot be accurately performed, the work becomes, in a very real sense, unintelligible, as the measurements between the five dynamics cannot be made aurally (perceptually).[94]

This, of course, supports the suggestion that the degree of compositional accuracy demanded by certain composers is simply unattainable by non-synthetic means. Generally speaking, it is true to say that works composed in an '*elektronische*' manner are designed to be heard *exactly* as the composer intended them, and that the Cologne school regarded the new creative frameworks, and synthesis in particular, as an unprecedented means of achieving this. 'Real' sounds – whether recorded onto a fixed medium or otherwise – were found to be inappropriate for this way of working, as they did not afford the composer a sufficient level of control over their intrinsic characteristics. Stockhausen evidently discovered this in 1952 whilst working – at the hospitality of a certain Pierre Schaeffer – on his *Étude*, whose compositional procedure he describes as follows:

---

[94] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 119.

> First, I recorded six sounds of variously prepared low piano strings struck with an iron beater (tape speed: 76.2 cm per second)... With scissors [I] cut off the attack of each sound. A few centimetres of the continuation, which was – briefly – quite steady dynamically, were used. Several of these pieces were spliced together to form a tape loop, which was then transposed to certain pitches using a transposition machine... I then chose, according to my score, one of the tapes having a certain sound transposition, measured the notated length in centimetres and millimetres, [and] cut off that length... Next, I chose another prepared tape, measured and cut off a piece, and spliced it onto the previous piece. Whenever the score prescribed a pause, I spliced a corresponding length of white tape onto the result tape... Upon hearing two synchronized layers, and even more so hearing three or four layers, I became increasingly pale and helpless: I had imagined something completely different! [...] Anyway – on this CD released in 1992 – the world can now hear my *concrète Étude* of 1952, which for many years I had presumed lost until I finally found it again in a pile of old tapes.[95]

This quotation is interesting in several aspects. Firstly (and as a slight aside), in describing the work in question as 'my *concrète Étude*,' Stockhausen appears to suggest that any composition using 'real' recorded sounds as its source material qualifies as a piece of *musique concrète* purely on that basis. Nowadays (although this misconception is still frequently iterated) it is widely acknowledged that the situation is more concerned with the method of working as opposed to the materials, but it is worth considering that in 1952 research into such aesthetic matters was at a very early stage of its development. It seems likely that many of the aesthetic differences associated with *elektronische Musik* and *musique concrète* are the result of years of *subsequent* research, and therefore care must be taken not to invoke this knowledge anachronistically.[96] It is, in part, for this reason that alternative (and in a sense more generic) terminologies will be proposed in section 2.5, and it may already be clear that, despite his use of an inherently top-down framework ('real' sound sources), in composing his '*concrète Étude*,' Stockhausen was actually working in a characteristically 'bottom-up' manner. Secondly, it is clear that Stockhausen, in constructing this piece, was working from a predetermined score (a facsimile thereof is provided in the source previously cited[97]) and therefore had in mind a

[95] Stockhausen (1991). *Stockhausen 3 - Elektronische Musik 1952-1960.* (Compact Disc - Stockhausen-Verlag: 3). Quotation in accompanying booklet: 95-97.
[96] Indeed, much of Schaeffer's work seems to have been directed toward the development of a systematic categorisation (or *solfège*) of 'real' sounds, a practice that, paradoxically, might nowadays be regarded as more characteristic of the Cologne school.
[97] Stockhausen (1991). *Stockhausen 3 - Elektronische Musik 1952-1960.* (Compact Disc - Stockhausen-Verlag: 3): 97-100.

precise structure that he wished to articulate with the sonic material. Note also the sense of 'scientific precision' embedded in the language that Stockhausen uses, for example in stating (perhaps superfluously?), 'Tape speed: 76.2 cm per second.' Thirdly, it is evident that the use of 'real' sounds for this purpose was, to a certain extent, unsuccessful, as Stockhausen acknowledges that the results were not as he had anticipated. This might be regarded as material evidence of the suggestion that 'real' and 'artificial' sounds, as a direct result of the processes necessary to facilitate their existence, inherently lend themselves to differing compositional approaches.

Arguments in favour of the '*elektronische*' method of composition often centre on claims to 'objective truth,' and composers accordingly seek to engage structures that can be demonstrated to transcend the subjective interpretation of the individual, that is, things which are irrefutably 'true.' This position was articulated by Eimert who, in the formative years of *elektronische Musik*, cited 'scientific fact' as a justification for the aesthetics associated with the Cologne school:

> In electronic serial music... everything to the last element of the single note is subjected to serial permutation... Examination of the material invariably leads one to serially ordered composition. No choice exists but the ordering of sine tones within a note, and this cannot be done without the triple unit of the note. A note may be said to 'exist' where elements of time, pitch, and intensity meet; the fundamental process repeats itself at every level of the serial network which organizes the other partials related to it... Today the physical magnification of a sound is known, quite apart from any musical, expressionist psychology, as exact scientific data.[98]

To paraphrase: it can be scientifically proven that the physical nature of sounds demonstrates finite, indisputable, relationships between certain fundamental parameters (time, pitch, intensity). The serial procedures by which composers organise musical materials are based on these scientifically observed relationships and therefore 'cannot be wrong.' Such claims relate to the innate need for (perhaps all) composers to differentiate their work from the purely arbitrary, and rigorous structuring principles such as serialism represent an effective means of achieving this goal (despite

---

[98] Eimert (1955). Cited in Manning (1993). *Electronic & Computer Music* (Oxford; Oxford University Press): 46-47.

those that might argue that the end results are musically unsatisfying). Ultimately, the use of objective (or perhaps 'super-human') structuring principles is, quite simply, one way to justify compositional decisions: the composition is 'good' because it is demonstrably based on 'the truth.'

This method of working, however, also has its drawbacks. It forces the composer into working with 'known quantities,' which can have a limiting effect on the results attainable. As a simple example, working strictly within the twelve semitones defined in the chromatic scale, it would be impossible to realise the clarinet glissando at the start of Gershwin's *Rhapsody in Blue* as it is usually performed. This manner of composition also burdens the composer with the responsibility of determining every last aspect of the sonic development. A laborious task indeed, and one which reportedly often yielded musically unsatisfying results. It can be argued that an appreciation of the works cited previously depends on the listener's ability to apprehend the 'super-musical' structures articulated by the sound material. If this does not happen, then the listener is left in a state of confusion:

> There appears to be a considerable discrepancy between postulation [what is composed] and reception [what is perceived by the listener], a discrepancy which must be of the very nature of the new art form... in that nothing pertaining to electronic music is analogous to any natural existent phenomenon of traditional music, associations have to be evoked from elsewhere. Instead of being integrated, they remain an ever increasing conglomeration of mentally indigestible matter.[99]

A 'discrepancy between postulation and reception' implies that structures *conceptualised* by the composer are not always reflected in what the audience *perceives*.[100] This indicates a mismatch between objective schema and the subjective evaluation of the resulting musical (sonic) phenomena. In this context, the following quotations are of interest:

> No evaluation of the musicality of sounds can be made on the basis of [...] its spectrum. A sound may possess a haphazard spectrum lacking in meaningful information as to how its component frequencies are ordered: this does not necessarily mean that its presence in a composition is not the result, distinct and

99 Stuckenschmidt (1955). Cited in Ibid.: 47.

100 There is also, of course, a sense in which the quotation implies a cognitive difference between 'new' (electroacoustic) music and traditional instrumental music.

removed from any physical reality, of sophisticated and highly meaningful musical construction.[101]

> A spectrogram is a type of literal spectral analysis at a chosen visual resolution: at too high a resolution detail becomes lost in a blur; at too low a resolution there is insufficient detail. But a sonogram is not a representation of the music as perceived by a human ear – in a sense it is too objective.[102]

Here, both Berenguer and Smalley suggest that *objective* (or 'measurable') and *subjective* (perceptual) aspects of sound are fundamentally different in nature. A sound that is 'well organised' in terms of the parameters that represent it, does not necessarily equate to a sound that is pleasing to the ear. Nor will the 'well organised' nature of the sound's constituent parameters necessarily be perceived by listeners ignorant of the processes at work. To put it another way, the arbitrary superimposition of a 'lattice' of discrete values onto continuously variable and interrelated parameters is exactly that: arbitrary. In most cases it serves a functional, and often visual, purpose – to observe the frequency content of a sound recording or visualise the contour of a melody line for example – but it is erroneous to assume that systematic structuring of such representative 'secondary' data will be reflected in a 'well organised' sound on resynthesis. In short, a piece of music that is based too heavily on 'objective' structuring principles will not necessarily be satisfying subjectively. On the other hand, there are those who would argue that a piece of music devoid of such objective structuring principles is not a 'composition' so much as an arbitrarily and whimsically assembled collection of unrelated noises.

It can be seen that the principles of *musique concrète* and *elektronische Musik* differed considerably on a very primary, and potentially irreconcilable, level. While devotees of *musique concrète* were committed to finding a place in musical discourse for the innate characteristics of natural sound sources, practitioners of *elektronische Musik* actively sought to escape the imprecise nature of 'real' sounds in favour of synthesis techniques more conducive to their agenda. As outlined in the preceding

---

[101] Berenguer (1998). "Music-space". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 220.
[102] Smalley (1997). "Spectromorphology - Explaining Sound Shapes". *Organised Sound*, 2(2): 108.

sections, what might at first appear to have been a trivial and inconsequential difference increasingly seems to suggest the existence of two deeply contrasting world-views (there is no reason, after all, to suppose that these differences manifest themselves only in matters of musical composition) whose *a priori* assumptions are in many ways diametrically opposed. Indeed this fundamental difference is embodied, to a certain extent, in the very nature of 'real' and 'artificial' sounds, the archetypal materials of choice for composers of *musique concrète* and *elektronische Musik*, respectively.

Despite the undeniable – and, in many respects, fundamental – differences between *musique concrète* and *elektronische Musik*, there is an important sense in which they were similar: both sought explicitly to explore the creative opportunities uniquely afforded by new audio technologies (creative frameworks). Both praxes existed *within* the abstraction layer brought about by the ability to encode auditory events into static and transitory streams, as illustrated in Figure 1 (page 11), and as such, could not have been realised by any other means. Additionally, these praxes did not make use of the new technologies in a merely functional manner: both appropriated the frameworks as a means to exploring previously inaccessible and uncharted musical territories. *Elektronische Musik* sought to construct relationships between the objective parameters of sound: this, of course, only became possible with the advent synthesis technology. *Musique concrète* adopted a more perceptual approach that, nonetheless, could not have been achieved in the absence of sound objects statically encoded onto physical recording media. It can therefore be stated that both *musique concrète* and *elektronische Musik* fall into the broader category of what is now known as 'electroacoustic music,' as defined in the previous chapter.

2.5. 'Top-Down' and 'Bottom-Up' Compositional Models

At this point it is appropriate to explain the very broad notion of 'top-down' and 'bottom-up' models of musical thought in more detail. It is hoped that these terminologies, as the respective poles of a hypothetical continuum, may prove useful as a means of evaluating the essential nature of a wide variety of art-music praxes, whether electroacoustic or otherwise. The situation of musical attitudes within this continuum will be referred to as 'aesthetic directionality.' Taking what has already been discussed as a starting point, it can be said that *musique concrète*, and those schools of compositional thought regarded as having inherited from it, represent top-down approaches to the creative process, while *elektronische Musik* and its subsequent counterparts exist closer to the bottom-up end of the spectrum.

Musique concrète, recalling our simplistic definition, essentially takes naturally occurring sound as its source material; sound objects that might be described as 'individual elements, stamped with [their] own gravitation values, possessing [their] own internal atomic cohesion.'[103] The compositional process, in this case, takes this *given* 'atomic cohesion' and builds 'down' from it until a musical structure, and eventually a finished composition, is realised. In contrast *elektronische Musik*, with its synthetic means, poses the problem of the (literally) blank canvas. Faced with this problem, composers are required to construct their own 'atomic cohesion' from the bottom up, and it is to this end that much (if not all) of the compositional energy is directed.

As discussed in section 2.1, the distinction between top-down and bottom-up modes of musical thought is not a new concept, in abstract terms. The following sections seek to collate some of the various literary invocations of dualisms that, it is proposed, are collectively indicative of a single underlying dichotomy in the context of electroacoustic music.

2.5.1. Invention versus Creation; Building versus Construction

Savouret eloquently articulates the top-down/bottom-up dichotomy by invoking various binarisms:

[103] Clozier (1998). "Composition, Diffusion and Interpretation in Electroacoustic Music". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 236.

> Creation implies that *something is made from nothing*, that starting from nothing one arrives at something [recall the 'blank canvas' mentioned above]. The creator-composer brings about a musical construction, which presupposes that he is working by self-imposed rules [...] from material which nowadays can be created [...] by means of synthesis. Construction is a project. Conception and realization are distinct, finalized phases, pointing towards a final object yet to come... Invention (*invenire* in Latin) means to come in, to find, *to dis-cover something that already exists.* Georges Braque distinguishes constructors, those who fill a frame, from builders like Cézanne. The inventor-composer picks up something on the way, something that is already there, *he enters into a relationship with an existing entity...* Conception and realization work intimately together in a perceptual bouncing back and forth in and on a chosen terrain... In such a compositional intention there is no object lying somewhere in the future, [...] *there is no objective...* In creation-composition, it is principally studio time that is required, whereas in the case of invention-composition it is principally community time that has to be managed. [My italics.][104]

He goes on to propose the analogy of 'figured bass' versus 'melody harmonisation':

> Figured bass requires the student to construct music upon the [...] lower voice by adding three upper voices. Thus *from bottom to top*, the student must fill in an empty space starting from those few sparse bass notes whose expressive potential is often close to zero. Everything so to speak remains to be done from scratch. The bass offers very little resistance, the creator-to-be can create to his heart's content... Harmonization is a very different matter: the melody is there, and one cannot escape the upper voice... Be it good, bad or indifferent, there is a style in the capricious succession of notes from on high, very different to the inept mumblings of the figured bass. In the melody, the way the notes follow upon each other already suggest a certain possible harmony. The student, *working from the top downwards*, is going to unmask what is already there, he doesn't have to fill in the near-void of the preceding exercise, but rather bring out, accentuate the shape, add supplementary notes as proof of what is already present in the song. [My italics.][105]

In these quotations Savouret invokes a subtle distinction between 'building' (which is associated with the top-down approach) and 'construction' (bottom-up; where once there was nothing). Note that in the previous quotation, one method works 'from bottom to top,' while the other proceeds 'from the top downwards.' Savouret refers to the top-down '*concrète*' composer as 'inventor-composer,' and the bottom-up '*elektronische*' composer as 'creator-composer.'

---

[104] Savouret (1998). "The Natures of Diffusion". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 348-349.
[105] Ibid.: 349.

2.5.2. Abstracted Syntax versus Abstract Syntax

Emmerson uses the expressions 'abstract' and 'abstracted' to indicate the differing processes by which musical dialogue is obtained in bottom-up and top-down methods, respectively. As previously discussed, the bottom-up composer begins with the proverbial 'blank canvas,' and proceeds to define abstract principles, or algorithms, abstract insofar as they precede the actuality of the sounds themselves and, in and of themselves, have nothing to do with sound *per se*. These abstract principles are subsequently used to define certain aspects of the musical dialogue, from the 'micro' level, concerned with the intrinsic characteristics of individual sounds, through to the 'macro' level, which defines the structure of the music at a higher level. The top-down composer, on the other hand, takes the very actuality of sounds themselves *as* the starting point, performs a subjective evaluation of their spectromorphologies, cultural implications, and so on, and uses this as the basis for a musical dialogue. In this case the musical dialogue has been abstract*ed* from the observed characteristics of *already existing* phenomena. In this respect the composer is, in a sense, uncovering, or 'dis-covering' (recalling Savouret's words), that which is already there. This is not composition 'from nothing' but, quite simply, composition 'from something,' with 'something' being concrete sounds and their perceived qualities. Emmerson states that the use of abstract and abstracted syntax as compositional tools are not mutually exclusive but, rather, 'are arbitrary subdivisions of a continuous plane of possibilities, the outermost extremes of which are ideal states which are probably unobtainable.'[106] As described previously, it is proposed that top-down and bottom-up ideologies demarcate a continuum, and this notion is supported here by Emmerson.

---

[106] Emmerson (1986). "The Relation of Language to Materials". In Emmerson [Ed.] *The Language of Electroacoustic Music*: 25.

2.5.3. Organic versus Architectonic

Harrison differentiates top-down and bottom-up approaches to the composition and performance of electroacoustic music by way of the terms 'organic' and 'architectonic,' respectively, using Stockhausen's *Kontakte* as an example of how these two paradigms overlap and interact throughout the composition/performance/reception process and in doing so reinforcing much of what has already been discussed:

> The apparent need for 'objective justification' of musical utterance, exemplified by the threads of analysis and 'measurement,' is one of the central creeds of Western art music... The high modernist agenda of serialism (of which *elektronische Musik* was, interestingly, a part) was heir to this tradition and continued the prevailing view that the 'text' of the score, amenable to 'out of time' analysis, was the 'true' representation of the composer's thoughts because it allowed for more accurate measurement of the distances between musical events... Musical events have no intrinsic interest; they exist primarily to articulate the distance between them, on the measurement of which rests the notion of 'structure.' This seems to be evidence of what I call 'architectonic structure' and is diametrically opposed to the 'organic structure' generated by the materials and compositional strategies of *musique concrète*... Despite the rigour and complexity of its concept, *Kontakte* was evidently assembled by ear, Stockhausen making countless experiments in the studio, testing the appropriateness of each 'moment,' modifying his intentions in the light of what he heard and selecting only those sonic results which worked *perceptually* [my italics] in a structure which *evolved* into its present form during the process of composition, rather than being preplanned... Not withstanding its impeccable elektronische Musik credentials in its synthesis method, I would argue that *Kontakte* can therefore be considered a classic piece of *musique concrète*... How was it that such an (allegedly) concept-oriented composer could be satisfied with an (apparently) arbitrary process of structuring a work, and: how do we reconcile the original compositional intent (concept, poiesis) with what we hear when we listen to the actual work (percept, esthesis)? There is a strong implication ([...] still to this day underpinning the very basis of much computer music and algorithmic composition) that if the conceptual backdrop is sufficiently strong, then a good piece is virtually guaranteed. Yet this is contradicted by *Kontakte*... This is [...] indicative of the (at least equal) importance of percept alongside concept in composition.[107]

In thus describing the distinction between 'organic' and 'architectonic' structuring processes, Harrison clarifies another aspect of 'what is top-down' about the top-down approach to composition. One might regard one's empirical (emotional or 'involuntary') response to an auditory stimulus as 'given,' insofar as the response has not been *preconceived*;

---

[107] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 120.

it has occurred just as naturally as the characteristics of the sound that evoked it. In the top-down model, this response does not require any objective justification: it is irrefutably 'true' on the basis that it *was* experienced; this *is* the justification. We might refer to this as the experience of 'subjective truth' in a musical context (as opposed to the more familiar expression 'objective truth'), and it is this process that Harrison cites as having been an important aspect in the composition of Stockhausen's *Kontakte.* When musical material is assembled 'by ear' in this way, what is 'true' (or, one might say, what is 'right') is determined *perceptually* rather than *conceptually*, or to put it another way, *subjectively* rather than *objectively*. That is, the subjective response to compositional materials – as 'given' rather than 'preconceived' – becomes the 'top' (the 'actuality') from which musical structure can be built 'downwards.' Harrison refers to this kind of structuring process as 'organic.' Antithetically, if musical materials are organised 'from the bottom up' according to an abstract preconceived structure, then subjective responses are secondary, a mere happenstance of the objective truth embodied in the 'text,' and are sometimes to be avoided altogether.

Harrison's suggestion that *Kontakte* is a piece of *musique concrète* (as opposed to a piece of *elektronische Musik*) also articulates a sense in which (just as 'beauty is in the eye of the beholder') the nature of music is 'in the ear of the listener.' As an obvious disciple of the top-down model, it is not surprising that Harrison seems to evaluate musical works in terms of their 'top-down' (that is, perceptual) characteristics, regardless of whether these played an important role in the composition of the work. Conversely, one might equally seek to evaluate a 'top-down' work in terms of its apparent 'bottom-up' (objective) characteristics, irrespective of the fact that these were not of any major concern to the composer. Thus, top-down and bottom-up approaches are just as much acts of *listening* as they are compositional acts.

2.5.4. The Relationship between Text and Context

As discussed in section 2.4, bottom-up works tend to be conceived in a manner that necessitates precision performance: in a compositional aesthetic that relies on deterministic structures, it is important that the structural relationships be recreated clearly and accurately for the audience, and it is equally important that no subjective 'interpretation' should cloud the objective nature of the music. This has strong implications for performance practice, for the presentation of any text (such as a piece of music) must invariably take place within a certain context: in a particular venue, with particular performers, using particular equipment, and for a particular audience. It follows that in the performance of the bottom-up work, which is 'monolithic' and has only one 'correct' performance, the context must be brought into line with the text: everything must be engineered to minimise 'inaccuracies' and 'errors' within the performance. In performing the top-down work, however, it is the text (the music) that must be shaped to fit the context, the music having been realised according to *subjective* truths, which are of course infinitely plural and highly context-sensitive. This idea is proposed by Savouret:

> I would now like to turn to creation-based [bottom-up] composition... The text is created from nothing: from it all things flow, and it must be considered as imperative. Here, the context in which it is projected is negligible, and may even be seen as totally subjugated. *Contrary to invention-based* [top-down] *composition*, it is the context that must submit: nothing is less certain nor more straightforward. [My italics.][108]

The impact of top-down and bottom-up approaches on the *performance* (diffusion) of electroacoustic music, with further development of Savouret's observations, will be discussed more fully in Chapter 3. Crudely speaking, in the performance of bottom-up music, the singular truth of the musical text must be brought to bear upon a context that can be contrived to suit, whereas in top-down music the *actuality* of the (particular) performance context is brought to bear upon the (infinitely

[108] Savouret (1998). "The Natures of Diffusion". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 351.

plural) musical text. The precise means by which this is achieved will also be discussed in Chapter 3.

It is proposed that all of the binarisms described in the previous sections – some of them more specific than others – point to the same underlying concept: a fundamental difference between top-down and bottom-up musical philosophies. This dual paradigm can be summarised as follows. In top-down composition, artistic justification is sought subjectively, that is, the 'truth' of the music is to be determined by the *human* process of perception. In this way, *subjective truth* can be revealed. Contrastingly, bottom-up composers typically seek to justify their work with reference to objective truths, or (pseudo-) scientific facts, which are beyond the fallibility of subjective opinion or, one could say, *super-human*. As compositions realised from the bottom up tend to be exacting in nature, it follows that their performance should take the form of a transparent realisation, and this explains the need to eliminate the possibility of 'errors' being introduced by performers, or by any other intermediate factors for that matter. There can be *only one* 'correct version,' and the context in which the work is performed must be engineered in such a way that the essence of the piece, as a singular architectonic object, can be perceived. On the other hand, top-down composers, having fundamentally based their work on the process of subjective perception, tend to positively embrace the notion of interpretation, endorsing the idea that their work should be in some way 'appropriated' by the performer, or otherwise 'submitted' into the context. In this respect, the top-down work can be regarded as 'plural' (different according to context), while the bottom-up work is essentially intended to be absolute, or 'monolithic.' Alternatively, one might describe the top-down work as always being 'relative' to its context (the performance venue, the performer, the audience, *et cetera*), while the bottom-up work has been *determined* – 'where once there was nothing' – by the composer, and is therefore absolute, and not 'relative' to anything. This is an analogue of the compositional process itself (as illustrated using the examples of *elektronische Musik* and *musique concrète*): the structure of bottom-up works is *determined* by prefabricated, absolute, abstract relationships (as in serialism for example), while the structure of top-down works unfolds *relative* to the phenomenal nature of already-existing material, and in subjective response to it.

Acousmatic music, described by Harrison as follows, represents a contemporary incarnation of the top-down model of musical thought:

> Acousmatic music on the whole continues the traditions of *musique concrète* and has inherited many of its concerns. It admits any sound as potential compositional material, frequently refers to acoustic phenomena and situations from everyday life and, most fundamentally of all, relies on *perceptual* realities rather than *conceptual* speculation to unlock the potential for musical discourse and musical structure from the inherent properties of the sound objects themselves – and the arbiter of this process is the ear. [Harrison's italics][109].

This is clearly indicative of a top-down approach because the compositional process is based on 'perceptual realities rather than conceptual speculation.' The trouble is that devotees of the opposing bottom-up approach would probably argue that it is subjective perceptions that are speculative, and that objective concepts are 'reality.' So what we are dealing with, really, is a dispute regarding how, exactly, we determine musical 'truth.' Is it to be found subjectively, or to put it in its most basic terms, 'If it sounds good and feels right,' or is it to be found in what we consider, objectively, to be true – in the mathematical nature of sound for example? By extension, is a piece of music to be regarded as a singular, monolithic, entity, or as a 'starting point' from which plural understandings might emerge *relative* to the many contextual factors that may impact on its reception? Many composers demonstrate an awareness of the relative benefits of both possibilities:

> There are two, apparently opposing, schools of thought. In the first, the choice of music must fit in with the characteristics of the venue and must be right for the audience... The second school of thought, however, believes that a strong well-made musical product will surmount all obstacles and survive in all conditions. I subscribe to both schools of thought.[110]

[109] Harrison (1999). "Keynote Address: Imaginary Space - Spaces in the Imagination". *Australasian Computer Music Conference* (Melbourne; 4-5 December 1999).
[110] Barrière (1998). "Diffusion, the Final Stage of Composition". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 205-206.

> Having dined at both tables, I am not denying that the two compositional attitudes are complementary...[111]

> This is [...] indicative of the (at least equal) importance of percept alongside concept in composition.[112]

Nonetheless it is also true to say that such composers will tend to express a specific affinity with one of the two philosophies:

> In preparing a concert, it is necessary, I believe, to begin by studying the hall... The hall's acoustic type must be noted: dry or reverberant; the size and volume of the space; its geometrical form; lighting, colours, the surface materials of the walls and the seats, the general atmosphere of the place [in other words, the context is considered before the text].[113]

> ...however, I do now understand better why I prefer the melody harmonization [Savouret is referring to the top-down ethos].[114]

> ...and here I declare my acousmatic [and, as previously discussed, top-down] allegiance...[115]

The terminologies used to describe the opposition of top-down and bottom-up philosophies are varied – 'two apparently opposing schools of thought,' 'figured bass versus melody harmonisation,' 'acousmatic music versus computer music or algorithmic composition' – but it is proposed that these are all different interpretations of the same, fundamental, underlying dichotomy. Clozier neatly encapsulates the notion of two opposing musical ideals as follows:

> Either the music is of a type that will not tolerate even the slightest variation without undergoing a negative change to its very essence, every instant having to be identical to the original model, or else it is of a type that allows itself to be enriched through interpretation and communication.[116]

Let us return briefly now to one of the technological defining characteristics of the electroacoustic idiom – the encoded audio stream – and pose the

---

[111] Savouret (1998). "The Natures of Diffusion". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 349.
[112] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 120.
[113] Barrière (1998). "Diffusion, the Final Stage of Composition". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 205.
[114] Savouret (1998). "The Natures of Diffusion". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 349.
[115] Harrison (1999). "Keynote Address: Imaginary Space - Spaces in the Imagination". *Australasian Computer Music Conference* (Melbourne; 4-5 December 1999).
[116] Clozier (1998). "Composition, Diffusion and Interpretation in Electroacoustic Music". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 252.

question, 'What exactly *is* it that we are encoding?' One answer might be as follows: the encoded audio stream is an abstract representation of real auditory events that contains precise information regarding the spectral content of the original phenomenon, and how this evolves over time; it is these objective data that are encoded. Alternatively: the encoded audio stream is an abstract representation of real auditory events that allows their sounding characteristics, and the ways in which we perceive them, to be manipulated; encoded within the streams is the potential for real, perceived, auditory experience. Clearly, both statements are true. The implication, however, is that bottom-up and top-down practitioners each recognise different potentials in the nature of encoded audio streams. The top-down practitioner regards the encoded audio stream as an abstraction of the perceptual qualities of the original auditory event, whereas the bottom-up practitioner regards it as an abstraction of the objective parameters that describe the auditory event.

Table 3, below, summarises the characteristics of top-down and bottom-up philosophies schematically.

Top-Down	Bottom-Up
Human	Super-human
Subjective	Objective
Composed to be interpreted/appropriated	Composed to be realised/disseminated
Realist (pragmatic)	Idealist
Plural	Absolute (monolithic)
Relativistic	Deterministic
Corporeal (physical)	Cerebral (intellectual)
Empirical/perceptual	Logical/conceptual
Qualitative	Quantitative
Phenomenological	Rational
Built/invented	Constructed/created
Organic	Architectonic
Abstracted forms	Abstract forms
Text submits to context	Context submits to text
Encoded audio streams are abstractions of the *subjective* (perceptual) qualities of real auditory events	Encoded audio streams are abstractions of the *objective* structures that define real auditory events

Table 3. A brief summary of the opposing characteristics of 'top-down' and 'bottom-up' approaches to musical discourse.

2.6. Examples of Top-Down and Bottom-Up Approaches in Contemporary Electroacoustic Music

Thus far, although passing reference has been made to more current work, top-down and bottom-up paradigms have largely been explained with reference to *elektronische Musik* and *musique concrète* which, as previously suggested, are now widely accepted as extinct in their pure form. It has also been proposed that the aesthetic standpoints engendered by both praxes present characteristics that remain in evidence to this day. A few choice examples will therefore be helpful.

Acousmatic music, described very briefly in the previous section, has already been cited as a contemporary example of the top-down approach to composition. Accordingly, one would expect composers within this genre to hold an interest in the perceptual, subjective realities of sound and its use within a musical context, and perhaps to be less interested in the abstract, objective schema that might be used to determine the musical dialogue. The following quotation is an extract from Simon Emmerson's description of his tape-only work *Points of Departure* (completed 1997):

> The sonic resources are all from the harpsichord and kayagum, from single pitches to clusters and resonances. A characteristic of both these instruments is their sharp attack and decay morphology. Indeed their playing techniques both suggest a struggle to extend this short event, to project it (diffuse it even), to prolong its sweetness as long as possible. From the toccata tradition of the harpsichord, the many subtle vibrato types of the kayagum and the arpeggio and tremolo flourishes of both we see this extension of simple linearity into denser textures – although always contrasted with the simple and often isolated plucked sound. As a kind of 'replacement' of these performance techniques, all the most obvious types of electroacoustic sound extension have been used: granular sampling, time stretching, reverberation, fast reiteration, all in many variants and combinations.[117]

Compare this with Michael Obst's description, taken from the same publication, of *Solaris* for ensemble and live electronics:

> *Solaris* is based on a science fiction novel by the Polish author Stanislav Lem. It seemed logical to include mathematical principles in both the composition and in the live electronics in order to construct a specific musical atmosphere. The compositional structure, sound treatment and sound spatialization were to provide a special ambiance for the story itself. The opera consists of three parts, each of which focuses on one mathematical principle: in the First Act stochastic treatment, in the Second Act interpolation [...] and in the Third Act mathematical functions (like the sine or the exponential function).[118]

In *Points of Departure* Emmerson seeks to express something of the essential nature of two musical instruments – the harpsichord and the kayagum – in terms of their cultural heritage, performance practice, and perceived sounding characteristics. His source materials consist of various recordings of each instrument being played, and descriptions are given of the morphologies of these sound objects, 'sharp attack and decay' for example. Although Emmerson mentions several audio processing technologies (granular sampling, time stretching, *et cetera*), no mention is made of how the parameters fed to these various processing algorithms were

[117] Emmerson (1998). "Intercultural Diffusion: 'Points of Continuation' (Electroacoustic Music on Tape)". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 283.
[118] Obst (1998). "Sound Projection as a Compositional Element: The Function of the Live Electronics in the Chamber Opera 'Solaris'". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 321.

devised; no mathematical principles relating numbers and functional processes to each other are described (compare this description, for example, with Stockhausen's account of his *Studie*[119]). On the contrary, Emmerson's description of his compositional approach suggests that the methods used were grounded firmly in the *perceptual* characteristics of each of the instruments' particular sounds. Stated simply, it can be suggested that Emmerson took various phenomenological aspects as 'given' (the characteristic sounds of the harpsichord and kayagum, and the respective cultural heritages of each, for example), and worked 'downwards' from these in order to arrive at the finished work. It can be deduced that compositional decisions were made on the basis of perceptual or *subjective* realities, rather than having been made with reference to a predetermined 'set of rules.' Being so firmly grounded in the *actuality* of sound, it is therefore possible – even only on the basis of the short quotation given – to imagine (to a certain extent) what the perceptual (sounding) qualities of the sound materials might be like.

Obst's working practices – and again, this can be deduced from only a very short quotation – are demonstrably different. In contrast with Emmerson, the compositional constituents of *Solaris* are abstract mathematical principles. Obst describes the numerical relationships between the parameters of the various processing algorithms in some detail (recall Stockhausen's 'tape speed: 76.2 cm per second'), for example:

> Here [are] the boundary conditions for the flanging and the pitch-shift: Flange: Frequency of the sine 0.25 Hz <-> 0.33 Hz, Feedback 0.8; Pitch-shift: Delay 5 ms, Transposition: -64 cents <-> +64 cents. Feedback 0.75.[120]

Recalling section 2.2.2, Obst's use of the expression 'boundary conditions' strongly suggests an approach to the compositional process comparable to that adopted by practitioners of *elektronische Musik*, and reference to

---

[119] See section 2.2.2.
[120] Obst (1998). "Sound Projection as a Compositional Element: The Function of the Live Electronics in the Chamber Opera 'Solaris'". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 322.

numerical, mathematical, or parametric precision in general is a common characteristic of the bottom-up compositional approach.

It would be very convenient to say that Obst makes no reference whatsoever to the perceptual characteristics of the resulting sounding material, but he does go on to say:

> The acoustic result is a continuous ribbon of sound which changes constantly without any periodicity. The long strings of the piano in the low register with their many partials provide a very rich sound.[121]

However, from Obst's writings it can be deduced that the spectromorphologies of his sound materials are, if you like, a 'by-product' of the mathematical algorithms used to produce them (Obst implies this himself in his use of the expression 'acoustic *result*'). This is in contrast with Emmerson's piece, whose very starting point was based on a perceptual evaluation of the already-existing 'acoustic results' of naturally occurring physical phenomena. Note also Obst's use of the objective/scientific descriptor 'many partials,' and compare this with Emmerson's perceptual evaluation, 'sharp attack and decay.' The distinction is subtle – indeed both expressions could plausibly be used as descriptors of the same phenomenon – but the contrasting use of language suggests that Obst and Emmerson each regard the encoded audio stream as an encoding of different attributes. For Obst, the stream is an encoding of the objectively measurable parameters of auditory events; for Emmerson, it encapsulates the potential for a perceptual evaluation of sounds.

Taking into account everything that has already been said, it should be clear that Emmerson's *Points of Departure* represents an electroacoustic work conceived in the spirit of top-down composition, while Obst's *Solaris* shows obvious signs of having been conceived in a bottom-up manner. It should also be restated that it would be erroneous to assert that either piece is 'absolutely top-down,' or 'absolutely bottom-up.' For one thing, there is not really sufficient detail as to the respective working procedures from which to draw such an assumption. Further, it has already been noted that Obst

[121] Ibid.: 322.

does indeed make reference to the perceptual qualities of his materials, and it is therefore sensible to assume that at least part of his compositional procedure may have been somewhat top-down in nature. Equally, it may well be the case that in composing *Points of Departure* Emmerson also made use of non-perceptual structuring schemes at some level, although there is no indication of this in his text. Nonetheless, it is reasonably safe to propose that, in terms of a continuum *ranging* from bottom-up through to top-down, Obst's piece is oriented *toward* the bottom-up, and Emmerson's *toward* the top-down.

2.7. Extending the Top-Down/Bottom-Up Paradigm

Having defined a context within which compositional practice may be evaluated in terms of its position along the top-down/bottom-up axis, it might be proposed that this paradigm in fact transcends the boundaries of electroacoustic music and may be manifest, albeit more abstractly, in earlier 'classical' music repertoire (and very probably in circumstances beyond music altogether). One might regard certain strict classical forms, such as fugue, as indicative of the bottom-up approach to composition, whereby an abstract set of rules is antecedent to the realisation of the musical fabric itself. Lévi-Strauss makes the following observation:

> Bach and Stravinsky appear as musicians concerned with a 'code,' Beethoven – but, Ravel too – as concerned with a 'message'...[122]

Clearly, a 'code' is an objective, abstract construct, whereas a 'message' is something that is intended to be *perceived*. Lévi-Strauss is effectively suggesting that Bach and Stravinsky were bottom-up composers, whereas Beethoven and Ravel were top-down.

Certainly, there are works within the classical repertoire that seem to have certain 'monolithic' aspirations, in the sense that they point toward a single 'correct' performance of the work. Highly prescriptive and precise performance directions might be regarded as indicative of this approach

[122] Lévi-Strauss (1964). *The Raw and the Cooked: Introduction to a Science of Mythology* (London; Pimlico): 30.

(Stockhausen's *Klavierstuck I*, already mentioned in section 2.4, might be re-cited in this context). On the other hand, there are works whose scoring clearly indicates, or indeed necessitates, a much stronger element of interpretation, perhaps even inviting the performer to put him or herself 'into' the work. A fitting example might be Erik Satie's *Trois Gnossiennes*, (1890). These piano works are scored without bar-lines or time signatures, and only one of them (the first) has a key signature. This suggests a non-monolithic compositional approach and invites the performer to *interpret* (as opposed to 'objectify') the score. Furthermore, there is a strong case for suggesting that annotations such as '*postulez en vous-même* (wonder about yourself)' (No. 1), '*sur la langue* (on the tip of the tongue)' (No. 1), '*sans orgueil*' (don't be proud)' (No. 2), '*munissez-vous de clairvoyance* (be clairvoyant)' (No. 3), and '*de manière á obtenir un creux* (so as to be a hole)' (No. 3), are performance directions that, realistically, can only be evaluated through subjective interpretation.[123]

Clearly much more could be said about the top-down/bottom-up paradigm within the context of western classical music, but such discussion would be digressive. This section does, however, serve to illustrate an important point: aesthetic 'directionality' (from the top down, or from the bottom up) in a musical context is not a foregone conclusion of the creative frameworks employed.

As previously discussed, it can be said that the compositional use of 'real' sounds is somewhat more conducive to the top-down approach, and the use of synthesised sounds perhaps lends itself more readily to the bottom-up approach. This, however, does not mean that any piece of electroacoustic music consisting mainly in synthesised sounds has 'by definition' been conceived in a bottom-up manner. Similarly, a piece comprising mainly 'real' sounds does not 'have' to be top-down. The fact is that the creative frameworks are, in a sense, 'merely' technological constraints. Before the advent of the technologies described in section 1.7, composers *simply did*

[123] Satie (1999). *Six Gnossiennes pour Piano* (Paris; Éditions Salabert).
Three further *Gnossiennes* were published posthumously, hence '*Six Gnossiennes*.'

not have the opportunity to work 'directly' with their materials: compositions, by and large, had to be written down in some form of musical notation. In this sense one might regard western classical notation as a compositional means, a framework with certain peculiar attributes, or even as a 'technology' of sorts. Some might argue that such notational syntaxes are, of themselves, bottom-up constructs (pitch, rhythm, timbre, and dynamic are, in the vast majority of cases, 'latticed') and as such would tend to spawn bottom-up compositions; Harrison speculates that, 'probably most [western classical] music before electroacoustic music was architectonic [bottom-up] almost by definition.'[124] This argument certainly has some validity, but is challenged to a certain extent by the suggestion that top-down and bottom-up models of musical thought predate the specific frameworks of electroacoustic music.

It is worth reiterating that every creative framework, by virtue of its particular characteristics, exerts certain unique pressures upon its user, and that such pressures can, of course, include a bias toward either top-down or bottom-up approaches to the use of the framework in question. Rephrasing Harrison's statement, it can certainly be argued that western classical notation is a framework that is inherently biased in favour of bottom-up approaches to the compositional process. However, we must also consider the fact that composers can only work within the creative frameworks available to them. In the absence of more conducive frameworks, the top-down composer (for example) has no choice but to 'work against' the inherent directionality of a less-than-ideal framework. As noted by Wishart:

> It is of course possible to subvert the various systems but it is a struggle against the design concepts of the instrument or software [read: creative framework].[125]

Outside of the strictly electroacoustic context, therefore, Clozier's words (previously cited towards the end of section 2.5) still ring true:

> Either the music is of a type that will not tolerate even the slightest variation without undergoing a negative change to its very essence, every instant having

[124] Harrison (2004). "An Interview with Professor Jonty Harrison". Personal interview with the author, 8th September 2004. See Appendix 1 for full transcription.
[125] Wishart (1996). *On Sonic Art* (Amsterdam; Harwood Academic): 27.

to be identical to the original model, or else it is of a type that allows itself to be enriched through interpretation and communication.[126]

Satie has been cited as a composer whose attitudes seem to have embodied the top-down approach to composition: outwardly, the *Gnossiennes* seem to enshrine 'subjective realities' rather than 'objective realities.' In many ways the scores force the performer to 'enrich through interpretation.' It can be deduced that in composing in this way, Satie had to work somewhat 'against the grain' of the (notational) frameworks available to him, and indeed these pieces clearly exhibit a deliberately subversive attitude towards these same means. In this sense it can be seen that composers working within this notational paradigm might experience more difficulty in composing 'from the top down' than composers working 'from the bottom up,' and this clearly illustrates the directional bias inherent within this particular framework.

2.7.1. Electroacoustic Music as a means to the Potential Unification of Top-Down and Bottom-Up Models

It can now be stated that, in electroacoustic music, the creative frameworks present equal opportunities for both top-down and bottom-up methods, and this has arguably not been the case at any other time in the history of western classical music. As a result of the technologies described in section 1.7, composers have direct access to auditory events in an abstract (encoded) – and, if necessary, static – form. This allows composers to work directly with their materials, as opposed to working via some kind of notational analogy.

Of course, the way in which composers choose to work directly with their materials can essentially be either top-down or bottom-up, with neither approach presenting significantly greater compositional difficulties than the other. This fact can be attributed to the dual nature of the encoded audio stream, which can at once be regarded as an

[126] Clozier (1998). "Composition, Diffusion and Interpretation in Electroacoustic Music". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 252.

abstraction of objective and/or perceptual information. Furthermore, and as a direct result of this, combined approaches are possible. Some aspects of a composition might be realised in a fundamentally 'perceptual' manner, while other aspects of the same composition might be the result of more 'objective' schemes. Indeed we might regard this possibility as one of the unique characteristics of electroacoustic creative frameworks (i.e. audio technologies). The ability to statically encode sound onto a recording medium and work directly with recorded materials, whether real or synthesised, is a characteristic of current compositional means that lends itself easily to top-down modes of composition. Simultaneously, the simple fact that algorithms of any sort (whether manifest in analogue circuitry or digitally coded software) must usually at some level be provided with parameters (*ergo* necessitating a denumerable set of discrete values) is a characteristic of current compositional means that lends itself more readily to composition from the bottom up. So, while it would be untrue to suggest that the top-down/bottom-up paradigm sprang into existence only after the advent of electroacoustic music, it would certainly be true to say that this continuum can be more fully explored – in a musical context – than ever before, by electroacoustic means. This observation was supported by Harrison, in a personal interview with the author, as follows:

> JM: [Do] you think that the binarism between the 'more-or-less organic' [top-down] and 'more-or-less architectonic' [bottom-up] ways of thinking, has existed for a long time in music?
>
> JH: Well it *has*. But the boundaries were moved when you became able to capture real-world sounds that existed beyond the 'normal' frame in which music happened, I think. [...] That pushes the boundaries suddenly, markedly, away from where they were. So yes, there's always been a range of composers operating, from the very intellectualised, to the more improvisatory (say Chopin or Liszt). You can situate composers somewhere along that axis, but until you get beyond having to use either notation and/or instruments designed for the purpose to reproduce the sounds you want – i.e. until you get recording – it seems to me that the boundary can only go so far; it's fixed. The minute you can record stuff that goes beyond that, then, 'Whoosh!' the boundary shoots back about a million miles!
>
> JM: So, in a sense, the advent of recording [and, by implication, the other technologies discussed in section 1.7] has opened up the full possibilities of that spectrum between the 'organic' [top-down] and the 'architectonic' [bottom-up]...

> JH: I would say so. [...] [Before that,] you had to specify what it was you wanted from [instrumental players], which meant that you were, of necessity, tending towards an architectonic thing. However improvisatory your aesthetic might be, you still have to write down that you want *this* note played at *this* level, for *this* long, on *this* instrument [laughs]. To go with *this* other note, *this* loud and for *this* long, on *this* other instrument! *Et cetera.* [...] So, yes, I think it's always been there, it's just that the degree of it is much more marked, this binary thing.
>
> JM: Absolutely. But I think that this is something that's not often acknowledged in the literature. In electroacoustic music it almost seems like the implication is that now we are doing things that have never been done before – which is obviously true, to a certain degree, in that there are things that it is now *possible* to do that it wasn't possible to do before – but perhaps in the way people, historically, have been thinking musically, there may be more similarities than the literature acknowledges?
>
> JH: ...But it is now very much more extreme, I think.[127]

To summarise: it can be proposed that the continuum between top-down and bottom-up philosophies, in a musical context, predates the existence of electroacoustic music. However, pre-electroacoustic creative frameworks have, by their nature, always tended to limit the scope of this continuum because composers have only ever had indirect access to auditory phenomena, via performers. In the context of bottom-up works (as described in section 2.2.2), this limits both the framework within which architectonic relationships can be defined, and the accuracy with which those relationships that are possible can be articulated. In the case of top-down approaches, the creative process is restricted in a different manner, but – paradoxically – by the very same factors. Because the musical discourse must be realised by performers, so the performers must be given some indication of what is to be performed. In practical terms this means that a system of differentiated performance actions must be defined (or, to re-cite Harrison, 'you want *this* note played at *this* level, for *this* long, on *this* instrument'), and this is somewhat contrary to the top-down ideal. Such restrictions are far less in evidence in electroacoustic music – owing to the direct access that composers have to their materials – and therefore both top-down and bottom-up approaches to composition can be explored more fully. This being the case, it is important that means be made available for the performance

[127] Harrison (2004). "An Interview with Professor Jonty Harrison". Personal interview with the author, 8th September 2004. See Appendix 1 for full transcription.

(diffusion) of, potentially highly contrasting, top-down and bottom-up works. This will become a criterion in the evaluation of sound diffusion systems in Chapter 4.

2.8. Summary

The purpose of this chapter has been to focus on the more specifically aesthetic concerns of electroacoustic music and, in doing so, add a further dimension to the mainly technological perspective offered in Chapter 1. The following conclusions are apparent:

2.8.1. Top-Down and Bottom-Up Approaches to Electroacoustic Music

As a starting point, the historical praxes of *musique concrète* and *elektronische Musik* have been examined. These contrasting praxes took place within a common technological framework – broadly speaking, both made use of the technologies outlined in section 1.7 – yet, aesthetically, the two approaches were radically different. The essential natures of these two approaches can be characterised as top-down (*musique concrète*) and bottom-up (*elektronische Musik*), respectively, and it is concluded that these two philosophies are still evident in contemporary electroacoustic music.

The abstract characteristics of both top-down and bottom-up approaches were summarised in Table 3 (page 84). In the briefest and most generalised terms possible, top-down can be described as a *per*ceptual approach, while bottom-up would be best described as *con*ceptual.

2.8.2. Creative Frameworks can be 'Directionally Biased'

In section 2.3, the essential nature of 'real' and 'artificial' (synthesised) sounds was briefly examined. From this it was suggested that 'real' sound, by its very nature, is particularly conducive to top-down working procedures, while 'artificial' sound is more in-tune with the bottom-up method of working. Accordingly, it is not surprising that *musique* *concrète*, with its focus on the creative possibilities of 'real' recorded sounds, should have developed into a top-down idiom, while *elektronische Musik*, with its synthetic means, should be bottom-up. This indicates a dialectical relationship between the nature of the music, and the nature of the creative frameworks employed in its realisation.

In more general terms, it can be concluded that all creative frameworks ('real' sounds, 'artificial' sounds, notational systems, technologies, software, sound diffusion systems, *et cetera*) can potentially be 'directionally biased,' and therefore gravitate to an extent towards either the top-down or the bottom-up end of the spectrum. Notation in western classical music, for instance, has been cited as an example of a creative framework that is biased toward bottom-up working methods. This can be seen as an extension to some of the arguments proposed in the previous chapter, whereby technologies are regarded as having affordances and therefore engender particular creative opportunities and ways of working. Therefore, the aspirations of composers can either be reinforced or impeded to a certain extent by the frameworks within which they choose to realise their work. In section 2.4 the compositional process of Stockhausen's *Étude* was briefly accounted. This represents a good example of a composer with characteristically bottom-up aspirations working within a framework that is inherently biased toward top-down methods. Unsurprisingly, therefore, Stockhausen expresses a certain degree of frustration that the results were not as he had anticipated. This, of course, does not mean that music written within a top-down framework is *always* top-down, nor that music written within a bottom-up framework is *always* bottom-up, but merely that the frameworks can have certain inherent biases and therefore exert certain directional pressures on composers.

2.8.3. Electroacoustic Technologies allow Equally for Top-Down and Bottom-Up Compositional Approaches

The practice of electroacoustic music exists, to a considerable extent, *within* the abstraction layer illustrated in Figure 1 (page 11), that is, much of the practice itself exists inside the abstract domain of encoded audio streams: all of the technologies associated with the idiom (described schematically in section 1.7), at some level, share this common communication protocol. Furthermore, electroacoustic means offer composers, for the first time in a musical context, the opportunity to work directly with their materials, as opposed to working via symbolic representational means such as notation. The *overall* framework of electroacoustic music can therefore be regarded – perhaps uniquely – as *both* top-down *and* bottom-up, because frameworks that are essentially top-down (such as 'real' recorded sound statically encoded onto recording media) and those that are essentially more bottom-up (such as sounds synthesised algorithmically) can be easily integrated. Consequently, the electroacoustic idiom presents equal opportunity for both top-down and bottom-up modes of composition. It is therefore reasonable to propose that electroacoustic music, as a whole, will be represented by top-down and bottom-up works in approximately equal measure, with works that embody these two ideologies in combination perhaps also being common.

Overall, the function of this chapter has been to present a context within which electroacoustic music can be regarded, in aesthetic terms, as a binary art-form. The purpose of this investigation has been, in conjunction with Chapter 1, to ascertain the technological and aesthetic demands that may be placed on sound diffusion systems used for the public performance of electroacoustic music. In the next chapter it will be argued that the underlying top-down/bottom-up binarism is carried forward into the performance context, resulting in two contrasting performance practice ideologies and, as will become apparent in Chapters 4 and 5, two fundamentally different kinds of diffusion system. If, however, works of varying technical and aesthetic demands are to be performed side-by-side in concerts of electroacoustic music, then sound diffusion systems that can cater for diversity in these areas are required.

MISSING PAGE

3. Sound Diffusion

3.1. Introduction

This chapter focuses on the notion of 'sound diffusion,' that is, the public performance of electroacoustic music to an audience. It was concluded in Chapter 1 that all electroacoustic music is dependent, at least in part, on mediation via loudspeakers, owing to its reliance on encoded audio streams. Accordingly, it can be stated that the process of sound diffusion must also involve loudspeakers and, furthermore, that the process of sound diffusion – if for this reason alone – is a fundamentally necessary one.

We also know from Chapter 1 that any given piece of electroacoustic music will involve certain further technical demands; these were outlined in sections 1.7 and 1.8. The ways in which these demands are met (or, indeed, are not met) must therefore also be considered part of the process of sound diffusion. Furthermore, it has been proposed that, aesthetically, electroacoustic works can broadly be sub-divided into those that are essentially top-down and those that are essentially bottom-up. These contrasting aesthetic demands must therefore also be catered for.

The ultimate purpose of this chapter is to set up a context within which specific sound diffusion systems can be evaluated in terms of how successfully they meet the technical and aesthetic demands imposed upon them; the evaluation itself will be one of the primary objectives of Chapters 4 and 5.

Discussion will begin with a summary of some of the acoustic difficulties inherent in presenting sound via loudspeakers in (often relatively large) performance venues. From this it will be concluded that presentation of encoded audio channels to the audience *in a way that is congruent with 'what the composer intended'* is, at some level, a unanimous objective of the process of sound diffusion. It will be further proposed that this overbearing concern tends to take precedence regardless of aesthetic directionality, that is, it applies to both bottom-up and top-down composers alike.

An additional function of the present chapter will be to propose two original abstract concepts intended to clarify the purpose of sound diffusion from both practical and theoretical perspectives. These interrelated concepts are that of the *coherent audio source set* (CASS), and that of the *coherent loudspeaker set* (CLS). The CASS is a theoretical framework that allows multiple (one or more) monophonic encoded audio streams to be conceptually 'grouped together' and treated, collectively, as a single, homogeneous entity, thus maintaining a constant relationship between the constituent channels. The two discrete channels – 'left' and 'right' – of a stereophonic recording are an intuitive example. The CLS, essentially, implies a group consisting of one or more loudspeakers arranged in an arbitrary formation. The constituent channels of a CASS are related to each other via a *conditional coherent bond*, that is to say, the set is only coherent on the proviso that certain conditions are met. One of the conditions of the coherent bond is that the CASS must be reproduced over an appropriate CLS (or CLSs). In this newly-established context, it will be suggested that the fundamental aim of sound diffusion is to present a CASS (or multiple CASSs), via a CLS (or multiple CLSs), to an audience, in a manner that ensures that the conditions of the coherent bond(s) are satisfied.

This suggestion may be open to criticism, however, on the grounds that it implies that the process of sound diffusion is one upon which all practitioners of electroacoustic music agree. This, of course, could not be further from the truth. In rearticulating the proposition, one might posit that the overall *credo* of sound diffusion is to effectively communicate the discourse of a piece of electroacoustic music to the public, and that – owing to the nature of the idiom – this process must involve the presentation of encoded audio channels (which can conceptually be grouped into CASSs) over loudspeakers (which can conceptually be grouped into CLSs). Of course, 'the discourse of a piece of electroacoustic music' will depend enormously upon whether that piece has been conceived and realised in an essentially top-down, or in an essentially bottom-up, manner.

As described in Chapter 2, the top-down practitioner regards the encoded audio stream as an abstract representation of the perceptual realities of sound, whereas the bottom-up practitioner treats it as a carrier of more objective data. The present chapter will therefore also discuss the differing attitudes of top-down and bottom-up practitioners with respect to the act of sound diffusion itself, on the basis of this distinction. It will be proposed that both top-down and bottom-up philosophies are carried forward into diffusion practice, implying a 'composition-diffusion continuum' that results in two highly contrasting performance practice ideologies. An analysis of the impact that this has on defining top-down and bottom-up attitudes towards sound diffusion will be offered, with particular reference to notions of interpretation and performance 'context.' Ultimately it will be concluded that the debate in question centres on the relationship between text and context, where 'text' refers to the piece of music itself, and 'context' connotes everything else – the audience, performance venue, diffusion system, performance methodology, and so on. Specifically it will be argued that in the case of top-down diffusion, the text is adapted (performed) in a manner appropriate to the context (and is, in this respect, flexible), whereas in bottom-up diffusion the context is adapted to suit the requirements of the text, which in this case is regarded as singular. Specific examples of both possibilities will be given.

3.2. What is Sound Diffusion?

'Sound diffusion' refers to the process involved in presenting electroacoustic music to an audience.

As previously observed, electroacoustic music is an idiom that relies extensively on encoded audio streams, which, of themselves, are of little use unless they are eventually to be decoded via the process of reproduction over loudspeakers (see section 1.6). It can therefore be assumed that the process of sound diffusion will necessarily involve at least one loudspeaker, and probably many more. Other technological demands are less certain, but it is likely that one or more of the technologies outlined in section 1.7 will be implicated, and may collectively take the form of one or other of the technological permutations summarised in section 1.8. In perhaps the simplest possible example, playing a compact disc on a hi-fi system with loudspeakers can be regarded as an act of sound diffusion. Sound diffusion is, therefore and above all, absolutely necessary, simply because it represents the process that ultimately renders electroacoustic music audible:

> Electroacoustic music can only have its being in the act of diffusion.[128]

An important aspect of Clozier's statement lies in his use of the phrase '*act* of diffusion.' This tellingly implies that we are dealing with an *active* process as opposed to a passive one. While – in the crudest and most generalised terms – sound diffusion can be regarded as the process in which works of electroacoustic music are 'played back' via loudspeakers, this is ultimately a misleading generalisation because the notion of 'playing back' – erroneously – implies a *passive* act. An important function of the present chapter, therefore, will be in determining the nature of the *active* and *intentional* process of sound diffusion. In what respect, exactly, is it an active process? And what, exactly, are the considerations that this active process entails? In short, what *are* the 'acts?' This enquiry will begin with a summary of some of the acoustic difficulties associated with the performance of electroacoustic music.

3.3. Acoustic Issues in the Performance of Electroacoustic Music

The performance of electroacoustic music to an audience clearly requires some kind of venue large enough to accommodate multiple listeners, and one that is appropriate for public performance. It can therefore be said that, in the vast majority of cases, sound diffusion involves the presentation of electroacoustic works in spaces that are relatively large in comparison with, say, domestic or studio listening environments. This raises a number of

[128].Clozier (1998). "Composition, Diffusion and Interpretation in Electroacoustic Music". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 235.

issues that are unique to the performance of electroacoustic music; issues that are further compounded by the broad technical demands and aesthetic standpoints embodied within the idiom. There are a number of acoustic and psychoacoustic considerations to be addressed, mostly stemming from the fact that such performances take place in (relatively) large venues to (relatively) large numbers of people. A number of these issues are summarised by Doherty.[129]

The sheer variety of venues in terms of their size and geometry – and therefore their acoustic characteristics – is an important consideration:

> The variety of venues where electro-acoustic music is played is enormous. This ranges from lecture hall-size spaces to large churches... This kind of presentation creates many problems for listeners, not least that large spaces severely interfere with stereo imaging, tending to destroy positional information and mask detail.[130]

Harrison documents the same problem as follows:

> If a stereo piece is played over a stereo pair of loudspeakers (even large speakers) in a large hall, the image will be even less stable and controllable than in a domestic space, and will certainly not be the same for everyone in the audience... Listeners at the extreme left or right of the audience will receive a very unbalanced image; someone on the front row will have a 'hole in the middle' effect, whilst a listener on the back row is, to all intents and purposes, hearing a mono signal![131]

Harrison's latter example (the listener at the back of the hall) is confirmed by Doherty, who additionally notes problems of phase cancellation and loss of stereo image integrity as a direct result of large distances between listeners and loudspeakers:

> In a large space, temperature and humidity variations and air movement create unwanted and continually varying changes in the phase of a signal as measured at a listening point... This results in variations in the phase relationship between the left and right loudspeaker output wherever you sit. These cause varying phase cancellations of different frequencies, as well as the destruction of any meaningful spatial information. Acousticians and sound engineers generally agree that in large spaces, particularly towards the back, frequencies

[129] Doherty (1998). "Sound Diffusion of Stereo Music Over a Multi-Loudspeaker Sound System: From First Principles Onwards to a Successful Experiment". *SAN Journal of Electroacoustic Music*, **11**.
[130] Ibid.: 9.
[131] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 121.

> above 3000 Hz cannot carry any sense of stereo in the phase relationship between left and right loudspeakers.[132]

And let us not forget the inverse-square law:

> As the sound spreads out from a source it gets weaker. This is not due to it being absorbed but due to its energy being spread more thinly... Every time the distance from a sound source is doubled the intensity reduces by a factor of four, that is there is an inverse square relationship between sound intensity and the distance from the sound source.[133]

In the context of sound diffusion, the intensity of sound experienced by a listener positioned twelve metres away from a loudspeaker will be one quarter of that experienced by a listener seated only six meters away from the same loudspeaker, simplistically speaking. Perhaps one might address this problem, simply, by driving the loudspeakers harder? Unfortunately, the solution is not as simple as that, as Doherty goes on to explain:

> For every octave down that you wish to reproduce you need four times the power used on the first octave to achieve similar perceived loudness. For example, to produce a sound as loud as that produced by one watt at 1 kHz at a frequency of around 32 Hz you would need around one thousand watts (assuming similar loudspeaker sensitivities). This means that the lower octaves use the bulk of audio power. Without lots of power driving the lower regions and moving lots of air, the loudspeaker output will tend to be bass light. Couple a lack of power in the bass with phase cancellation of higher frequencies, and you end up with an over-emphasis of mid frequencies, particularly noticeable in loud passages.[134]

In summary, large distances between audience members and loudspeakers – particularly in large and reverberant spaces – are problematic, and the problems are compounded when this distance is increased. These problems include loss of phantom imaging integrity and frequency-dependent colourations of the sound, as well as additional issues relating to the amount of power required to achieve satisfactory levels of sound intensity for distantly located listeners. The latter of these issues in particular forces us to deal with the non-linear relationship between power and frequency

[132] Doherty (1998). "Sound Diffusion of Stereo Music Over a Multi-Loudspeaker Sound System: From First Principles Onwards to a Successful Experiment". *SAN Journal of Electroacoustic Music*, **11**: 9.
[133] Howard and Angus (1996). *Acoustics and Psychoacoustics* (Oxford; Focal Press): 29-30.
[134] Doherty (1998). "Sound Diffusion of Stereo Music Over a Multi-Loudspeaker Sound System: From First Principles Onwards to a Successful Experiment". *SAN Journal of Electroacoustic Music*, **11**: 9-10.

response. Aside from the problems posed by simple distance from the loudspeakers are issues such as those raised by Harrison: it is *impossible* for all of the members of an audience to be centrally located, and off-centre listeners will experience problems. All of these observations offer up an obvious conclusion: the accurate reproduction of encoded audio streams via loudspeakers is very likely to be jeopardised in most performances of electroacoustic music unless appropriate measures are taken to minimise this. As noted by Harrison:

> Whatever a composer has put on a tape is potentially at risk in a large space unless positive steps are taken to reinstate what would otherwise be lost.[135]

It is in this important respect – referring back to section 3.2 – that sound diffusion is an *active* process rather than a passive one.

It seems clear (and perhaps this is obvious) that the communication of a piece of electroacoustic music through the process of sound diffusion must be, in some sense, *coherent* with respect to the intentions of the composer. Although it seems reasonable to suppose that the 'intentions' of composers will be enormously variable – the 'message' of the music is likely to vary dramatically from composer to composer, and from work to work – in the context of sound diffusion (where works doubtlessly exhibiting profound technological and aesthetic differences will, nonetheless, have to be performed side-by-side), it is certainly worth considering the *common* interests of composers working within the electroacoustic idiom. What might these be?

In Chapter 1 it was proposed that – despite differences in specific technological profile – the presence of encoded audio streams intended for later decoding via loudspeakers is a defining characteristic of electroacoustic music. In this respect, the use of encoded audio streams represents a common interest for electroacoustic musicians, as does the subsequent use of loudspeakers as a means to decoding them. It is proposed that this applies regardless of whether the work in question is top-down or

[135] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 124.

bottom-up. In section 1.6.2 it was further noted that the use of *multiple* encoded streams, and their subsequent decoding via *multiple* loudspeakers – through the phenomenon of phantom imaging – ultimately leads to the emancipation of spatiality in an auditory context. Unsurprisingly, the 'uniquely spatial' potential of such audio technologies has been cited – in section 1.13.2 – as an important avenue for creative exploration within electroacoustic music. It is therefore proposed that the *accurate reproduction of phantom images* is, amongst other things of course, a common concern in the diffusion of electroacoustic music.

Because the existence of phantom images is dependent on the co-ordinated use of *multiple* encoded audio streams to be decoded via *multiple* loudspeakers, it seems logical that there should exist a conceptual framework within which streams can be dealt with in a multiple, and co-ordinated, manner. Surprisingly – excluding 'convention' – there does not seem to be any such framework. It is for this reason that the concepts of the 'coherent audio source set' (CASS) and the 'coherent loudspeaker set' (CLS) are proposed; these will be described in sections 3.5 to 3.8.

3.4. Diffusion Example: Two Movements from Parmegiani's *De Natura Sonorum*

The following sections are designed to indicate – by way of example – the kinds of actions and outcomes that a performer might wish to perform in the diffusion of a work of electroacoustic music. This is intended as a supplement to the otherwise more abstract description of the act of sound diffusion as presented in this chapter. A brief account of some (but by no means all) of the logistical practicalities inherent in the performance of these diffusion actions will also be given.

Two movements from Bernard Parmegiani's stereophonic tape-only work *De Natura Sonorum*[136] (1975) will be described; these will be the second movement – entitled 'Accidentals/Harmonics' – and the third – 'A

[136] Parmegiani (1991). *De Natura Sonorum*. (Compact Disc - INA: INAC3001CD).

Geological Sonority.' Audio recordings of these movements can be found on the accompanying audio CD.

These particular movements have been selected for a number of reasons. Firtly, they are highly contrasting in nature, and therefore afford the opportunity to describe quite different diffusion techniques in each case. Secondly, the movements are reasonably 'monothematic' in terms of the kind of diffusion actions they demand: each movement therefore presents the opportunity to explore one specific 'style' of diffusion in some detail. Finally, the first of these two movements runs *segue* into the second, necessitating a description of the techniques required to perform the movements consecutively.

As a classic work of acousmatic music, it can be stated that *De Natura Sonorum* is oriented towards top-down ideologies. Accordingly, appropriate diffusion actions for the performance of these movements have been arrived at mainly on the basis of a perceptual evaluation of the musical materials and their development and interaction with each other. Top-down diffusion theory will be described more fully, in abstract terms, in section 3.12.

In the analysis that follows it will be assumed that the loudspeaker array illustrated in Figure 2, below, is being used. This is based on the BEAST array for the diffusion of stereophonic tape works as described by Harrison[137], who provides a detailed rationale for the placement of each loudspeaker pair.

[137] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 121-123.

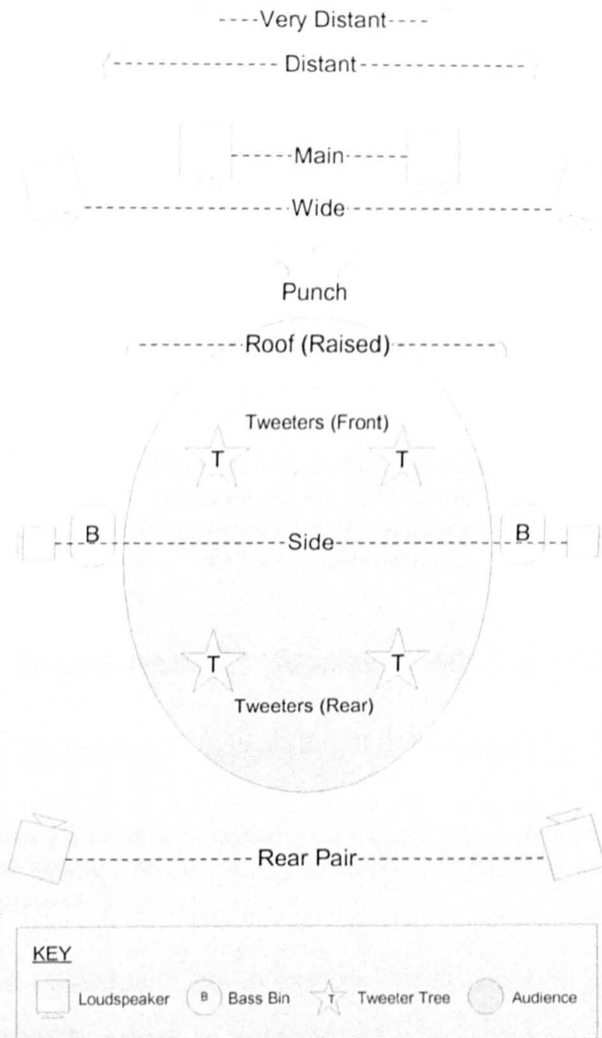

Figure 2. Loudspeaker array assumed for the diffusion of movements from Parmegiani's *De Natura Sonorum*

At present, the vast majority diffusion concerts are carried out using mixing desk style systems whereby the level of each loudspeaker must be controlled with an individual fader. The BEAST system (to be described in section 4.8) operates on this basis and for the purposes of this discussion it will be assumed that the mixing desk configuration illustrated in Figure 3 is being used. Again, this is closely based on a configuration described and rationalised by Harrison.

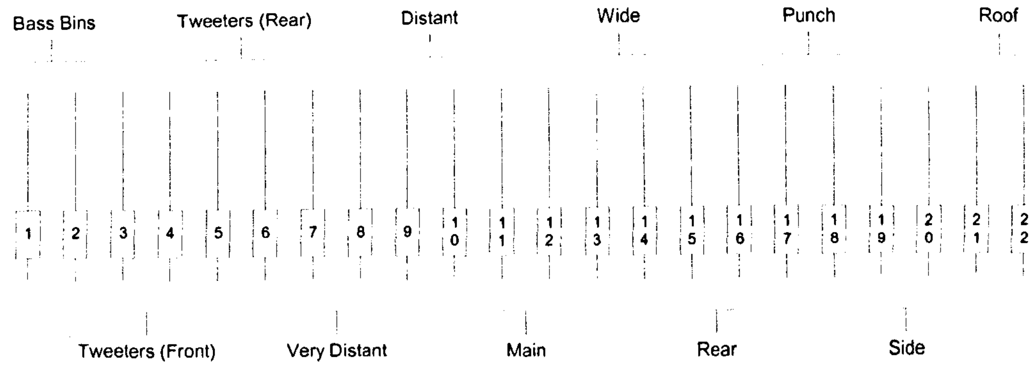

Figure 3. Mixer configuration assumed for the diffusion of movements from Parmegiani's *De Natura Sonorum*

3.4.1. Movement 2: 'Accidentals/Harmonics'

'Accidentals/Harmonics' is described by Parmegiani as follows:

> Often very brief events of instrumental origin are brought in to modify the harmonic timbre from the continuum that they undercut or on which they are superposed.[138]

At the most basic level, the movement comprises one continuous drone of uniform pitch, which is punctuated by sporadic and sudden gestural interjections. With regard to sound diffusion – and, again, at the most basic level – it would seem appropriate for the continuous drone material to be relatively static in terms of spatial articulation, and for the gestural interjections to be more spatially dynamic. In order to go into more detail, the movement will be divided up into seven sections, appropriate diffusion actions for each of which will be described in turn.

The first section introduces the pitched drone, tentatively at first, but this quickly becomes more firmly established in discrete steps (as opposed to smoothly or gradually) brought about by superimposed entries that add to its mass and change its timbral character, as described by the composer. In diffusion, the drone could begin distant, with the abrupt introduction of sequentially closer pairs of loudspeakers

[138] Parmegiani (1991). *De Natura Sonorum*. (Compact Disc - INA: INAC3001CD). Quotation from accompanying booklet.

coinciding with each one of these entries. At 0'29", for example, a very brief pitched interjection affects a change of quality in the underlying drone, which at this point attains a distinctive spatial morphology that oscillates rapidly from left to right within the stereo field. This would be an appropriate point for the drone to reach the wide pair of loudspeakers (or even the sides) as this would emphasise the intrinsic left-to-right movement. In terms of its physical demands this progression poses no major difficulties because individual fader pairs are introduced sequentially, with plenty of time available to prepare for each new entry. As we will soon see, however, when diffusion actions are to be performed in more rapid succession – and particularly if these actions involve the simultaneous use of more than two faders – the logistical difficulties can become considerable.

The drone is interrupted for the first time at 0'59", marking the beginning of the second section. At this point, a loud and abrupt 'drum roll' occurs, halting the drone – albeit very briefly – in its tracks. This is an important moment also because it introduces a timbre – unpitched and percussive – that has not yet been heard in the hitherto entirely pitched sound-world of the movement. Really, this entry is unexpected enough to be effective without any particular diffusion action, but could be emphasised with, for example, a sudden shift to the main or (even more emphatic) punch pair of loudspeakers. This would be achieved by raising faders 17/18 to coincide with the sudden percussive entry. The results would be even more effective if this was accompanied by a simultaneous lowering of all of the other fader levels, although this would clearly be much more difficult to achieve in practice.

The drone immediately reasserts itself, and this is followed by two short 'sub-sections,' the first of which is characterised by unpitched percussive entries that are – to use the composer's terminology – *superposed* on the drone (i.e. they do not 'interrupt' it as such) and the second of which is characterised by similar entries that *undercut* the drone (i.e. they *do* interrupt it). The sub-sections are separated by two brief pauses, occurring at 1'09" and 1'25" respectively. In diffusion such pauses can be used to affect a 'scene change' if necessary, setting up a different part of the performance space for the following section if this is appropriate. There is no particular need to treat these sub-sections differently in diffusion: the composed differences fixed onto the medium will be enough to make the difference between the two sufficiently clear. In both cases the sudden interjections can be timed to coincide with a slight emphasis on the main or punch pair of loudspeakers (in the case of interjections consisting of a single brief impulse) or – in the case of longer gestures such as that occurring at 1'12" – a more spatially dynamic articulation in which the gesture is 'thrown around the space' among several loudspeaker pairs. The mixer configuration is such that gestural articulation among the 'main eight' loudspeakers (mains, wides, distants, rears) is fairly straightforward. Articulations incorporating other loudspeaker pairs – however effective this might prove to be – would be logistically more difficult to achieve. In either case, it is very important that the drone be clearly established as spatially static, and that any sudden fader movements are carefully timed to coincide with the percussive or gestural interjections. The main challenge here, therefore, will be in simultaneously lowering the levels of multiple faders in order to re-establish the drone at a static spatial location. If this is achieved, however, the psychoacoustic effect should be sufficiently persuasive to convince the audience that the drone remains 'stationary,' while the gestural interjections articulate different parts of the performance space.

The third section – which follows the brief pause at 1'24" and lasts until around 1'52" – is dominated by the drone and is, accordingly, rather more static, in terms of its spatial morphology, than the previous section. Interjections are also less frequent, and – excluding the gesture with which this section begins – consist only in individual percussive hits. Two of these impulses – occurring at 1'35" and 1'44" – affect changes in the dynamic profile of the underlying drone. The first initiates a *crescendo*, the second, a *subito piano* followed by a *crescendo*. At these points it would be effective for the drone to gradually grow in terms of the space it occupies, closely tracking each *crescendo*, starting in the distant pair and steadily expanding through mains, wides, and perhaps even to sides. The only real inconvenience here is the gap between faders 13/14 (wides) and 19/20 (sides), which could make the realisation of a smooth front-to-back transition somewhat awkward. The percussive hit at 1'44" would be effectively diffused via the punch pair of loudspeakers, this coinciding with the instantaneous 'shrinking' of the drone back into the distant pair in preparation for the second 'growing' *crescendo*. This will be more difficult to achieve because – once again – it involves the sudden and simultaneous movement of multiple faders. This is one of the relatively few passages where gradual fader movements will be necessary for any extended length of time.

Transition into the fourth section is affected by a *marcato* saxophone-like entry consisting of two rapid repetitions of the same note (1'52"). This motif occurs twice, the implied space of the second occurrence (1'54") being more distant than that of the first. This can be used in diffusion to re-establish the drone via the distant pair of loudspeakers in readiness for the following section. The most prolonged in the movement – lasting until 3'03" – diffusion of this section will largely consist in articulating the frequent gestural movements around the performance space whilst maintaining those brief moments of stasis where the drone dominates the musical texture. The drone is relatively quiet (and perhaps distant) throughout this section and this should be reflected in the diffusion. There are, of course, many points at which more specific actions could be effective. The *pizzicato* double-bass notes (2'00") could be nicely emphasised with the subtle use of bass bins, highlighting the fact that up to this point there has been little pitched material with frequency content as low as this. In addition, notable gestures such as the reversed breathy gesture at 2'35", the pitched saxophone-like crescendo from 2'44" to 2'46", and the sweeping upward glissando beginning at 2'51", could be used to 'open out' the spatial articulation or otherwise shift the spatial emphasis in various ways. The upward glissando could be used, for instance, as a means to introducing the 'roof' pair of loudspeakers into the performance if these have not been used previously.

The fifth section begins with a deep, bassy impact at 3'11" (the use of bass bins would be particularly effective here), which initiates the establishment of new material that is pitched, resonant, fluttering, and reverberant in quality. (This new material has, in fact, been pre-empted in the previous section, very subtly at first (2'39", 2'53") and then more obviously (3'03", 3'07"). These points should be used as a means to establishing the diffusion behaviour that will dominate the following section and could, perhaps, be diffused to distant loudspeakers.) Throughout the section this material alternates and interacts with the drone, which has also become more reverberant and less intimate in quality, but somewhat faster and more continuous in terms of gestural movement. This section could be diffused via a rapid interplay between reverberant/distant space (e.g. 3'11" to 3'14") and a closer space (e.g. immediately following 3'14"). The section is also quite 'expansive' in character, and should probably therefore make use of many loudspeaker pairs; interplay between outlying loudspeakers (e.g. very-distants, distants, sides, loudspeakers located in the roof, rears) and those located closer to the audience (mains, wides, and the punch pair) would be very effective, building to a peak at 3'35". This will be particularly tricky because it involves the manipulation of two different groups of faders that are not conveniently juxtaposed on the mixing desk. The outlying loudspeakers are routed through faders 7/8, 9/10, 15/16, 19/20 and 20/21 respectively, while those located closer to the audience occupy faders 11/12, 13/14 and 17/18. A rapid interplay between these two distinct spaces will therefore be extremely difficult to execute, the actions achievable in practice being seriously restricted by the physical ergonomics of the system.

At this point, the sound should completely fill the space as the musical dialogue becomes more continuous with the prominent establishment of the drone at the forefront. Shortly, the drone is accompanied by much gestural and textural activity, which should be articulated quickly around the performance space. Because many loudspeakers will be in use at this point, this will involve frequent but subtle 'nudges' to the faders. There is rather a lot of composed left-to-right movement in this section, and so frequent use of the wides and sides could be used to emphasise this. There is also much high-frequency content in the rapid textural material, which would very effectively enlarge the perceived 'size' of the performance space if diffused via tweeter trees above the audience. This sixth section – running from 3'35" to 4'11" – should be the 'largest' in the movement, in terms of the number of loudspeakers in simultaneous use.

The seventh and final section, running from 4'11" to the end of the movement at 4'46", exists as a preparation for the subsequent movement, which follows *segue*. It should therefore be used in diffusion as a means to preparing and establishing the correct space for the following movement, which – as will be seen shortly – is considerably different in character. The section begins with an abrupt halting of the previous rapid gestural activity, brought about by a distant, resonant drum-like entry. Again, such a rapid shift to distant loudspeakers could prove challenging. This is followed by a series of emphatic, fast-moving statements, each abruptly terminated by a similar percussive interruption. The implied spatial articulation is therefore one of rapid movement, followed by sudden stasis, followed by rapid movement, and so on. Here, the gestural material is reverberant, implying that the spatial articulation should be large and expansive, probably employing large numbers of loudspeakers for its diffusion. The percussive interruptions are 'distant' in character, and could be effectively articulated via sudden shifts to distant and very distant loudspeakers. As described previously, however, the actions achievable in practice are likely to be limited by the physical constraints imposed by the diffusion system itself. The hiatus immediately preceding the final emphatic impulse-like entry at 4'39" should be used to prepare the mixing desk for the bass drone, which is introduced at this point to underpin the beginning of the following movement. The way in which this bass drone should be diffused will be described shortly.

In summary, 'Accidentals/Harmonics' is characterised by moments of relative stasis – where the musical material is dominated by the pitched drone – interspersed with percussive impacts and gestural statements in which much spatial articulation is implied. Sudden changes in timbral quality or gestural behaviour can normally be successfully diffused with sudden fader movements, provided the timing is accurate enough. Where timbre is static, on the other hand – as is often the case here when the pitched drone dominates the musical dialogue – this tends also to imply spatial stasis, and sudden fader movements run the risk of yielding perceptually implausible and/or distracting results. The main challenge, therefore, in the diffusion of this movement rests in maintaining the perceptual separation between the underlying drone and the gestures that articulate it, and in doing so mediating and clarifying the relationship between the two throughout. This challenge is, of course, augmented by the fact that the drone frequently changes in character, as do the articulating gestures, as does the relationship between the two!

3.4.2. Movement 3: 'A Geological Sonority'

'A Geological Sonority' is markedly different in character from the previous movement. Here, the tempo is considerably slower, and the musical material develops through what is essentially an interplay between multiple layers of broad, continuous, drone-like material, which is conveyed in long, drawn out, sweeping gestures. The effect is rather like being repeatedly 'washed over' by multiple waves of sound. Parmegiani states that the movement 'resembles flying over a landscape in which the different 'sound' levels will emerge on the surface one after the other.'[139]

The movement begins with a single low-pitched bass drone, as established at the end of the previous movement. The intensity and regularity of the subsequent 'waves' of sound gradually increases to a climax at around 3'00", after which the original low-pitched bass drone is re-established. At this point, the pace of the music rapidly diminishes, finishing more or less as it started with the solo bass drone.

The first 'wave' grows out of the solo bass drone with the very gradual introduction (from 0'20" to 1'00") of a second textural layer, which – although not unpitched – has a noisier spectrum than the underpinning bass drone. At 1'25" this second drone enters a slow downward glissando whilst gradually decreasing in dynamic level, and finally disappearing at 1'48". The glissando is accompanied by the staggered introduction of a further three distinct textural layers that serve to build the dynamic level and thus bring the wave to a 'crest': a drone whose pitch is in between that of the bass drone and the second layer introduced, and two noisy drones, one of high tone-height (a hiss) and the other much lower (a rumble). The perceptual experience of a wave – beginning in the distance, passing over the audience, and eventually disappearing in the opposite direction from whence it came – is the result of two important factors. Firstly, the wave gradually builds in dynamic level, reaches a peak of intensity, and then dies away. Secondly, the gradual downward glissando is rather like the Doppler effect experienced by listeners when a moving sound source passes by. In this case, the onset of the glissando happens when the dynamic level reaches its peak, further emphasising the effect.

One approach that could be adopted in the diffusion of this movement, therefore, is the articulation of the wave-like behaviour of the musical material. In order to evoke a sense of 'flying over a landscape,' it would

[139] Ibid. Quotation from accompanying booklet.

seem appropriate to have the waves travel from front to back, thus suggesting that the direction of 'travel' is forwards with respect to the direction in which audience members would normally be facing. In this way, the different features of the aural landscape will emerge out of the distance in front of the audience, gradually approach, passing around and over the audience before disappearing into the distance behind.

In the context of this interpretation, the bass drone with which the movement begins can be regarded as the 'horizon,' a static and dominating presence against which movement takes place. Accordingly, the bass drone should be diffused in a way that establishes it as static and omnipresent: we have not yet begun our journey over the aural landscape, and the task of the diffuser at this point is to hint at the imposing grandeur of what we are yet to experience. The nature of the material inherently lends itself to this kind of treatment: psychoacoustically, lower frequencies are more difficult to localise spatially than higher frequencies and therefore – even diffused via mains and wides only – the bass drone is likely to sound as though it is coming 'from everywhere' to a certain extent. The impact of the bass drone could be augmented by the use of bass bins to emphasise the very lowest frequencies. These will be the most difficult to localise, thus reinforcing the sense of infinite space. As mentioned earlier, the correct fader positions for the diffusion of the bass drone should already have been prepared in time for the final impulse that occurs at 4'39" in the previous movement. In this way, a smooth transition – avoiding any clumsy fader movements as the drone emerges – is assured.

When the first wave begins, the gradual emergence of the second, semi-pitched, drone could be used very effectively to establish 'forward motion' over a landscape as the dominating gestural statement of the movement. Because this drone is somewhat noisy in spectral profile, so it will be inherently easier to localise spatially. If loudspeaker pairs are introduced carefully to coincide with its arrival and implied trajectory, therefore, it will be possible to establish the onset of this first wave as spatially independent from the accompanying bass drone (which is, recall, more difficult to localise anyway). The first wave should be diffused first to the very-distant pair of loudspeakers, very gradually moving forwards towards the audience via distants, mains, and wides as it builds in dynamic level. The point at which the first wave passes the audience is at around 1'38", and this is quickly followed by the emergence of a second wave on the horizon. The white-noise material that fades in and out between 1'38" and 1'48" should coincide with the introduction of tweeter trees, first the front pair, then the rear. As these are located directly above the audience, this would very effectively articulate the moment at which the first wave passes by. Use of the roofs and sides just prior to this would also be very effective in evoking a sense of the wave filling the space before shrinking into the distance behind the audience.

In terms of the fader movements involved in the realisation of this extended gesture, the progression from very-distants through to wides is straightforward because the corresponding fader pairs (7/8, 9/10, 11/12, 13/14) are located sequentially from left to right on the mixer. Furthermore, the onset of the wave is very gradual, and therefore only very slow fader movements will be required. At this point, however, the diffusion becomes logistically more complicated. As the crest of the wave passes the audience, the consecutive use of front and then rear tweeter trees is required, and these are located towards the extreme left of the mixing console on fader pairs 3/4 and 5/6. This is to be accompanied by the progression of the wave into the sides (19/20) and roofs (21/22), which are located at the opposite end of the console. Simultaneously, the mains and wides – 11/12 and 13/14 – should be gradually lowered, bringing the wave into the centre of the auditorium. When the wave passes, all of the fader levels should be gradually lowered further, with those located furthest forward in the hall being lowered first. This should coincide with the gradual introduction of rears on faders 15/16. Thus, the realisation of a (perceptually) relatively simple front-to-rear articulation is complicated to a certain extent by the physical routing of loudspeaker pairs to mixing desk faders. This problem is noted by Harrison:

> I've done a lot of things in France [...] where they tend to start at one end of the mixer with the speakers that are the furthest away from you in front, and then [gestures from left to right] the next pair, the next pair, the next pair, and this end [gestures to the right] is at the back. That's absolutely tickety-boo if you want to do lots of front-to-back or back-to-front things, but if you want to get the most significant speakers near the audience to be going [makes antiphonal sounds and hand gestures], you'll find that there's a pair there, a pair there, a pair there, and a pair there [indicates four completely separate areas of an imaginary mixing desk]; you can't get that interaction, which is why we don't do that in BEAST. We put those main speakers in a group on the faders so that you can do it easily. But then of course it means that if you want to go front-to-back, it's awkward. So, nothing's perfect! Unless you could suddenly re-assign the faders mid-diffusion. *You* [referring to the M2 Diffusion System, to be described in Chapter 5] could do that, couldn't you? [Laughs].

While the approach and passing of the first wave (and, to a certain extent, the second) is clearly a 'solo' event – nothing else happens concurrently – subsequent waves arrive in an increasingly overlapping fashion. As one wave is disappearing into the rear loudspeakers, the second wave should already be building in the very-distant and distant pairs and this, of course, compounds the challenges inherent in diffusing this movement effectively. The task of the diffuser is to convey a sense of perpetual forward motion by carefully timing the introduction of loudspeaker pairs – sequentially from the front to the rear of the hall – to coincide with the spectromorphology of each wave. In contrast with the previous movement, there will be no sudden fader movements here. Textural layers emerge progressively and seamlessly, and this should be articulated by the subtle, fluid, and perpetual cross-fading of loudspeaker pairs. This is relatively straightforward when the corresponding fader pairs are adjacent on the mixing console, but where this is not the case, difficulties arise. Clearly, a degree of compromise and prioritisation of the perceptual results will be necessary if the wave-like behaviour of the musical material is to be conveyed successfully. In this respect, the points at which each 'crest' passes the audience will be the most important and in most cases these can be diffused as described previously, with the tweeter trees articulating the points of transition from 'in front' to 'behind.'

In terms of perceived dynamic level, the first three minutes of the movement are, basically, an extended *crescendo*. This is accompanied by a gradual increase in gestural motion and intensity, reaching a peak at around 2'58", where three consecutively more emphatic waves occur in quick succession. The bass drone has long since disappeared (at around 1'40"), indicating that we have lost our awareness of the horizon and become completely immersed in our experience of the landscape. This gradual build could be emphasised in diffusion by the progressive use of an increasing number of loudspeaker pairs, such that we become more and more immersed within the landscape. By 3'11", most of the loudspeakers will be in simultaneous use, with only very subtle shifts in emphasis between loudspeaker pairs being used to articulate the slowly lapping, wave-like, motions. Of course, with every additional loudspeaker the diffuser has an additional fader to manipulate. Basically, this means that the degree of accuracy with which diffusion actions can be executed is inversely proportional to the number of loudspeakers involved. Accordingly, passages involving the use of most of the loudspeakers in the array are likely to suffer the most, the physical constraints of the system and the limitations of the diffuser's dexterity being most obvious under such circumstances.

From 3'11" onwards, the gestures become steadily more extended and meandering, although the dynamic level tails off less abruptly. When the gestural movement begins to abate, this need not be immediately accompanied by the use of fewer loudspeakers. Although the musical material becomes somewhat more drawn-out, reverberant, and slightly lower in dynamic level – almost as though we are beginning to regard the landscape from an increasing distance – the landscape is still large and dominating; it is not yet 'disappearing' as such. Although we are increasing our distance from the surface of the landscape (as implied by the blurring of spectral detail and increased reverberance), we begin to realise that it extends infinitely around us in all directions and is therefore, if anything, getting larger. This effect is intensified by the reintroduction of the bass drone 'horizon' (at around 3'30"; this should coincide with the reintroduction of bass bins), as we gradually become aware once again of our static point of reference. Throughout this passage it would probably be effective to let the material 'play itself,' with most if not all of the loudspeakers remaining in use. Extensive use of the roofs and sides would prove particularly effective throughout this passage. The perceived effect here might be something akin to vertigo, or agoraphobia.

The final wave passes the audience at around 4'05". This point of departure is particularly crucial because it is so exposed. If the sensation of perpetual motion that has been so central to the articulation of the musical material is to be finally affirmed, it is especially important that this final wave disappears into the distance behind the audience, leaving behind the omni-present bass drone with which the movement began. Because we are currently using most of the loudspeakers in the array, this will involve careful timing in the gradual dropping out of loudspeaker pairs to coincide with the implied trajectory of this final wave. In terms of fader movements, 7/8 and 9/10 (very-distant and distant) should be the first to drop out, followed by 11/12 and 13/14 (main and wide). As was the case with the introduction of the very first wave, this should not be problematic as these faders are arranged sequentially from left to right on the mixing console. When the final wave passes, front-to-back movement should be emphasised by dropping out the front pair of tweeter trees first (3/4), followed by the rear (5/6). This could be accompanied by a slight emphasis on the sides and roofs (19/20,.21/22), which should subsequently be lowered to affect the transition into the rears. Meanwhile, preparations should also have been made for the underpinning bass drone to be re-established in the mains, wides, and bass bins. The point at which the wave actually passes the audience should be a sufficiently distracting juncture at which to perform these fader movements. Again, we can see that the performance of a seemingly straightforward front-to-back articulation is, in reality, rather convoluted in terms of the actual fader movements involved. Given that the articulation involves the introduction (or dropping out) of loudspeaker pairs in a very specific order, this would seem to be an area in which optimisations could be made.

When the final wave disappears, we return once again to a position of stasis articulated by the low-pitched bass drone. As the bass drone fades out (4'25" to 4'34") loudspeaker pairs can gradually be dropped out one by one, perhaps ending the movement on mains and bass bins only (we normally perceive the horizon, after all, as being in front of us). In the context of this interpretation, it would be inappropriate for the bass drone to disappear into the distance, as such, because its function has been to act as the static background against which movement takes place. A sense of 'shrinking' – or perhaps 'focussing' into the main pair of loudspeakers – as the drone fades into silence would seem perfectly legitimate, however.

3.5. Coherent Audio Source Sets (CASS)

A coherent audio source set (CASS) is a single entity consisting of a number of discrete but mutually dependent encoded audio streams, which are linked by a conditional coherent bond.

As discussed in section 1.6.2, multiple encoded audio streams can be used co-operatively to encode spatio-auditory attributes with respect to sound source directionality. Two-channel stereo is a good example. Here, two discrete encoded audio streams are designated 'left' and 'right.' (The designation is essentially arbitrary: there is no real reason why the two channels could not represent 'up' and 'down,' for example). From this point onwards, the two audio channels, although physically discrete, are fundamentally interrelated. Collectively, they abstractly represent an imaginary axis that extends, 'spatially,' from left to right. If, at any given moment, correlated audio signals of equal amplitude are present in both channels simultaneously, this represents a sound source positioned at the exact centre of this imaginary axis. A signal in the 'left' channel only represents a sound source positioned at the left-most extremity of the axis, while a signal in the 'right' channel only represents a sound source positioned at the right-most extremity. Intermediate positions – assuming that the two signals are sufficiently correlated – are represented by varying ratios of signal amplitude in each of the two channels.[140] Because it is the *ratio* of the amplitudes of the two component channels that abstractly represents spatial positioning along an imaginary axis, so no representation of this spatial attribute exists unless both channels are present. In this respect, although both channels are technically independent, each is meaningless without the other and this is one respect in which a coherent bond exists between the two encoded audio channels. It therefore makes absolute sense that the two channels, in combination, should be regarded – and indeed treated, where appropriate – as a single entity. The concept of the coherent audio source set exists to make this possible.

Why is the coherent bond between the constituent channels of a CASS conditional? One condition, as discussed, is that all of the channels are present at all times. If this condition is not met, then that which is abstractly represented by the set as a whole (spatial location in the present example) is lost. Another important condition, however, concerns the decoding of the CASS channels. It is important to note (again, using stereo as a simple example) that the amplitude ratio of the two channels of a stereophonic CASS is an abstract representation of spatial positioning along an *imaginary* axis. The *real* axis will not exist until such time as the encoded audio signals are decoded via reproduction over loudspeakers. The real axis must be facilitated by the positioning of loudspeakers: in this case, positioning one loudspeaker to the left of the other would seem appropriate. In the stereophonic CASS we have a hypothetical representation of 'left' and 'right,' and intermediate points. Now we have a real 'left' and a real 'right,' which are differentiated in terms of real, physical, space. If the channel arbitrarily designated 'left' in the CASS is decoded via the loudspeaker that is physically positioned to the left of the other loudspeaker, and the 'right' CASS channel is decoded by the other loudspeaker, which is physically positioned to the right (this is, of course, another condition of the coherent

[140] This account of amplitude-based phantom imaging techniques is clearly simplified, but is nonetheless adequate for the purposes of the present discussion.

bond), then – and only then – will we have a real spatio-auditory phenomenon that accurately reflects the data encoded within the CASS. In other words, this is where the spatial (or any other) representation contained within the CASS is concretised. To clarify, the primary condition of the coherent bond between CASS channels is that they are decoded via an appropriate coherent loudspeaker set; this will be discussed more fully in section 3.7.

In and of itself, however, a CASS is simply a conceptual framework that allows multiple encoded audio streams (whether analogue or digital, static or transitory) to be arbitrarily grouped together; nothing more. This grouping implies that the component channels should be treated, collectively, as a single homogeneous entity, thus maintaining a constant relationship between them. In doing so, anything that is represented collectively by the constituent channels (whether this is spatial information, as in the previous example, or any other attribute) can be preserved. Some diagrammatic examples are given in Figure 4, below.

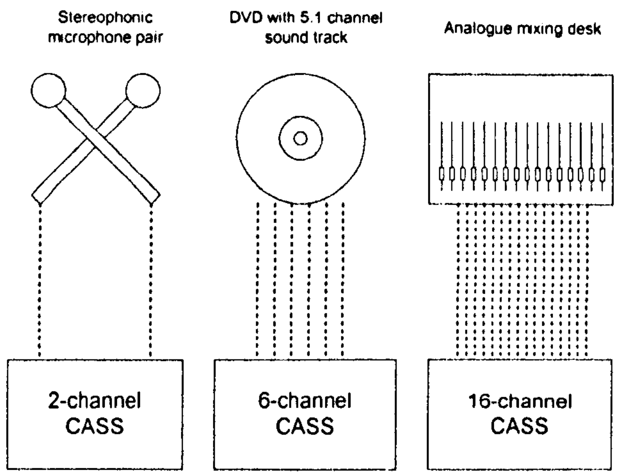

Figure 4. Some examples of multiple discrete encoded audio streams grouped into coherent audio source sets.

In the examples given above, all of the audio channels available within a given medium are grouped together into a single CASS. This procedure is fairly common in dealing with fixed medium electroacoustic works. It would be usual, for instance, for the two audio channels available from a compact disc to be conceptually grouped together and treated as a single, two-channel, CASS. Similarly, in octaphonic works, it is common for the eight channels encoded onto, say, an ADAT tape, to collectively represent a single, homogeneous, octaphonic spatio-auditory image.[141] It is critical to state at this point, however, that a coherent audio source set need not simply be defined as the total number of audio channels available on any given medium. In a personal interview with the author[142], Harrison described his octaphonic tape piece *Rock 'n' Roll*, in which the eight discrete channels present on the recording medium are subdivided into *two* coherent audio source sets: one stereophonic and one six-channel. The piece therefore comprises one coherent two-channel image, and one coherent six-channel image. The integrity of the stereo image will be lost if it is not treated as a CASS with its 'own internal atomic cohesion' (to reiterate the Clozier citation given in section 2.5), and the same can be said of the second, six-channel, image. In diffusing this piece one would expect that attempts should be made to maintain the integrity of both coherent audio source sets, but that the precise relationship between the two need not necessarily be so strict.

Truax describes a comparable approach:

> I would like to suggest that the multiple-channel system can be understood as an extension of stereo practice. Eight-channel tape, for instance, can be thought of as four contrapuntal stereo layers... This can be done [...] with multiple channel inputs where each soundtrack can be kept discrete and projected independently of all others.[143]

Here, Truax is effectively suggesting that, rather than regard the eight channels of an octaphonic fixed medium as one single CASS, one might compose four independent two-channel sets. Each two channel CASS represents a homogeneous stereophonic image in its own right; the *temporal*

[141] Some examples of the various coherent loudspeaker sets used to decode such CASSs will be described later, in section 3.7. For the time being, we will focus only on the CASSs themselves.
[142] Harrison (2004). "An Interview with Professor Jonty Harrison". Personal interview with the author, 8th September 2004. See Appendix 1 for full transcription.
[143] Truax (1998). "Composition and Diffusion - Space in Sound in Space". *Organised Sound*, 3(2): 145, 141.

relationship between the four stereo sets is fixed onto the medium, but no spatial relationship is defined at this stage. Harrison's and Truax's respective subdivisions of the eight available fixed medium channels into coherent audio source sets, along with some further arbitrary examples, are illustrated in Figure 5, below.

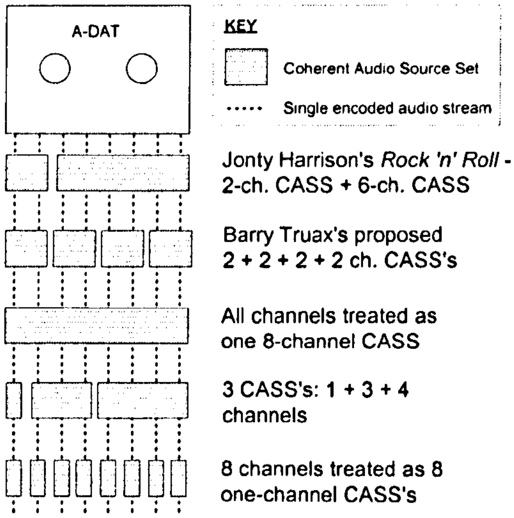

Figure 5. Some arbitrary examples of the subdivision of eight ADAT audio channels into coherent audio source sets.

These examples serve to demonstrate the true value of the expression 'coherent audio source set,' insofar as coherent audio source sets can be independent of the number of discrete audio channels supported by a fixed (or any other) medium. Nonetheless it should be restated that (at present) the majority of fixed-medium works do indeed treat all of the available channels as one homogeneous CASS.

Of course a coherent audio source set need not consist of audio channels encoded onto a fixed medium at all: the concept applies to *all* encoded audio streams, whether static or transitory, analogue or digital (see section 1.6). Consider the case of an instrumentalist, whose performance is being encoded via a near-coincident stereophonic pair of microphones. The two transitory-analogue encoded signals – one for each microphone – can (and probably should) be treated as a coherent audio source set.

A coherent audio source set may also comprise arbitrarily grouped encoded audio channels from a number of different sources. Arguably, this could be said to be the case if two transitory microphone signals are collectively regarded as a single stereophonic CASS, as illustrated in Figure 4, above. Alternatively, the sources of CASS channels could be more disparate. For instance, the of total four encoded audio channels from two compact discs could be conglomerated into a single, four-channel, CASS. In this case, the CASS might also be said to possess one level of hierarchy: two stereophonic (sub-)CASSs aggregated into one four-channel CASS. Such possibilities will be described more fully in the next section.

3.6. CASS Attributes

Having introduced the concept of the CASS, we can now summarise some of its key characteristics:

3.6.1. Consists of an Arbitrary Number of Discrete Encoded Audio Streams Treated Collectively as a Single Entity

A CASS consists of an arbitrary number of discrete but fundamentally inter-related encoded audio streams. Very often this number will remain constant, for any given CASS, throughout the duration of a work, but this does not necessarily have to be the case. It would theoretically be possible, for instance, for a CASS to be composed of, say, six fixed medium channels and then, at some point later in the work, be augmented to eight channels. Although the author is unaware of any instances of this at the time of writing it is, at least, a theoretical possibility. Notwithstanding this, in many (most) cases a CASS will comprise a *fixed* number of audio channels.

3.6.2. Independent of the Number of Encoded Audio Streams provided by any Single Audio Source

In section 1.6 it was observed that electroacoustic music deals almost exclusively in audio streams that are statically or transitorily encoded.

The two encoded signals from a stereophonic microphone could justifiably be regarded as a single stereophonic CASS, thus defining a constant relationship between the constituent channels; this is fairly intuitive. What is perhaps less intuitive is that, conceptually, a CASS need not consist of the sum total of independent audio channels from any single 'audio source' (such as a microphone, or multichannel ADAT) and may therefore consist of less, or more, channels than this. Some examples of the former case were discussed in section 3.5: the eight channels of an ADAT tape could be conceptually grouped into four two-channel CASSs, each of course containing *fewer* channels than the total available on the (in this case, fixed medium) source. For a CASS to contain *more* channels than the number available from any single source, it would clearly have to be composed of audio channels from more than one source, and the 'nested' CASSs might therefore be regarded as hierarchically subordinate to the 'aggregated' CASS. This latter case will be further clarified in section 3.6.6.

3.6.3. Constituent Channels Contain Phantom Imaging

One of the most important defining characteristics of a CASS is that phantom imaging techniques may be present (read: almost certainly *will* be present) across its constituent channels. The relationship between these channels is therefore not arbitrary, but constant. It is primarily for this reason that, in the vast majority of cases, all of the channels of a CASS must be present at all times during reproduction if the coherent bond between them is to be maintained to any reasonable degree.

3.6.4. Order of Channels is Spatially Determinate

Another characteristic of the CASS – strongly related to phantom imaging – is that, typically, it abstractly defines a hypothetical 'sounding space' which is delineated by the constituent audio channels. To use the amplitude panning method of phantom imaging, it would be true to say that signal in one CASS channel abstractly denotes sound at a *different* location from that implied by signal in another channel. To put it another way, the distribution of signals across the channels of the CASS (and perhaps also the spectral characteristics of those signals, which would be more relevant to a binaural stereophonic CASS for example), has a *constant relationship* with the relative 'location' of the sound at any given moment upon reproduction. This is equivalent to the notion of 'order' in coherent loudspeaker sets, which will be described in section 3.8.2.

3.6.5. Reliance on a Specific Coherent Loudspeaker Set

Following on from the previous item in particular it is true to say that, ultimately, the efficacy of a CASS upon reproduction depends on the positional relationships between the loudspeakers representing each constituent channel being, more or less, constant and, to a certain extent, conforming to the notionally 'ideal' configuration for that particular CASS. This does *not* mean that the geometrical positions of the loudspeakers must be absolute, but rather that, at any given moment during diffusion, the loudspeakers being used to recreate the CASS must be spatially arranged relative to each other in a formation that at least in some way resembles the relationship assumed by the CASS. Obviously what exactly constitutes a 'related' spatial formation in this context varies from composer to composer, but (for example) for a stereo CASS to be fully realised it would need to be broadcast over two loudspeakers between which effective phantom imaging could take place to some extent. This particular condition of the coherent bond between CASS channels demonstrates the close relationship between CASS and CLS, and will be discussed again in section 3.7.

3.6.6. Hierarchical Organisation

One aspect of the CASS model that is particularly interesting from the point of view of diffusion, is that it is conceptually hierarchical. That is to say, a CASS could theoretically be composed of two independent sub-CASSs. Consider the case of an electroacoustic work for two-channel tape and a live, stereo-miked, instrumentalist. Here, in the most obvious case, we are dealing with two independent, two-channel, coherent audio source sets. Each CASS exhibits phantom imaging across its constituent channels as well as all of the other conditions of the coherent bond; this state of affairs is represented pictorially in Figure 6.

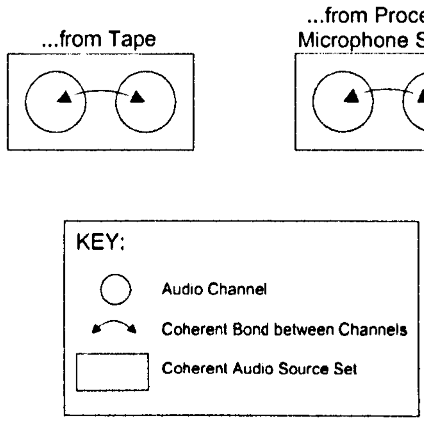

Figure 6. Graphical representation of two stereophonic CASSs in an electroacoustic work for processed instrument and tape.

Alternatively, the total four channels (compact disc plus two microphone signals) could effectively be treated as one single CASS. In doing so, a constant arbitrary relationship – or *relative bond* – between the two homogeneous stereo images would be defined: perhaps the left and right channels of each are mixed; perhaps the tape channels become the 'fronts' and the microphone channels the 'rears' of a quadraphonic-style setup; perhaps the two stereo images are mixed but the lefts and rights of each are reversed with respect to each other; the possibilities are numerous. In encapsulating these two CASSs within an aggregated CASS, the relative bond between the constituents would be maintained in diffusion, effectively treating the two notionally independent images as a single entity. This would result in a CASS with one level of hierarchy, as illustrated in Figure 7, below. This specific example would allow for combined tape and live processed material to be diffused as a single source and represents an example of a CASS containing more channels than those provided by any single medium/source.

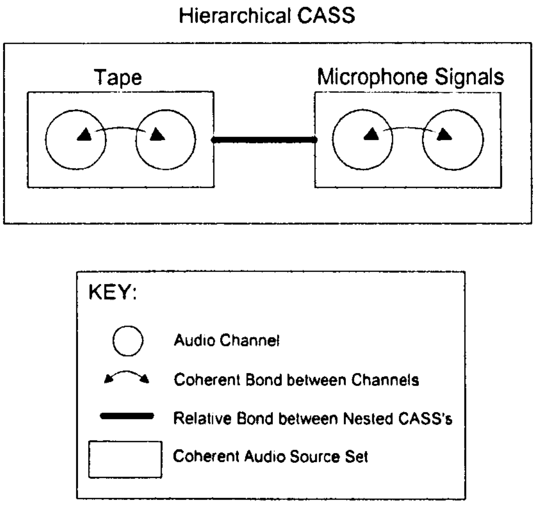

Figure 7. A CASS with one level of hierarchy, consisting of two nested CASSs linked by a constant relative bond.

Equally (although perhaps more bizarrely) the four channels in the previous example could be treated as one three-channel CASS and an independent monophonic CASS. Of course, this does not mean to say that this configuration would necessarily be appropriate, but merely that it is a theoretical possibility. Essentially, to summarise, the number and nature of the audio signals within a CASS is *arbitrary*, and therefore any collection of signals could, theoretically, be treated as a CASS. It is the responsibility of the performer (diffuser) to make appropriate decisions in this regard.

It seems reasonable to suggest that, under any circumstances, the existence of a coherent audio source set is the result of an intentional act, and that the set will have certain indigenous properties that are important to its effective realisation upon reproduction; such is the conditional bond between its constituent channels. A stereophonic electroacoustic work, as a specific two-channel CASS, will most likely contain subtly executed phantom imaging that has been carefully prepared in the studio. Similarly, an engineer recording a piano recital in the studio will take care over the precise positioning of microphones, monitoring the signal and making adjustments and refinements until the perfect stereo image is obtained. Such procedures are not arbitrary acts, but precise and intentional ones and, in the context of electroacoustic music performance, it is important that such coherent audio source sets are handled sensitively, and in a manner that affords respect to the craft of their creation: on this aspect most practitioners would agree. To give a crude example, it would probably be inappropriate to fail to reproduce the centre channel of a 5.1 channel CASS, or to arbitrarily present images intended for the surround speakers at the front and *vice versa* (of course some composers may not object as strongly as others to such an approach). It would be equally inappropriate to reproduce a tape piece intended for a standard 5.1 configuration over an array of loudspeakers positioned completely differently with respect to each other: the resulting phantom images would bear little or no resemblance to those intended by the composer.[144] Overall, it is true to say that the coherent bond between CASS channels is conditional because in order for the desired effect to be realised, or even approximated, the coherent audio source set must be *treated* as such, that is, it must be dealt with in a way that acknowledges the criteria on which its coherency and homogeneity rests. An important aspect of this concerns the relative positions of the loudspeakers used to decode the CASS. It is primarily in this respect that the notions of CASS and CLS are fundamentally interrelated.

3.7. Coherent Loudspeaker Sets (CLS)

A coherent loudspeaker set (CLS) is an arbitrary group of loudspeakers over which a CASS can be broadcast without unduly disrupting its coherency.

[144] Of course the notion of 'reproducing what the composer intended' is not quite as simple as this statement seems to imply, because – as discussed in Chapters 1 and 2 – composers have a propensity to intend very different things with respect to each other, and also tend to have very different ideas with respect to how their compositions should be realised in diffusion. We will return to this matter later.

Stereo is, once again, a simple and convenient exemplary paradigm. In a stereophonic CASS (that is to say, in most two-channel stereophonic recordings) there are two independent encoded audio streams. It is usual in such cases for the two independent channels to be used collectively to encode spatial attributes with respect to a single axis, as described in section 1.6.2: this is why, collectively, they constitute a CASS. Because the audio streams are encoded, it is assumed that – in order for the encoding to have served any useful purpose – they will at some point be decoded via the process of reproduction over loudspeakers. But in order for the intended spatial effect to be correctly recreated (i.e. for the coherency of the CASS to be preserved), some fairly specific conditions with respect to the number, and relative positioning, of loudspeakers, must be observed. If this is the case, then the group of loudspeakers in question can be regarded as a CLS. This is the crucial point of contact between CASS and CLS and represents the most important respect in which they are related.

In composing a stereophonic work, the composer is likely to have based his or her compositional actions, to some extent, on a certain set of assumptions with respect to how the CASS will be reproduced over loudspeakers in performance. In this case, two matched loudspeakers are probably assumed. These should be subtended at an ideal angle of between sixty[145] and one-hundred-and-twenty[146] degrees (depending on whose research is observed) with respect to the ideal listening location; certainly one loudspeaker should be positioned further to the 'left' than the other. Ideally the listener should be facing directly forwards and situated at a point in front of the loudspeakers and equidistant from each of them. The 'left' channel of the CASS should be fed to the left-most loudspeaker, and the 'right' channel to the right-most, and the two constituent signals of the CASS should be subjected to equal amounts of amplification (although some diffusers – particularly if they are top-down practitioners – might dispute this final point; this will be discussed later). These conditions collectively represent

[145] Malham (1998). "Approaches to Spatialisation". *Organised Sound*, 3(2): 174.
[146] Küpper (1998). "Analysis of the Spatial Parameter: Psychoacoustic Measurements in Sound Cupolas". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 190.

the criteria for a CLS that would be appropriate for the reproduction of a typical stereophonic CASS, as illustrated in Figure 8.

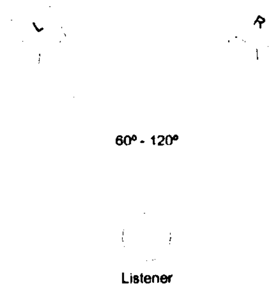

Figure 8. 'Ideal' CLS conditions with respect to maintaining typical stereophonic CASS coherency.

The list of CLS conditions for the successful reproduction of a stereophonic CASS could be augmented further still, but it is worth posing the question, 'What are the *most important* conditions?' This is, of course, a matter of opinion above all else. For certain, if all of the criteria described are met (and, it should be noted, this will probably only be possible in the case of a single ideally positioned listener) then the 'best' phantom imaging results will be achieved. But is it really necessary to meet *all* of these criteria? Before considering the answer to this question, some further CLS configurations will be considered.

Another CLS configuration in common use is the 5.1 array. In composing for such an array it is likely to have been assumed that five – ideally matched – loudspeakers and a sub-woofer will be made available during performance, that their relative positions with respect to the audience and to each other will more-or-less conform to the accepted standard for 5.1 sound reproduction – as illustrated in Figure 9 – and that the CASS signals will be routed appropriately to their corresponding loudspeaker outputs.

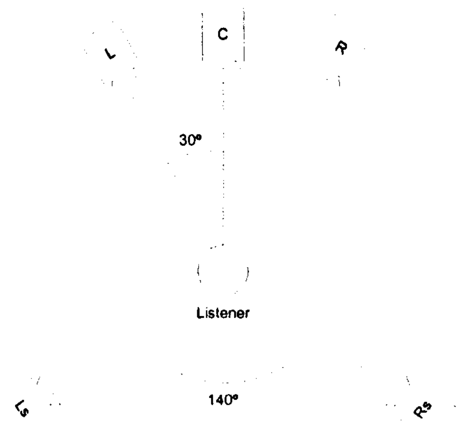

Figure 9. Relative loudspeaker positions in a standard 5.1 CLS (excluding sub-woofer).

Eight-channel (octaphonic) CLSs are also fairly common in electroacoustic music, although, in contrast with stereophonic and 5.1 configurations, there is less unanimity with regard to the relative loudspeaker positions in an octaphonic array. Harrison's octaphonic tape work *Streams* assumes the 'main eight' arrangement of loudspeakers as frequently deployed by Birmingham Electro-Acoustic Sound Theatre (BEAST) in concerts of electroacoustic music; this configuration is illustrated in Figure 10.

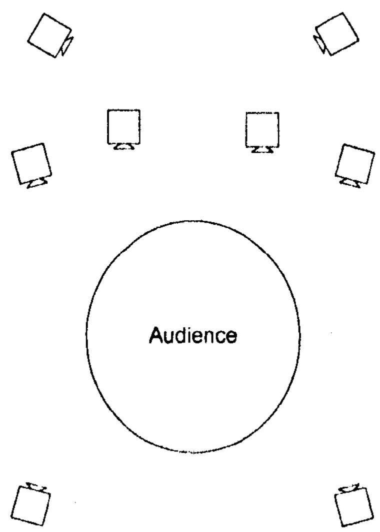

Figure 10. BEAST 'Main Eight' octaphonic loudspeaker configuration.

Wyatt states that 'three different loudspeaker positionings for eight-channel systems are in common use.'[147] As well as reiterating BEAST's 'main eight' octaphonic configuration, Wyatt describes two further possible arrangements of eight loudspeakers: one in a circular formation around the audience and one rectangular array. These configurations are shown in Figure 11 and Figure 12 respectively. Again, in order for the spatial imaging to remain notionally 'true' to what was composed in the studio (the notion of a theoretically 'true' reproduction will be discussed at length later) the loudspeaker positions and signal routings during performance must be the same as those assumed by the CASS.

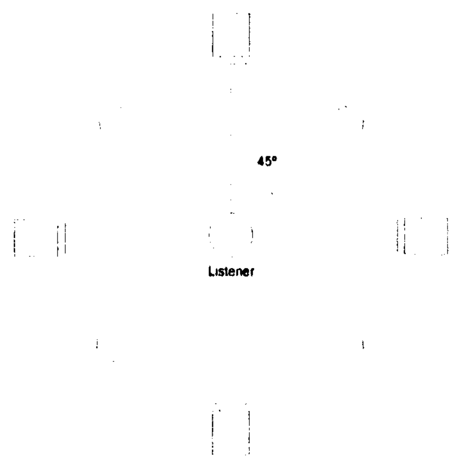

Figure 11. Circular octaphonic loudspeaker array as described by Wyatt.[148]

[147] Wyatt, *et al.* (1999). "Investigative Studies on Sound Diffusion / Projection". *EContact!* **2**(4). Electronic journal available at: http://cec.concordia.ca/econtact/Diffusion/Investigative.htm.
[148] Ibid.

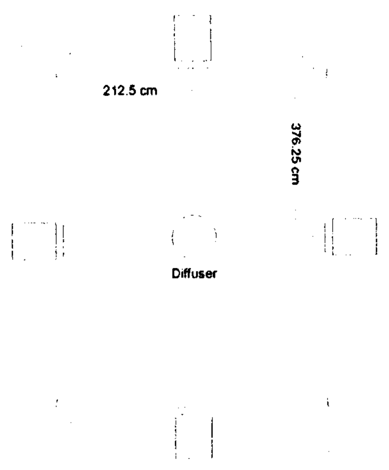

Figure 12. Rectangular octaphonic loudspeaker array as described by Wyatt.[149]

3.8. CLS Attributes

Let us now return to the question posed previously: what are the *absolutely indispensable* characteristics of a CLS? One could argue that – at the most basic level – the most important criteria are: that the number of loudspeakers contained within the CLS at least equals the number of channels contained within the CASS to be reproduced;[150] and that each loudspeaker is placed in a different spatial location from the others (it is, of course, physically impossible to negate this latter criterion). Realistically, however, such minimal criteria would probably be considered inadequate by most electroacoustic musicians.

---

[149] Ibid.

[150] There are exceptions to this generalisation. Ambisonic B-format (assuming first-order encoding) consists of four encoded audio streams that can be transcoded into an (essentially) arbitrary number loudspeaker feed signals. If the Ambisonic B-format is regarded as the CASS, then the number of loudspeakers will almost invariably exceed the number of channels contained within the (B-format) CASS. However, in order to be reproduced, the Ambisonic B-format must be transcoded into a number of signals suitable for decoding via a specific loudspeaker array, and in this case the number of loudspeakers must indeed be equal to the number of channels present in the (transcoded) CASS. In this respect, Ambisonic B-format can be viewed as a further abstraction of the CASS itself. Ambisonics aside, a case *could* (at the very least theoretically) be made for using a CLS with fewer channels than the CASS in question. This would, of course, result in the loss of a certain amount of spatial information, but could be regarded as equivalent to producing an alternative 'mix' of the electroacoustic work, designed for reproduction over a smaller number of loudspeakers. The author's own *Graffiti 2* (2003), for example, exists in both two-channel stereo and 5.1 versions.

3.8.1. Specific Shape

It therefore seems reasonable to suggest that there should also be a certain consistency with respect to the *relative positioning* of individual loudspeakers within a CLS. In the case of a stereophonic CLS it is proposed that the two loudspeakers should collectively delineate a single axis. The existence of this axis may be (and usually is) reinforced by training each of the loudspeakers on a single 'focal point,' normally located at some point along a second axis that is at right-angles to the centre of the 'main' axis: we will return to this notion later. In 5.1, the five full-range loudspeakers are not usually positioned arbitrarily, but such that they conform to a particular 'shape' (see Figure 9). In terms of its essential characteristics, this particular CLS formation can be seen to demarcate a plane (as opposed to a single axis) with a trapezium-shaped outline. It should also be noted that, within this trapezoid architecture, there are two 'sub-groups' of loudspeakers, which occupy the parallel sides: one group of three 'frontal' loudspeakers, and one group of two 'rear' loudspeakers. In Figure 9 all five full-range loudspeakers share a common, central, focal point, but it is also common for each of the two sub-groups to have focal points at different places along a central 'front-to-back' axis. In this case, it could be argued that the 'frontal' and 'rear' loudspeakers represent two hierarchically subordinate CLSs contained within one, hierarchically superior, CLS.

The 'shape' of a coherent loudspeaker set, therefore, refers to the physical positioning of the constituent loudspeakers, relative to each other, within a venue. The direction in which loudspeakers are facing can also be a criterion of the shape of a CLS, although this is less important than the relative positions. For example, in some cases a common 'focal point' will be desirable, while in other cases it may be less necessary. In Figure 9, the five full-range loudspeakers in the 5.1 CLS share a common focal point. In Figure 13 the loudspeakers do not share a common focal point, but they do still conform to the generally accepted shape of the 5.1 CLS. It is also important to note that the 'correct' shape for a CLS is determined by certain aspects of the coherent bond of the CASS that is to be decoded through it. It would, of course, also be true to say that the conditions of the CASS are formulated in response to the shape of the CLS that is used during the compositional process. Stereo and 5.1 configurations are convenient examples because they are conventional, but in a sense they are misleading as they suggest that the nature of coherent loudspeaker sets determines the nature of coherent audio source sets. In abstract (non-conventional) terms this is not the case: the relationship between CASS and CLS is one of co-dependency, and in this respect CASS and CLS define each other.

Figure 13. Five loudspeakers within a CLS. The loudspeakers do not have a common focal point but the CLS still conforms to the conventional 5.1 shape.

3.8.2. CASS Channels must be Appropriately Routed to CLS Loudspeakers

Another characteristic of both CLSs described so far is that the constituent loudspeakers exist in a spatially-differentiated *order*: in stereo, we have 'left' and 'right'; in 5.1 we have 'left,' 'right,' 'centre,' 'left-surround,' 'right-surround,' and 'LFE' (the latter of which is not, in fact, spatially differentiated: according to certain researchers, the rationale behind this is questionable[151]). In both cases the order is

[151] It is commonly stated that, below a certain threshold frequency, localisation of sound sources in terms of directionality becomes impossible. On this basis, it can be suggested

rationally connected to the relative positions of the individual loudspeakers relative to the overall CLS shape: the 'lefts' are positioned further '*to* the left' than the 'rights'; 'centre' is central in relation to 'left' and 'right,' and so on.

In section 3.6.4 it was stated that the *order* of the constituent channels of a CASS is spatially determinate: there is a constant relationship between the individual encoded streams that defines the intended spatio-auditory attributes on decoding. In order for these spatio-auditory attributes to be decoded successfully, the loudspeakers will need to be correctly positioned in relation to each other, and the individual CASS channels will need to be routed to the correct CLS loudspeakers. An exemplary description of the relationship between a stereophonic CASS, and a stereophonic CLS, was given at the beginning of section 3.7.

3.8.3. Reliance on a Specific Coherent Audio Source Set

The previous two sections articulate an important relationship: above all, a CLS is effectively meaningless in the absence of a CASS whose conditions match those provided by the CLS. The inverse, as described in section 3.6.5, is also true: a CASS is meaningless unless there is a CLS whose characteristics will allow for it to be decoded acceptably. In this respect the two concepts are co-defining and, therefore, depend on each other.

3.8.4. Multiple CLSs can Exist within a Single Loudspeaker Array

It is conceptually useful to differentiate coherent loudspeaker sets from loudspeaker *arrays*: the two expressions are not synonymous. A

---

that the spatial positioning of the LFE loudspeaker – because it only emits very low frequencies – is unimportant. According to certain sources, however, recent research implies that this may not, in fact, be the case:

Malham (2001). "Toward Reality Equivalence in Spatial Sound Diffusion". *Computer Music Journal*, **25**(4): 34.

loudspeaker array should be understood to mean the total number, and formation, of loudspeakers present within a sound diffusion system; a CLS can be regarded as a hierarchical subdivision of this. An appropriate example would be the 'main eight' loudspeaker *array* commonly utilised in the BEAST sound diffusion system; this is illustrated in Figure 14, below, and will be described further in section 3.12.1. It is true to state that this array was originally conceived to be appropriate for the decoding of *stereophonic* CASSs. Accordingly, the total number of loudspeakers in the array (eight) is sub-divided into four coherent loudspeaker sets consisting of two loudspeakers each: in Figure 14 these are labelled 'distant pair,' 'main pair,' 'wide pair,' and 'rear pair,' respectively. Notice that each of the CLSs is similar in shape, meaning that each individual CLS is appropriate for the coherent reproduction of a two-channel stereophonic CASS. This kind of practice will later be defined as 'pluriphony.'

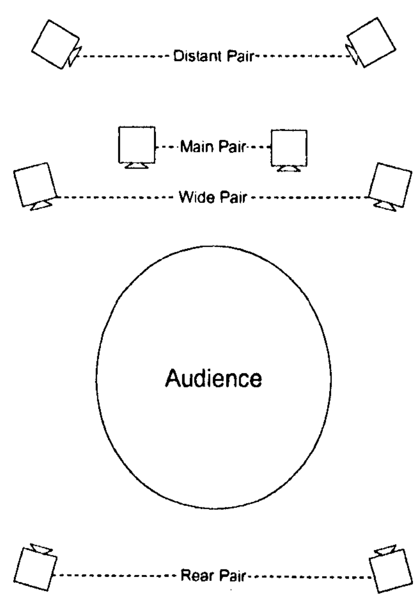

Figure 14. The BEAST 'Main Eight' array of loudspeakers, consisting of four two-speaker CLSs.

This does not mean that a loudspeaker array *has* to be subdivided into several independent CLSs: it could, indeed, be treated as a single CLS. If a stereophonic CASS is broadcast via a single stereophonic CLS, and there are no other loudspeakers present in the array, then the array represents a single CLS only. Similarly, if a 5.1 composition is reproduced over a standard 5.1 array, with only one loudspeaker present for each CASS channel, then the array, in its entirety, represents a single CLS that conforms to the conditions of the coherent bond between the constituent channels of a 5.1 CASS. These scenarios were illustrated in Figure 8 and Figure 9, respectively. As mentioned previously, Harrison's octaphonic work *Streams* treats the BEAST 'main eight' array as one single, eight-channel, CLS, as opposed to four groups of two.

3.8.5. A Single Loudspeaker can be a Member of Multiple CLSs

Because the coherent loudspeaker set is simply a conceptual grouping of loudspeakers not necessarily consisting of the total number of loudspeakers in an array, so any given loudspeaker within an array can theoretically be a member of multiple coherent loudspeaker sets. Figure 15 illustrates a circular eight-channel loudspeaker array[152] in which the front-centre loudspeaker is treated as a member of several arbitrarily chosen stereophonic CLSs.

[152] As proposed by Wyatt: see page 136.

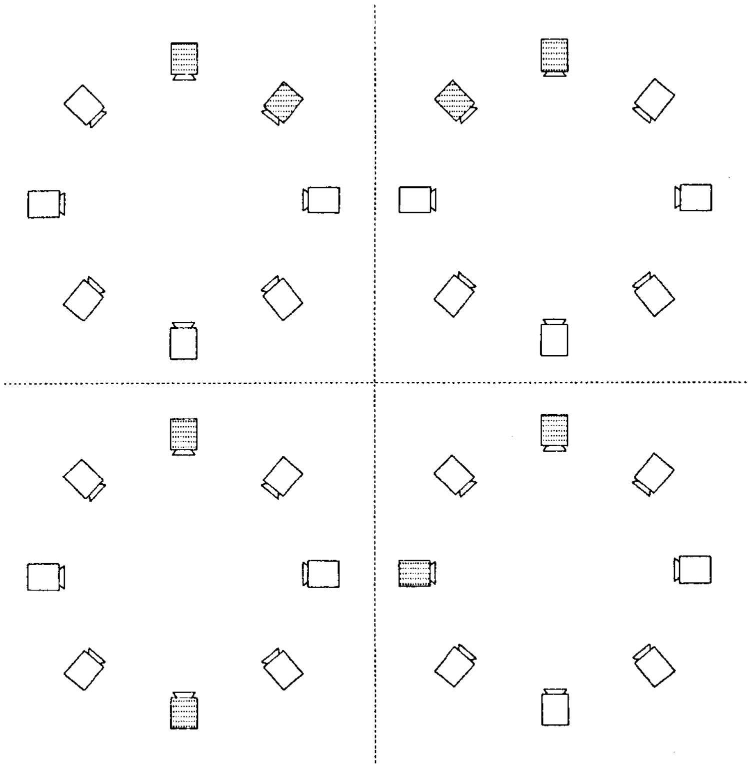

Figure 15. Four out of the theoretical fourteen possible stereophonic CLSs involving the centre-front loudspeaker in a circular octaphonic loudspeaker array.

In this case, each individual loudspeaker can be a member of seven different coherent loudspeaker sets, this number being doubled if inverted left-to-right audio source routings are considered. Whether or not all of these hypothetical CLSs would actually be appropriate for the accurate broadcast of stereophonic CASSs is, of course, a matter of debate, and may well vary depending on the piece being diffused and the attitudes of the diffuser.

3.8.6. Hierarchical Organisation

Related to the previous item, coherent loudspeaker sets can be organised hierarchically (and as described in section 3.6.6, this is also the case with coherent audio source sets). For instance, in defining a 5.1 CLS, one has conceptually grouped a number of independent loudspeakers together on the basis that they are likely to be utilised for the broadcast of several similarly grouped encoded audio channels. It would, theoretically, be possible to further define 'sub-groups' within that CLS; for example, the three frontal, and the two rear, loudspeakers. This particular example is illustrated in Figure 16 below.

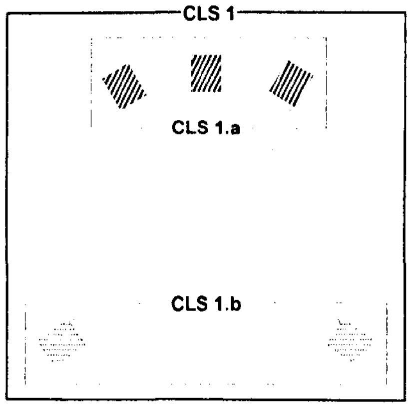

Figure 16. Five loudspeakers collectively constituting a CLS, within which there are two further sub-CLSs.

Such an approach could be useful, for example, if the piece in question treats the three frontal speakers and the two rear speakers as essentially separate multichannel images in some parts, but as a single five-channel image in others.

3.9. Sound Diffusion: A Re-Evaluation in terms of the Concepts of CASS and CLS

Models of the coherent audio source set and the coherent loudspeaker set have been proposed as useful tools in helping to define the relationship between audio sources (those technologies used to *en*code, manipulate, and transmit audio streams) and loudspeakers (those technologies used to *de*code audio streams). In mediating this particular relationship it seems reasonable to suggest that these concepts should be particularly useful within the context of sound diffusion, whose basic task is also to mediate between audio sources and audio decoders. Accordingly, it can now be proposed with more clarity that the process of sound diffusion involves the decoding of one or more coherent audio source sets via one or more coherent loudspeaker sets, with one of the principal aims being to *maintain* the coherency of the source set(s) throughout the process. As outlined in section 3.3, this objective is often complicated by acoustic issues. The remaining sections within the present chapter will be devoted to an examination of the differing approaches by which top-down and bottom-up practitioners seek to negotiate these difficulties as a means to achieving the ultimate objective of musical communication.

3.10. 'Diffusion' versus 'Panning'

It is a commonly held misconception that 'diffusion' and 'panning' (i.e. the act of artificially applying spatial attributes to encoded audio streams) are in some way synonymous. In an interview with the author, Harrison identified this misconception at a very early stage:

> JM: What does 'sound diffusion' mean to you?
>
> JH: What it means to me, is to do with not *just* space. That's the first thing to say. I think one of the problems is that when people talk about diffusion and spatialisation, it's as though it's something being *added* to the music, that's not already inherent in the music. To me, whether spatial, or any other thing that you apply to the material coming from CD or tape, it should be in the spirit of the music. So space, of course, comes into it, but it is not the only aspect. [...] The problem is that a lot of people think that diffusion and spatialisation are synonymous, and I don't think they are; I think they're different things.[153]

Here, Harrison touches upon what could be one of the fundamental reasons behind the hostility of bottom-up composers towards the top-down style of diffusion: the misconception that something is being inappropriately and arbitrarily 'added' to their carefully constructed music. In this context it may be useful, therefore, to differentiate between the predominantly performance-oriented practice of *diffusion*, and the conceptually different practice of *panning*.

A brief examination of the essential nature of 'panning' might lead one to propose that it represents part of the process of *constructing* a coherent audio source set. For example, a monophonic sound object might be placed and moved around within a stereophonic CASS using the pan-pots on a

[153] Harrison (2004). "An Interview with Professor Jonty Harrison". Personal interview with the author, 8th September 2004. See Appendix 1 for full transcription.

mixing desk, or their software equivalents. The final stereophonic CASS might ultimately comprise several monophonic sound objects panned across its two constituent channels in this way. By extension, a stereophonic source recording may be panned within an octaphonic CASS via a similar procedure. Indeed any number of source materials may be panned within the same CASS, in a process that basically entails placing or moving audio sources (the individual 'sound objects' that constitute the CASS) within a virtual space represented by a greater, or at least equal, number of channels. The process of diffusion, on the other hand, is more likely to be concerned with the presentation of one or more 'complete' coherent audio source sets already containing multiple 'panned' sound objects across their constituent channels. In this respect panning is a 'lower level' practice than diffusion.

This does not mean that the process of 'panning,' as such, is restricted to the confines of the studio: in certain performances of electroacoustic music, particularly where live audio sources are involved, an element of 'live panning' (as distinct from 'diffusion') may be entirely appropriate. Intuitively, and generally, it would seem to make more sense for singly identifiable sound sources to be 'panned,' and coherent audio source sets to be 'diffused.' As a hypothetical example, consider the case of an electroacoustic work for quadraphonic tape and (monophonic) synthesised sounds to be diffused over an array of twelve loudspeakers as illustrated in Figure 17, below. The tape part – a four-channel CASS – can be *diffused* over the twelve loudspeakers in such a way that its coherency is maintained throughout the performance. In doing so, it is likely that coherent sets of four loudspeakers will be used. Those loudspeakers highlighted in different colours in Figure 17 are appropriate examples. The monophonic signal from the synthesiser, however, can essentially be *panned* around the output channels of what is, effectively, a twelve-channel CASS, decoded in real-time via a single coherent loudspeaker set of twelve. The tape part is treated as a CASS, while the synthesiser part is treated more like a 'point source.' A scenario such as this demonstrates the usefulness of differentiating the processes of 'panning' and 'diffusion' as described.

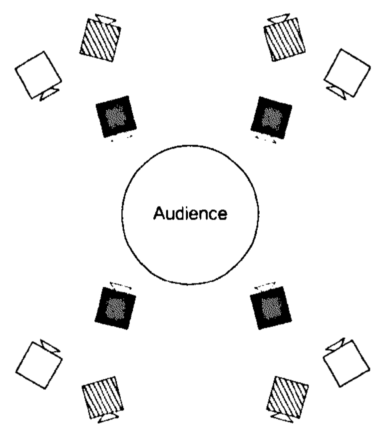

Figure 17. Hypothetical 12-speaker diffusion array, subdivided into three 'coherent loudspeaker sets' of four loudspeakers each. A four-channel CASS could be *diffused* via these three CLSs, while a monophonic source could be *panned* within a twelve-channel CASS represented by a single CLS consisting of all of the loudspeakers contained within the array.

Clearly there is there is the potential for a certain degree of cross-over between these two expressions. As previously mentioned, a stereophonic source could conceivably be 'panned' around, say, an eight-channel CASS, and in doing so one would expect that some attempts be made to maintain the original coherency of the stereophonic sound source within the CASS. This might be thought of as another example of the hierarchical nature of the CASS as described in section 3.6.6. Additionally the possibility of 'panning' – perhaps more readily identifiable as a *compositional* practice – as an appropriate aspect of the *diffusion* of certain electroacoustic works, represents a sense in which both composition and performance exist within a continuum: this will be more fully described in section 3.13.

3.11. Uniphony and Pluriphony

Roads *et al* use the expression 'pluriphony' to denote the act of diffusing a stereophonic audio source via *multiple* stereophonic loudspeaker pairs.[154] In

[154] Roads, Kuchera-Morin and Pope (2001). "Research on Spatial and Surround Sound at CREATE". Website available at: http://www.ccmrc.ucsb.edu/wp/SpatialSnd.2.pdf.

abstract terms, this concept can be extended and generalised as a practice in which a coherent audio source set is diffused via *multiple* appropriate coherent loudspeaker sets. The relationship between CASS and CLS – and between individual CASS channels and individual CLS channels – can therefore be described as 'one to many.'

Roads *et al* do not explicitly define pluriphony in terms of its opposite, and it is for this reason that the term 'uniphony' is proposed. In contrast with the pluriphonic approach, uniphony can be described as a practice in which a coherent audio source set is presented via a *single* appropriate coherent loudspeaker set. The relationship between CASS and CLS (and, again, between individual CASS channels and individual CLS channels) is therefore 'one to one.' A diagram clarifying the difference between pluriphony and uniphony, using stereo and 5.1 as examples, is given in Figure 18, below.

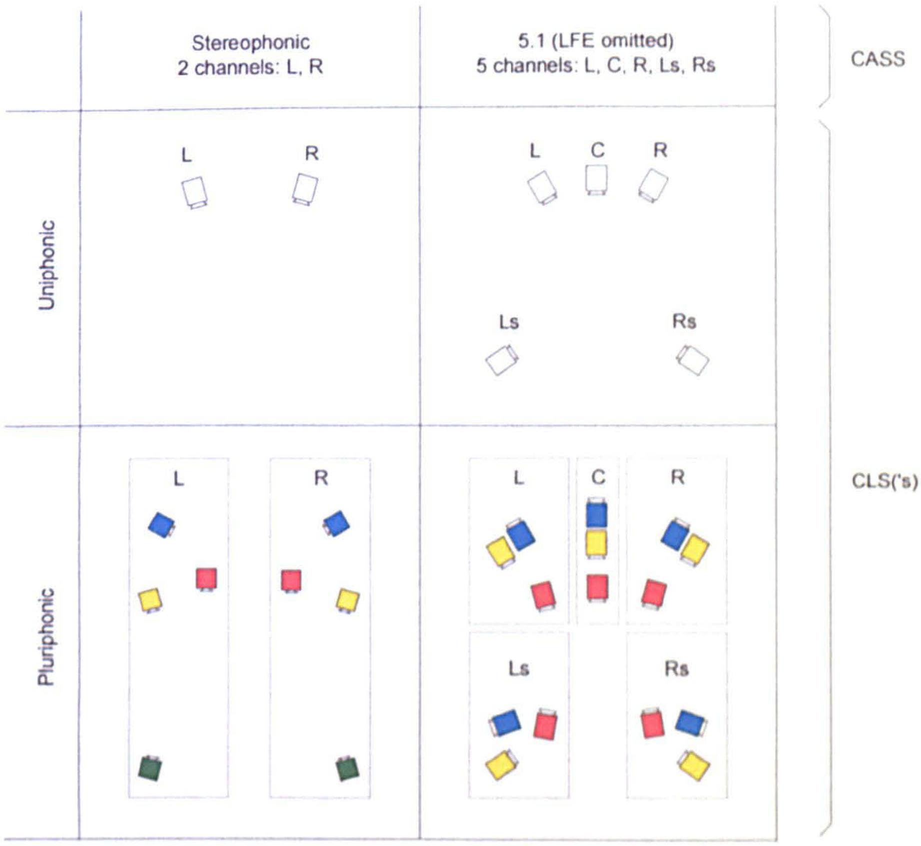

Figure 18. Diagram illustrating the difference between uniphonic and pluriphonic CASS-to-CLS relationships, using stereo and 5.1 (LFE channel omitted) as examples.

3.12. Top-Down and Bottom-Up Approaches to Sound Diffusion

At the end of section 3.2 an as yet unresolved question was posed: If sound diffusion is an *active* process whose fundamental aim is to ensure that an electroacoustic work is accurately communicated to an audience, then what exactly does this active process involve? What, specifically, *are* the acts? It can be deduced from the previous sections that maintaining the coherency of the CASS (in the broadest possible sense) is an issue *central* to the performance of electroacoustic music in diffusion. As noted in section 3.5, the successful broadcast of coherent audio source sets, even over the minimum number of loudspeakers (two for stereo, four for quadraphonic, *et cetera*), is dependent on a number of acoustic conditions. Most performance venues do not fully support these conditions and some of the reasons behind this were described in section 3.3. It is therefore clear that the public performance scenario seems, in many cases, to conspire against this central objective to a certain degree, and that *appropriate measures* must be taken to minimise the impact of this. It is proposed that these essential problems are generally recognised by most composers and practitioners regardless of their aesthetic directionality (top-down or bottom-up) and that it is in finding their ultimate solution that fundamental differences in approach become most obvious. The question therefore becomes, 'What manner of coherency are we attempting to preserve, and what might constitute appropriate measures to achieve this particular kind of coherency?' This is, essentially, the issue upon which the whole sound diffusion debate centres.

With certain conceptual frameworks in place, it can now be proposed that the exact nature of the 'act' of sound diffusion represents a continuation of the overall nature of the 'act' of composition itself. In Chapter 2 it was suggested that the compositional process can, broadly speaking, be characterised as either top-down or bottom-up. Accordingly, it is now proposed that the 'act' of diffusion can also be either top-down or bottom-up in nature. That is, the active measures taken to ensure the successful communication of the work are demonstrably congruent with the overall notions of either top-down or bottom-up practice. Much of the remainder of this chapter, and some of the following chapter, will be directed towards consolidating and exemplifying this assertion.

At this stage it is necessary to recall the continuum that exists between top-down and bottom-up models of compositional thought – Table 3 on page 84 is a useful refresher should one be necessary – and to consider what might constitute 'appropriate measures' from these two standpoints, and what might be the *nature* of the 'coherency' that each seeks to preserve. A considerable portion of electroacoustic work exists on stereophonic fixed media alone, and this also serves as a convenient and conceptually simple example through which to explain the basic premises of top-down and bottom-up attitudes towards sound diffusion. It is proposed that the same basic concepts, however, can be applied to electroacoustic music of any technological permutation.

3.12.1. The Top-Down Model

One possible solution to many of the problems outlined in section 3.3 is to use more than two loudspeakers, thereby theoretically presenting a less compromised auditory image of the stereophonic material by (amongst other things to be described presently) minimising the distance between audience members and loudspeakers. The use of multiple loudspeaker pairs also addresses the power issues described earlier. In abstract terms this solution entails diffusing a CASS over a number of loudspeakers that is *larger* than the number of audio channels that constitute the CASS, that is, pluriphonically. This kind of approach is advocated by the Birmingham Electro-Acoustic Sound Theatre (BEAST), whose 'main eight' configuration of loudspeakers (see Figure 14 on page 141) is described by Harrison as 'the absolute minimum for the playback of stereo tapes'.[155]

This circumstance raises the question, then, of how to present a stereo image across more than two loudspeakers. The top-down model of sound diffusion observes the fact that, outside of the controlled studio environment (which will most likely be relatively acoustically stable and populated by a single ideally positioned listener) acoustic detail will be lost. For example, dynamic subtlety that was obvious in the studio will be far less so in a larger, less acoustically ideal, performance venue. In response to this observation, Harrison states the following:

> The composer will have indicated relatively louder and quieter events. In performance I would, at the very least, advocate enhancing these dynamic strata – making the loud material louder and the quiet material quieter – and thus stretching out the dynamic range to be something nearer what the ear expects in a concert situation.[156]

Smalley concurs:

[155] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 121.
[156] Ibid.: 120-1.

> In a recorded format you can never achieve an ideal dynamic range that will suit all spaces and contexts; maybe it is not even ideal on two loudspeakers. And so you need to exaggerate or highlight the high end – lift the top levels up – and possibly drop the low levels down. Extending the dynamic range affects peoples' perceptions of the piece and permits and enhancing of the structural shape.[157]

Such practice already begins to address the question of how to utilise additional pairs of loudspeakers:

> For effects of distance (which on the original stereo tape are implied by careful balancing of amplitude and reverberation characteristics, and which are very susceptible to being swallowed by (actual) concert hall acoustics), it is useful to be able to move the sound from close to distant *in reality*, following the cue on the tape – hence the **distant** pair [of loudspeakers].[158]

Following the same logical process, it follows that a sweeping pan from the left to the right of the stereo field, although markedly obvious under ideal listening circumstances, may well seem like a disproportionately small gesture in a venue whose physical dimensions and acoustic characteristics are much larger. It therefore seems reasonable to exaggerate this gesture in diffusion, in order to properly convey the effect intended by the composer. Indeed any gesture that is notionally 'on the tape' may be lost in a large venue, and should therefore be exaggerated by the sound diffuser:

> If you have something which is [...] zapping around all over the stereo stage, then it seems to me perfectly legitimate to exaggerate that erratic behaviour over a much bigger loudspeaker system. Hence doing this kind of thing [makes short rapid hand movements] and wiggling the faders around in a way that will throw the sound all around the room, in an erratic manner. That seems to be perfectly in-keeping with the musical idea... Where you have a real sense of something going [makes a 'whooshing' sound and gesticulates from front to back], that's such a physical gesture that it's just *screaming* to be exaggerated! So you do it with the faders; you just 'do more.' [...] So inevitably [diffusion] grows from a need to reinstate, if you like, in a big space, that which is audible in a small space but would get lost in a big space.[159]

We might regard this as a process whose goal is to present musical material in a manner that is appropriate to the nature of the context, or

---

[157] Austin (2000). "Sound Diffusion in Composition and Performance - An Interview with Denis Smalley". *Computer Music Journal*, 24(2): 13.
[158] Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 121-122.
[159] Harrison (2004). "An Interview with Professor Jonty Harrison". Personal interview with the author, 8th September 2004. See Appendix 1 for full transcription.

as Smalley puts it, 'to expand the stereo image and to project it effectively in a large space.'[160] The very essence of top-down sound diffusion is that this process necessarily entails a subjective interpretation of the musical material. This is, of course, perfectly in-keeping with the ethos of top-down composition in general, which fundamentally bases its aesthetics on perceptual and subjective realities. Accordingly, proponents of the top-down model of sound diffusion tend to regard the practice, in the truest possible sense, as a continuation of the compositional process itself:

> Through analysis, familiarity and understanding of the work, an informed and experienced composer/diffusor/projectionist can present the diffused work as a continuation of the composer's musical intent in such a way to significantly expand the listening experience of that work.[161]

> Certain musics have such physicality already embodied in them, embedded in them, that not to continue that process into diffusion seems to be a travesty of the piece.[162]

As the compositional process in this case is effectively based on the premise of abstracting musical structure from 'concrete' materials via a process of perceptual evaluation, so the process of diffusion consists in a reiteration of this procedure. In diffusion, the work is treated as a concrete material, whose *subjective evaluation within a particular context* informs the way that the material is treated in performance. This can *only* happen via a process of interpretation.

Accordingly, aspects that might *increase the scope for interpretation* in the diffusion of electroacoustic music tend to be embraced by top-down composers, at least to a far greater extent than by their bottom-up counterparts. Loudspeakers that 'colour' the sound, for example, are actively endorsed by certain practitioners on this basis:

[160] Austin (2000). "Sound Diffusion in Composition and Performance - An Interview with Denis Smalley". *Computer Music Journal*, 24(2): 12.
[161] Wyatt, *et al.* (1999). "Investigative Studies on Sound Diffusion / Projection". *EContact!* 2(4). Electronic journal available at: http://cec.concordia.ca/econtact/Diffusion/Investigative.htm..
[162] Harrison (2004). "An Interview with Professor Jonty Harrison". Personal interview with the author, 8th September 2004. See Appendix 1 for full transcription.

> *Are we really all in agreement [...] that the diffusion instrument should be an ensemble of high fidelity loudspeakers, possessing linear response, and thus rather lacking in character? [...]* We are under enormous pressure to normalize so that compositions may be distributed with guaranteed conformity. And yet it is precisely those highly original diffusion instruments, made up of motley mixes of loudspeakers, that have given pleasure to so many of us.[163]

> One notices a certain rejection of the idea of an ensemble of high-fidelity rigorously homogenized loudspeakers, possessing a near-military precision of performance and behaviour in their devotion to the common cause of the composition. To this totalitarian concept of sound-projection, I prefer the high-infidelity of loudspeaker pairs that allow variable shading during the diffusion. To the autism of an ensemble of identical loudspeakers, I prefer the multiracial accents of a disparate gathering.[164]

Clearly the above quotations are indicative of a broader top-down view of electroacoustic music that extends beyond the mere specifics of its performance, and are also, characteristically, antagonistic towards the opposing bottom-up view (use of the word 'totalitarian' is particularly telling in this respect). In the broadest possible terms it can be stated that top-down approaches to sound diffusion tend to consist in adapting electroacoustic works to given performance spaces via a subjective process of interpretation. Often this involves the *pluriphonic* presentation of a coherent audio source set via multiple coherent loudspeakers sets. Some diffusion systems that embody the values discussed in this section will be described in the next chapter.

3.12.2. The Bottom-Up Model

Another approach to the issue of maintaining coherence in performance would be to tailor the *performance context* such that any undesirable colourations of the musical material are minimised. For example, if an overly reverberant venue interferes with the subtleties of the musical material, then measures can be taken to reduce the reverberation time, or even find a more appropriate, less reverberant, venue. If an audience's positioning adversely affects their perception of accurate phantom imaging, then steps can be made towards restricting the

---

[163] Boesch (1998). "Composition / Diffusion in Electroacoustic Music". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 221.
[164] Savouret (1998). "The Natures of Diffusion". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 347.

auditioning area so that (ideally) all listeners experience the same effect. (Calls for every member of the audience to be provided with an individual pair of headphones are not unknown).

Of course, this approach to the performance of electroacoustic music is dependent on a somewhat fixed notion of what a composition 'is,' and it should be clear that this standpoint follows on logically given what we know about the nature of bottom-up composition. If an electroacoustic work has been composed 'monolithically,' and particularly if there are notionally quantifiable relationships within its fabric, then it is of crucial importance that the 'truth' of the work be delivered to the audience in its pure, unadulterated form, to rearticulate a point raised in section 2.4. As discussed previously, this model of the musical composition is conceptually resistant to the notion of interpretation: here, there *is no interpretation*; the work is a closed, fixed entity; it can only be performed 'correctly,' or 'incorrectly,' with little margin for error. It is therefore perfectly understandable that composers inclined towards bottom-up models of composition and diffusion tend to be dubious about the methods of their top-down colleagues, (perhaps rightly) suspicious that such practice might destroy the precise relationships expressed by the musical material. This particular methodological conflict is described anecdotally, and very effectively, by Harrison, who quotes Jean Piché as follows:

> Phase-aligned fullband systems with enough power to fill larger spaces [...] neutralize positional phase-shift and offer a better rendition of the original compositional intent in the studio. [This] is certainly not perfect given the conditions of most [electroacoustic] concerts but at least it tries to provide a 'neutral' acoustical front to the audience where the music is of prime importance, not the 'artistry' of the diffuser... When I go to an electroacoustic concert, I want to hear the music in the best conditions possible as it was intended to be heard. I don't go to hear someone express himself on the sliders, when I know perfectly well that whatever is done there (with the possible exception of the composer him/herself) will be at best 'inspired' improvisation and at worst 'Bobby is loose on the sound system again...' Perhaps there will come a time when sensitive diffusion artistry can be codified, but for the time being, it seems more of a whim than 'sensitivity.'[165]

[165] Piché (March 1997) cited in Harrison (1998). "Sound, Space, Sculpture - Some Thoughts on the 'What,' 'How' and 'Why' of Sound Diffusion". *Organised Sound*, 3(2): 124.

An important aspect of the above quotation is Piché's use of the word 'neutralize,' as it epitomises a characteristically bottom-up attitude towards the diffusion of electroacoustic music, whereby the aim is to *neutralise* the performance context so that the truth of the music can be perceived transparently. Similarly the notion of interpretation, which top-down diffusers would consider an integral and essential part of performance diffusion, is frequently regarded as 'whim' by those of bottom-up persuasion. The reason for this is that interpretation, as such, can be regarded as a search for *subjective* truth, a notion which is generally not recognised by bottom-up composers with their objective aspirations but one which is central to the ethos of top-down composition. It therefore follows that Piché would have no quibble with top-down diffusion if it could be (his own words) 'codified,' and thus rendered in some way objective.

This argument, of course, does not solve the previously observed difficulties of accurately presenting a stereo image in a large and acoustically unstable performance venue. If bottom-up diffusers are unhappy with the idea of presenting their works pluriphonically (and in doing so 'letting Bobby loose on the sound system again'), then an alternative solution to this very real problem is required. This is perhaps why bottom-up composers seem to favour composition for multichannel fixed media. Like the top-down, pluriphonic, approach, this offers a solution to the problem of phantom images (*et cetera*) getting 'lost' in larger performance venues by providing larger numbers of loudspeakers. Unlike the top-down method, the multichannel solution (notionally) offers the composer complete control over the phantom imaging contained across the constituent channels. This uniphonic approach contrasts the tendency of top-down diffusers to prefer the pluriphonic approach as a method that offers more scope for interpretation.

There is also a case for stating that pluriphonic traditions have arisen, at least partially, as the result of technological circumstance. Historically, electroacoustic composers have, for the most part, been limited to two channels of audio stored on magnetic tape as the mode of delivery for compositions. Accordingly there were finite limitations with respect to the auditory results attainable. In the absence of fixed medium multichannel formats, composers were perhaps forced to rely on third party diffusers to facilitate the fully immersive sonic landscapes that they had envisaged, as well as to compensate for the dynamic range limitations of analogue recording media. It would perhaps be argued by the bottom-up composer that this practice is outmoded nowadays, with probably very close to one-hundred percent of all recently produced electroacoustic works having been realised and stored entirely in the digital domain, and given the increasing availability of multichannel capable software and storage media. In this context it can at least be *argued* that the convention of 'emphasising the contours' – be they dynamic or spatial – of works realised in the studio has out-lived the technical need to do so. Even Harrison – a self-declared ally of the acousmatic (top-down) school – admits that, from this purely technological perspective, certain aspects of top-down diffusion are 'less necessary' than has previously been the case.[166]

Of course top-down diffusers would also argue that, while matters may have been slightly improved with the advent of digital and multichannel technologies, it is still not possible to provide a 'definitive' version of a composition owing to the persistently variable acoustics of performance venues:

> I don't agree with the premise that you compose on four or eight speakers [...] and therefore all you need to do in the concert hall is replicate it. Because the point is, you *can't*. I think that's my basic problem. You *cannot* do it: it doesn't work. It may work *in theory*, and if the acoustic is sufficiently controlled, but how many halls have you been in that have that kind of controlled acoustic? Not very many.[167]

Notwithstanding this assertion, certain practitioners of the bottom-up school of thought remain dedicated to constructing 'standardised'

---

[166] Harrison (2004). "An Interview with Professor Jonty Harrison". Personal interview with the author, 8th September 2004. See Appendix 1 for full transcription.
[167] Ibid.

performance venues for the diffusion of electroacoustic music. Here, 'transparency' and 'neutrality' are often key considerations: high-quality, matched, phase-corrected arrays of loudspeakers; dry acoustics; carefully planned seating areas; and so on. Generally, variables that might detract from a transparent dissemination of the work – 'coloured' loudspeakers or whimsical 'interpretations' of the musical material for example – are to be avoided:

> It is certainly true that for many composers, intimate contact with the sound material of a work, the constant re-listening, refinement and steady progress to the realisation of the work's conception, *together with the possibility of its almost totally accurate reproduction*, excludes more and more the idea of an interpretation afterwards. [My italics.][168]

So, in summary, bottom-up approaches to sound diffusion tend to consist in adapting the performance circumstances to suit the nature of *given* electroacoustic works, most often via a process of objective neutralisation. This can involve minimising the reverberant characteristics of performance venues, and is often (but not always) characterised by a preference for the uniphonic presentation of one CASS via one CLS. Some specific systems will be discussed in the next chapter.

3.13. The Composition-Performance Continuum

Essentially what differentiates top-down and bottom-up approaches to the performance of electroacoustic music is whether actions are taken on the basis of the (perceptual) *subjective reality* of what a work sounds like when diffused over a particular system, in a particular hall, and to a particular audience (top-down), or on the basis of the notionally *objective reality* of what the piece '*is* like' (bottom-up). In the top-down case, the text (the piece) is submitted to the context (the venue, audience, *et cetera*) via a process of perceptual interpretation. In the bottom-up case, the context submits to the text, and this often involves carefully setting up an 'ideal' listening environment in which factors that will compromise the monolithic

[168] Tutschku (2002). "On the Interpretation of Multi-Channel Electroacoustic Works on Loudspeaker-Orchestras: Some Thoughts on the GRM-Acousmonium and BEAST". *SAN Journal of Electroacoustic Music*, **14**: 14-16.

nature of the music are minimised. Both approaches take measures to ensure that the coherency of the coherent audio source set (or sets) is communicated, but the exact nature of the measures taken differs markedly. In other words it can be argued that the ultimate goals of sound diffusion, in the most general terms possible, are 'universal,' but approaches differ according to which phenomena are regarded as 'given,' or 'absolute.' The bottom-up standpoint regards the *composition* as absolute, and seeks to address the issues of performance by bringing the context closer to it. The top-down view, essentially, regards the context as 'given' (at least with respect to any one specific performance) and attempts to move the composition 'towards' that context.

Stated simply, the bottom-up composer attempts to place the performance context within the scope of the work, while the top-down composer seeks to place the work within the scope of the context. This conceptual difference owes much to the writings of Savouret,[169] and is expressed diagrammatically in Figure 19.

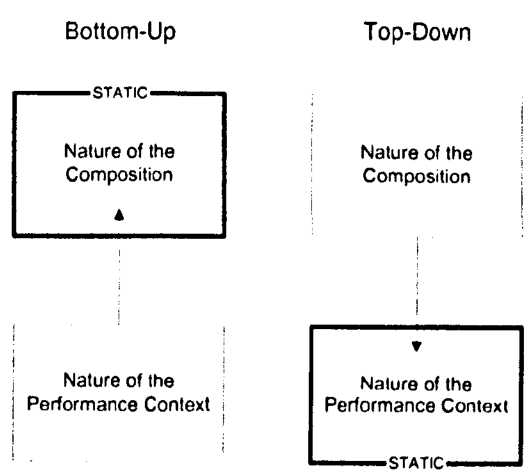

Figure 19. Simple diagrammatic illustration of the difference between bottom-up and top-down approaches to the diffusion of electroacoustic works.

[169] Savouret (1998). "The Natures of Diffusion". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music.*

An interesting phenomenon that supports this model can be found in differing attitudes towards the use of artificial reverberation in composition. Denis Smalley, for example, expresses a preference for producing compositions with very little artificial reverberation, based on the assumption that most performance venues will provide reverberant characteristics of their own and the work can therefore be diffused in a manner that takes advantage of this:

> In my pieces, I've not been keen on placing one reverberated space into another reverberated space, which is what can happen when taking your piece to a public space... That's probably the reason I don't use artificial reverberation...[170]

Others might tend to favour the use of artificial (and therefore controllable) reverberation in composition, intended for performance in a venue that is more acoustically *neutral*. In this case it would seem likely that proactive steps would need to be taken in order to ensure as little ambient reverberation in the venue as possible: Küpper's approach to diffusion via sound cupolas – which will later be described in section 4.12 – is a good example.

In both top-down and bottom-up cases the practice of sound diffusion can be regarded as something of a continuation of the compositional process itself. In bottom-up composition, 'musical structure is created by a process which is primarily the imposition of quantifiable values on fundamentally inert sound material.'[171] Similarly, in diffusion, the performance context must also be 'fundamentally inert' (or, as discussed, neutral) in order to transparently articulate the music. In top-down composition, the composer is guided by a perceptual evaluation of the materials within the (subjectively true) context of his or her perceptions, and this process is continued in the *context* of diffusion. It should therefore be clear that, in either case, it is not possible to define an obvious distinction between the processes of composition and performance. In both top-down and bottom-up cases, the

[170] Austin (2000). "Sound Diffusion in Composition and Performance - An Interview with Denis Smalley". *Computer Music Journal*, **24**(2): 16.
[171] Harrison (1999). "Keynote Address: Imaginary Space - Spaces in the Imagination". *Australasian Computer Music Conference* (Melbourne; 4-5 December 1999).

nature of the compositional approach taken, to a considerable extent, presupposes the nature of the performance, and *vice versa.* It can be proposed that sound diffusion, broadly speaking, is a process that seeks to publicly present coherent audio source sets in an effective manner via loudspeakers, such that the essential nature of the music is successfully relayed to the audience. This process is complicated because the 'essential nature' of electroacoustic works can be either top-down or bottom-up, and therefore the means of presenting them 'effectively' can also be either top-down or bottom-up. In general terms it can be seen that the conflict between these opposing attitudes is forced out into the open at the performance stage because, ultimately, attempts to reconcile these highly contrasting convictions simply cannot be deferred any further.

3.14. Review of Top-Down and Bottom-Up Techniques

In the briefest terms possible, it can be said that the process of sound diffusion might entail presenting a piece of electroacoustic music in a manner that suits the context, or alternatively tailoring the context itself in a way that suits the piece. Clozier's use of the expressions 'diffusion-interpretation' and 'diffusion-transmission'[172] represents an easily comprehensible means of expressing the difference between the two. 'Diffusion-transmission' suggests a passive process involving the 'transparent' communication of the musical work. 'Diffusion-interpretation,' contrastingly, infers that the performer plays a more active role in expressing the discourse of the work. These contrasting methods have here been described, respectively, as the top-down and bottom-up approaches to sound diffusion and are, in many cases, bound within a continuum that chronicles the development of an electroacoustic work from initial conception, through composition, to performance. That is to say, actions taken during the respective processes of composition and performance are often taken with respect to the same set of underlying beliefs. Table 4, below, reiterates the abstract defining characteristics of the top-down and bottom-up standpoints given in the previous chapter,

[172] Clozier (1998). "Composition, Diffusion and Interpretation in Electroacoustic Music". In Barrière and Bennett [Eds.] *Composition / Diffusion in Electroacoustic Music*: 235.

additionally appending certain traits specific to the approaches to the task of sound diffusion adopted by each of these groups.

Top-Down	**Bottom-Up**
The overall ethos is…	
Human	Super-human
Subjective	Objective
Composed to be interpreted/ appropriated	Composed to be realised/disseminated
Realist (pragmatic)	Idealist
Plural	Absolute (monolithic)
Relativistic	Deterministic
Corporeal (physical)	Cerebral (intellectual)
Empirical/perceptual	Logical/conceptual
Qualitative	Quantitative
Phenomenological	Rational
Built/invented	Constructed/created
Organic	Architectonic
Abstracted forms	Abstract forms
Text submits to context	Context submits to text
…which indicates that…	
Encoded audio streams are regarded as abstractions of the *perceptual* qualities of real auditory events	Encoded audio streams are regarded as abstractions of the *conceptual* structures that define real auditory events
...and therefore the techniques of sound diffusion are more likely to involve…	
'Coloured' loudspeakers	'Transparent' loudspeakers
Diffusion-interpretation	Diffusion-transmission
(Sensitively) improvised diffusion	Pre-planned or automated diffusion
Variable performance contexts	Controlled performance contexts
Preference for pluriphonic approach	Preference for uniphonic approach
Maximising interpretative scope	Minimising interpretative scope

Table 4. Revised criteria of the Top-Down/Bottom-Up model, with items relevant specifically to the practice of sound diffusion appended.

Once again it should be pointed out that these criteria are, essentially, gross generalisations, and that in reality opinions with respect to sound diffusion are never so neatly divided. Not *all* top-down practitioners, for example, would agree with *all* of the attributes designated as 'top-down.' Nonetheless, in generalising in this way it is possible to identify certain patterns that seem to have emerged in the performance practice of electroacoustic music, and this gives an overall indication of the kind of dichotomy we are dealing with. It would be more reasonable to suggest that most practitioners recognise the relative advantages and disadvantages of both standpoints, nonetheless ultimately expressing some degree of preference for one or the other.

If works from across the technological and aesthetic spectrum of the electroacoustic idiom are to be successfully performed side-by-side in concerts, then some kind of solution that – somehow – reconciles these two highly contrasting philosophies and methodologies and is acceptable to both top-down and bottom-up practitioners is required. Chapter 4 will focus, in part, on the evaluation of existing sound diffusion systems on this basis.

3.15. Summary

The present chapter has focused mainly on two things: the definition of 'sound diffusion' in abstract terms; and the attitudes of electroacoustic musicians with respect to this practice. Concepts of the 'coherent audio source set' (CASS) and the 'coherent loudspeaker set' (CLS) have been proposed as useful tools in defining the practice of sound diffusion from a technological perspective; it is also proposed that these concepts will prove useful in the design of future sound diffusion systems.

It can be concluded that sound diffusion involves the presentation of one or more coherent audio source sets, via one or more coherent loudspeaker sets, to an audience in a public performance situation. Ideally this should be done in such a way that the discourse of the musical work is effectively communicated to the audience. This central objective is, basically, unanimous among electroacoustic musicians but its realisation in practice is complicated by the fact that, acoustically, it is difficult to facilitate a completely 'accurate' decoding of encoded audio streams in most performance venues (as described in section 3.3). It is for this reason that sound diffusion must be an *active* process.

The central objective of sound diffusion is further complicated by the fact that top-down and bottom-up practitioners fundamentally disagree with regard to what exactly an 'accurate reproduction' of a work is. This is owing to the fact that the top-down practitioner regards the encoded audio stream as an abstraction of the *perceptual* qualities of real auditory events, whereas the bottom-up practitioner regards it as an abstraction of the *conceptual* structures that define real auditory events. Accordingly, the ways in which composers/performers seek to achieve this central objective are essentially two-fold, and can broadly be categorised as top-down and bottom-up, respectively. Usually, top-down composers demonstrate a preference for top-down approaches to sound diffusion, while bottom-up composers tend to prefer bottom-up methods. Some of the general characteristics of top-down and bottom-up approaches to electroacoustic music and sound diffusion were summarised in Table 4. In the most basic possible terms, top-down diffusion methods are motivated, broadly, by the subjective reality of sound as *per*ceived by the diffuser, while bottom-up methods are based on the objective reality of the work in question, as *con*ceived by the composer. In this respect the two approaches differ fundamentally in terms of the nature of what they are trying to communicate. In other words, top-down and bottom-up practitioners have highly contrasting views with regard to the nature and purpose of sound diffusion itself and, therefore, with respect to the role of sound diffusion systems in the performance of electroacoustic music.

In both top-down and bottom-up cases, the approach to sound diffusion represents a continuation of the essential nature of the compositional process itself. This being the case, the hypothesis is that sound diffusion systems will themselves be divisible into those that are broadly intended for the communication of works in an essentially top-down manner, and those intended for the communication of works in an essentially bottom-up manner. This bifurcate approach is, however, problematic given that top-down and bottom-up works are reasonably likely to be juxtaposed in concert programmes. Ideally, a sound diffusion system should be able to facilitate the presentation of both top-down and bottom-up works, and will therefore be required to accommodate both top-down and bottom-up attitudes and diffusion methods. In combination with those technological observations made in Chapter 1, this proposition will form the basis for a system of criteria designed for the evaluation of existing sound diffusion systems, which is to be the central focus of the next chapter.