The Audiovisual Object

Andrew Connor

PhD Creative Music Practice

The University of Edinburgh

2017

I certify that:

a) this thesis has been composed solely by me;
b) the work is my own;
c) the work has not been submitted for any other degree or professional qualification, and;
d) any included publications are my own work, except where indicated throughout the thesis and summarised and clearly identified on the declarations page of the thesis

Signed:

Date:

Abstract

Abstract

The ‘audiovisual object’ is a fusion of sound object and visual object to create an identifiable perceptual phenomenon, which can be treated as a ‘building block’ in the creation of audiovisual work based primarily on electroacoustic composition practice and techniques. This thesis explores how the audiovisual object can be defined and identified in existing works, and offers an examination of how it can be used as a compositional tool. The historical development of the form and the effect of the performance venue on audience immersion is also explored.

The audiovisual object concept builds upon theories of electroacoustic composition and film sound design. The audiovisual object is defined in relation to existing concepts of the sound object and visual object, while synaesthesia and cross-modal perception are examined to show how the relationship between sound and vision in the audiovisual object can be strengthened.

Electroacoustic composition and animation both developed through technological advances, either the manipulation of recorded sounds, or the manipulation of drawn/photographed objects. The key stages in development of techniques and theories in both disciplines are examined and compared against each other, highlighting correlations and contrasts.

The physical space where the audiovisual composition is performed also has a bearing on how the work is perceived and received. Current standard performance spaces include acousmatic concert systems, which emphasize the audio aspect over the visual, and the cinema, which focuses on the visual. Spaces which afford a much higher level of envelopment in the work include hemispheric projection, while individual experience through virtual reality systems could become a key platform. The key elements of the audiovisual object, interaction between objects and their successful use in audiovisual compositions are also investigated in a series of case studies. Specific audiovisual works are examined to highlight techniques to create successful audiovisual objects and interactions.

As this research degree is in creative practice, a portfolio of 4 composed works is also included, with production notes explaining the inspiration behind and symbolism within each work, along with the practical techniques employed in their creation. The basis for each work is a short electroacoustic composition which has then been developed with abstract 3D CGI animation into an audiovisual composition, demonstrating the development of my own practice as well as exploring the concept of the audiovisual object.

The concept of the audiovisual object draws together existing theories concerning the sound object, visual perception, and phenomenology. The concept, the associated investigation of how audiovisual compositions have evolved over time, and the analysis and critique of case studies based on this central concept contribute both theory and creative practice principles to this form of artistic creativity. This thesis forms a basis for approaching the creative process both as a creator and critic, and opens up a research pathway for further investigation.

Lay Summary

Lay Summary

When composing music, there are building blocks of notes and musical phrases. In film sound design, there are also building blocks of sound effects and sound textures. These can be described as sound objects. Animators start with a blank space and add visual objects as background, textures and foreground characters. Sound and vision combine to create audiovisual objects, which can be combined to create background and foreground, textures and characters, in a composition which is immersive and engaging.

Audiovisual objects and the environment in which they can be combined can be defined using theories that address how we perceive sound and vision as specific phenomena. The thesis makes use of these theories to provide a clear definition of the audiovisual object and its environment. The composition techniques used in the practical part of the degree concentrate on abstract soundscapes combined with animation, and the history of how both have developed are examined to establish contrasts and similarities. The space in which the work is performed also has an important effect on how audiovisual compositions are experienced, and the pros and cons of concert halls, cinemas, and other venues are considered.

Examples of successful audiovisual compositions are described and examined to show how the audiovisual object is central to their success. As this research degree is in creative practice, a portfolio of composed works is also included, with production notes explaining the inspiration behind and symbolism contained within each work, along with the practical techniques employed in their creation.

Dedication and Acknowledgements

Dedication

With many thanks for her support and advice over the course of this degree, this thesis is dedicated to Susan Graham. Also, in memory of my mother, who always believed completely in my ability to do anything I wanted to do, the thesis is also dedicated to Josephine Connor.

Acknowledgements

This degree has been many years in the undertaking, and there are many people who have helped over those years. Acknowledgements and heartfelt thanks are owed to my supervisors: Professor Raymond MacDonald, Dr. Robert Dow, Jared Taylor, Dr. Martin Parker, Dr. Tamara Trodd and Professor John Young.

Many thanks also go to those who were of great support throughout: Louise Rossiter, Pete Stollery, Alastair MacDonald, Juliette MacDonald, Andrew Hill, Andrew Willy, Andrew Dolphin, Ben Ramsay, Bret Battey, Trevor Wishart, Peter Nelson, Sean Williams, Jules Rawlinson, Ninian Dunnett, Lauren Hayes, Christos Michalakos, Karen Loomis, Matt Collins, Harry Whalley, Matt Brennan, Tami Gadir, Tom Western, Athos Tsiopani, Alan Mason, Rachel Everitt, Kevin Hay, Michael Mullin, Neil Kempsell, Donald Bell, Martin Tweddell, David Kyle, Chris Hall, Steve Cantwell, Abigail Lamb, and all the Animation students at ECA, past and present.

Finally, my deepest thanks for all their help and support go to my father, John Connor, and sister, Heather Wilson.

Table of Contents

Contents

Chapter 1

Introduction

Introduction to Thesis

This thesis is divided into four main chapters, exploring the audiovisual object and the environment within which it exists. The 'audiovisual object' is a fusion of sound object and visual object which creates an identifiable perceptual phenomenon. The audiovisual object can be treated as a 'building block' in the creation of audiovisual work based primarily on electroacoustic composition practice and techniques. This thesis explores how the audiovisual object can be defined and identified in existing works, and offers an examination of how it can be used as a compositional tool. The historical development of the form and the effect of the performance venue on audience immersion is also explored. Case studies are also given, with several audiovisual works examined in the context of the audiovisual object and its properties, and how these are exploited successfully within the works. As the degree is in Creative Music Practice, appendices have been included which contain production notes for the works included in the accompanying portfolio. While these appendices lie outwith the main thesis itself, they serve to support the use of the concepts covered in the thesis within my own practice, and explore the use of theory and inspiration behind the works composed over the course of my research and practice.

Chapter 2, *The Audiovisual Object*, examines the key concept of the audiovisual object, which builds upon theories of electroacoustic composition and film sound design. The audiovisual object is defined in relation to existing concepts of the sound object and visual object, while synaesthesia and cross-modal perception are examined to show how the relationship between sound and vision in the audiovisual object can be strengthened.

The chapter establishes the general idea of the audiovisual object, which is based on the key theories concerning the sound object, originally defined by Pierre Schaeffer in 1966. The sound object reflects a phenomenological reduction when approaching the listening process, defining the sound perceived as an individual identifiable phenomenon. The development of analogue and later digital sound recording technology has allowed the manipulation and processing of sound objects, and their use as building blocks in electroacoustic composition, in the same way as notes and phrases are used in classical music composition. The chapter also explores the visual object, establishing it as a perceptible phenomenon similar to the sound object – the image, drawn or photographed, can also be recorded, manipulated and reproduced, and then used as a building block in visual composition.

The chapter outlines the use of Gestalt theories of perception in defining sound objects and visual objects, examining the key principles of proximity, similarity, common fate, good continuation, closure, surroundedness, emergence, re-ification, multistability and invariance. The idea of exclusive allocation, where the boundary between objects outlines and defines them, and places them firmly in either the foreground or background, is also explored. The materiality of colour, and how it can be incorporated into the audiovisual object, is also explored.

The audiovisual object is a consequence of the fusion of one or more sound objects with one or more visual objects, also representing an individual, identifiable phenomenon. The perception of combined perceptual streams is explored, including the crossover of perceptual qualities between two streams as occurs in synaesthesia. The propensity of proper synaesthetes to show high consistency in correlations between perceptual streams offers an approach to audiovisual composition where limits and boundaries can be established, creating an artistic space within which to work. Similarly, the idea of cross-modal perception is explored, as perceptual streams are normally perceived in association with each other. The results of combining perceptual streams, and the resulting enhancement, dominance, or modulation of sound and vision are explored, as is the emergence of a cohesive audiovisual object from the combination. The emergence of specific audiovisual objects can be reinforced by the pre-cueing of the audience, including elements of the emergent combination in earlier parts of the composition.

The chapter also explores the temporal and spatial aspects of the audiovisual object and the reinforcement between sound and vision that can be achieved. The phenomenon of synchresis is explored, where fortuitous matches between sound and vision can be achieved, and its relationship with planned synchronisation between the two perceptual streams. A continuum matching disorganised sound and vision through to highly organised sound and vision is developed to include the effects they have on audience appreciation and the aesthetics of the compositional process. The audiovisual object exists within an audiovisual environment, and the ideas of background and foreground, texture and gesture are also examined.

Chapter 3, *The Development of the Audiovisual Object Palette*, addresses how electroacoustic composition and animation both developed through technological advances, either the manipulation of recorded sounds, or the manipulation of drawn/photographed objects. The key stages in development of techniques and theories in both disciplines are examined and compared against each other, highlighting correlations and contrasts.

The chapter first addresses how the development of technology has affected the progress of sound recording, manipulation and playback, with particular regard to the development of electroacoustic composition. The origins of this discipline come from Russolo and his manifesto on how noise and sound should also be regarded as part of music. The innovations afforded by electricity and electromagnetical forms of recording sound on tape allowed Schaeffer to develop his theories on reduced listening and the sound object, leading to *musique concrète* and the manipulation of recorded sounds in his Parisian institute. At the same time, in Cologne, Eimert and Stockhausen explored the use of electrical technology to both record sound and synthesize completely new material, using experimental analogue synthesizers to create their *elektronischen musik*. Currently, electroacoustic composers make use of both approaches, manipulating recorded sounds along with digital synthesis to achieve their compositions. This organisation of sound follows one of two paths, according to the composer and theorist Trevor Wishart – exploring the idea behind the sounds, or the heard relationships between the sounds.

The chapter continues with the exploration of the origin of animation, starting with the earliest films and the use of stop-motion effects. The rise of drawn animation was marked by Emile Cohl's *Fantasmagorie* in Paris in 1908 and Winsor McCay's *Little Nemo* in the USA in 1911. In Europe, the artistic environment of the avant-garde encouraged the development of auteur-led animation, which also led to the rise of visual music, whose principal proponent was Oskar Fischinger. His work explored the close correlation that could be achieved between visual composition of colour and movement and musical compositions. While the work produced by such auteurs explored the mutability of the medium, they were quickly supplanted by the rise of animation studios, where the auteur was replaced with a lead animator and a supporting team, in what became known as the sweatshop model. This also replaced the artistic approach of the individual animator with the overall direction and philosophy of the studio heads. The studio headed by Walt Disney, by virtue of early adoption of sound and engaging recurring anthropomorphic characters, established a dominance in the industry. His particular philosophy and adherence to 'real-world' physics in his animations became the standard due to his commercial success. While auteur animators continued to a greater extent in Europe, the post World War I depression and the advent of World War II severely limited the opportunities to develop their craft. Experimentation still took place outside the main studios, with experimental work by the Whitney brothers in California inspiring special effects in mainstream films, such as the Star Gate sequence in *2001AD: A Space Odyssey* (1968).

The development of both electroacoustic composition and animation through technological advances also resulted in correspondences between the two disciplines. In both, the composer or artist has complete control over the environment, starting with a completely clean slate. The chapter outlines Paul Wells' examination of orthodox and experimental animation, noting his comparison of experimental animation with a fluid 'musicality' of approach rather than the more structured dialogue driven approach of orthodox work. Both disciplines also make use of spatial positioning of constituent elements, creating a perspective of layers from background to foreground, which shift and adjust over time, establishing further similarities. This importance of correlation between sound and image is also acknowledged in the film industry, with professional sound designers championing the importance of both sound and vision in creating a fully integrated experience.

In Chapter 4, *The Performance Space and Audiovisual Immersion*, the physical space where the audiovisual composition is performed is discussed, as it also has a bearing on how the work is perceived and received. Current standard performance spaces include acousmatic concert systems, which emphasize the audio aspect over the visual, and the cinema, which focuses on the visual. Spaces which afford a much higher level of envelopment in the work include hemispheric projection, while individual experience through virtual reality systems could become a key platform.

The chapter establishes the early explorations of achieving a created immersive experience, starting with Wagner's concept of *gesamtkunstwerk*, a combination of inter-relating theatrical, sound and design elements making use of technology to achieve his vision of an immersive operatic theatre environment. His ideas were further developed within the Bauhaus in the early 20th century, with the architect Walther Gropius expanding on Wagner's principles in his designs for a *Totaltheater*, where projected images would surround the audience, adding to the spectacle on stage, and with acoustics designed to expand the sound to fill the auditorium, a concept that has resonances with the modern IMAX cinema and planetariums.

The chapter outlines the advantages and disadvantages of various performance space formats. Electroacoustic composers who venture into audiovisual work commonly stage their compositions within the electroacoustic concert format, where the sonic space surrounding the audience can be completely controlled. However, the small flat front screen commonly used in these settings counteracts this expansion of sound and focuses the attention too closely on a narrow focal point. As an alternative, cinema in its basic form offers a space that the audience associates more with the idea of an audiovisual performance in its joint aspects. However, in a simple cinematic space, the sound is not optimised to achieve the immersiveness sought after by the audiovisual composer, and the flat front screen again focuses the attention in too strict a direction. The limitations of a simple cinema screen have been ameliorated by the introduction of surround sound, with 5.1 surround sound now an accepted standard, expected in most cinemas. This allows a more diffuse sonic environment.

A development on the simple cinematic screen is the use of multiple screens, although these are most commonly encountered only as part of an art installation. A common form is that of a triptych, with a central screen between two left and right screens, angled to slightly enclose the audience. The enclosure effect adds to a feeling of immersiveness, but the discontinuity between screens can detract from this effect. More recent innovations in the cinematic form have seen the development of large curved screens, filling the forward view of the audience and creating a more immediately immersive environment. IMAX cinemas in particular have established this immersive format, although the cost of the technology involved places this beyond the financial capacity of most audiovisual composers. Another parallel development to the cinema is the planetarium, which makes use of hemispheric projection and multi-speaker arrays to overwhelm the sight and sound of the audience members. While this format is still primarily used for astronomical displays and informative shows, planetariums have also successfully been used as venues for audiovisual concerts, such as the Vortex concert series by Jordan Belson in the USA between 1957 and 1960. The planetarium may perhaps be the most attractive group audience format available for the audiovisual composer. The audiovisual composer may also combine their compositions with other artforms in an interdisciplinary collaboration, which can provide a further method of creating an immersive experience.

The chapter concludes with an examination of current digital technology advances, where virtual reality may allow a bespoke immersive environment for the audiovisual composer to exploit. From the creation of virtual worlds such as *Second Life* that can be explored through the computer screen, the most recent advances in screen and sound technology have led to individual virtual reality headsets such as *Oculus Rift* and *Project Morpheus*. While the driving force for this technology has been the creation of a platform for immersive interactive computer games, they also offer a new platform for audiovisual composers.

In Chapter 5, *Case Studies*, the key elements of the audiovisual object, interaction between objects and their successful use in audiovisual compositions are also investigated in a series of case studies. Specific audiovisual works are examined to highlight techniques to create successful audiovisual objects and interactions.

Three pieces are explored in depth to establish how synchresis and synchronisation have been used to create a satisfactory aesthetic experience, as well examining the works with reference to the Gestalt principles outlined in Chapter 2. The use of pre-cueing and anticipation in increasing the impact of a climactic audiovisual object is also explored. Each piece exhibits different aspects of the properties of audiovisual objects discussed elsewhere in the thesis, and so offer a good set of practical examples of how the close interaction between sound and vision, necessary to the successful creation of audiovisual objects, can be achieved.

Finally, Chapter 6, *Conclusions*, provides a summary of the main elements examined in the course of the thesis, and offers a discussion of how the future development of audiovisual composition may take shape, with particular interest in current developments in digital technology and the impact these will have on the composition process and the environments within which the works will be experienced.

As this research degree is in creative practice, a portfolio of 4 composed works is also included, with appendices containing production notes explaining the inspiration behind and symbolism within each work, along with the practical techniques employed in their creation. The basis for each work is a short electroacoustic composition which has then been developed with abstract 3D CGI animation into an audiovisual composition, demonstrating the development of my own practice as well as exploring the concept of the audiovisual object.

Chapter 2

The Audiovisual Object

Introduction

The key concept of my research is the 'audiovisual object', a fusion of sound object and visual object to create an identifiable perceptual phenomenon, which can be treated as a 'building block' in the creation of audiovisual work based primarily on electroacoustic composition practice and techniques. The audiovisual object can be combined with others, and with both sound objects and visual objects, within a composed audiovisual environment, to create an immersive, perceptually engaging composition.

The concept of the audiovisual object derives directly from that of the sound object, which in electroacoustic composition is a basic element, equivalent to a note or phrase in classical music. By combining sound objects together, an electroacoustic composition is formed. The audiovisual object is a development from the sound object, and the techniques and methods of electroacoustic composition can also be mapped on to the composition of audiovisual works, combining these techniques with those used by sound designers and video editors. Developments in technology now mean that audio and visual manipulation can be carried out on a much faster and more accessible scale than heretofore, allowing an individual to create both a sonic and visual composition, with the aim of tying them together in one holistic aesthetic work. My research and practice investigates the techniques, concepts and methods available in both audio and video composition, and how they can be combined to form a set of compositional tools for the audiovisual composer.

Pierre Schaeffer (1910-1995), a pioneer of electroacoustic and acousmatic composition, set out his definition of the 'sound object', further refined by Chion in his commentary of Schaeffer's work – the *sound object* is a separate entity from the *object creating the sound*. This is more apparent given modern technology, where sounds are easily reproducible away from the original source. In the 19th century, when you walked along a road, and heard a dog's bark, you would expect to turn around and see a dog. In the 21st century, the sound of a dog's bark could be produced by a dog, or by a mobile phone, an advert from a local shop, or from a radio or television programme overheard through an open window. We now have the technology to record, play back and manipulate sounds, making the sonic environment a much richer and more creative place in which to manipulate sound objects. The development of sound synthesis has also reinforced the concept of the sound object – it is now possible to create an analogue of the dog's bark, without ever having to make an original recording of a real dog barking.

The sound object is in effect the basic 'building block' of electroacoustic composition. By taking the sound object, either as it is, or processed in some way to highlight an aspect of its spectromorphology or spatial aspects that the composer wants to explore, it can be combined with other sound objects. This may result in another perceptible sound object, or a sequence consisting of several discernible objects. By combining these objects creatively and in an aesthetically satisfying way, an electroacoustic composition takes shape.

The visual object can be compared to the sound object as another discrete phenomenon that is self-contained and perceived by the viewer as an individual identifiable object. There have been a variety of approaches to defining the visual object – a widely used philosophical approach to defining an object (visual, auditory, or perceptible by other senses) is the idea of phenomenology. The original ideas concerning phenomenology were raised by Edmund Husserl (1859-1938), who posited that we discern the essence of an object by reducing it into a phenomenon, or set of phenomena, that we observe about the object.

Following on Husserl's ideas concerning phenomenology, an approach originally outlined by Max Wertheimer in the 1930s makes use of Gestalt principles to demarcate the essentials that cause an object to be perceived as such. These principles cover aspects of the visual phenomenon such as proximity, similarity, common fate good continuation, and closure, which define how we perceive an object as a coherent entity (or group of related entities). The principle of exclusive allocation is also key to our conception of the visual object - a boundary between two objects is always perceived as belonging to one and not the other, thus placing the object in the foreground rather than the background. These principles are examined in this chapter to explain how an audiovisual object can be defined and outlined, establishing it as a phenomenon that exists within both the audio and visual perceptual streams observed by the audience member, existing in both at the same time and for the same duration.

These principles have been extended by Julesz and Hirsch to compare the idea of visual objects to sound objects, observing where the principles can be directly compared and where they cannot. In their work, they defined sound as operating in the temporal domain, while vision was primarily spatial. Bregman has also explored these principles, and compared exclusive allocation between the visual and sound domains.

The audiovisual object consists of the fusion of sound and vision. It is in itself the basic building block of audiovisual composition, in the same way that the sound object is the basis of electroacoustic composition. In his 2005 paper, *A Glow on Pythagoras' Curtain. A composer's perspective on Electroacoustic music with video*[1], Diego Garro observed that, "Depending on the compositional strategies adopted, digital artists can explore the relationships between the phenomenology of the *objet sonore* and that of the moving image, towards the definition of what we may call '*objet audiovisuelle*'".

Husserl's original thoughts about phenomenology contained the concept of identifying an object in terms of its perceived properties as individual phenomena, but he also considered the existence of phenomena within the temporal frame - "They are continuously in flux, continuously coming and going, and a kind of flux lies even in the "lasting". Over and over, a now phase is there in the appearance of

[1] Garro, D. (2005). *A glow on Pythagoras' curtain. A composer's perspective on electroacoustic music with video*. EMS International Conference Series. Available at http://www.ems-network.org/spip.php?article169 [accessed 23/08/2015]

lasting, but the now is already a was, and a new now has stepped in for it."[2] Perceiving a phenomenon is not a simple single act – Husserl placed the experience of phenomenological perception on a variable timeline. The act of perception in the moment is *primary sensation* – however, there must also be *retention* of the earlier perceptions, so that a sequence of images or sounds makes a definable pattern or collected object – otherwise we would just perceive a meaningless succession of sounds and images. As we can also relive the memory of experiencing a phenomenon, there must also be an element of *recollection* in our perception, which can affect how we perceive in the moment. There is also an argument for *protention* or *expectation*, where previous experience affects how we anticipate the development of a set of sounds or images.[3] This is also important in how we compose audiovisual work, building themes and guiding the audience to anticipate the resolution of the piece – or to challenge that anticipation.

The Sound Object

The original concept of the sound object is detailed in Pierre Schaeffer's "Traite des Objets Musicaux", published in 1966, and recently translated into English by John Dack and Christine North. Schaeffer postulates that there are various forms in which you can listen to music. One of these, and the one that is most influential in the idea of the sound object, is *reduced listening*. Reduced listening is a phenomenological approach to hearing – the listener isolates the sound itself from any other influences, and treats it as a single phenomenon in and of itself. This is a correlate of the idea of the sound object – by removing the experience of listening to a sound from any association of the sound with its source, a phenomenological reduction, the listener perceives the sound object as an object in itself.

As Schaeffer puts it; "if the listening intention is directed *towards the sound itself*, as

---

[2] Husserl, E. (2008). *Introduction to logic and theory of knowledge*. Dordrecht, London: Springer.
[3] Clarke, D. & Clarke, E. (2012). *Music and Consciousness: Philosophical, Psychological, and Cultural Perspectives.* Oxford Scholarship Online.

with the instrumentalist – or the acousmatic listener indifferent to conventional language and the anecdotal origin – signs and values are surpassed, forgotten, renewed, for a unique, unusual, but nevertheless irrefutable perception: *having ignored the source and the meaning, we perceive the sound object.'*[4]

This is much more appreciable with modern technology than in the early years of electroacoustic composition. We can now record and replay sound in an environment quite distinct from the original source. Hearing a reproduction of a sound from a loudspeaker has the immediate effect of drawing a veil between the listener and the original sound source. The discipline of reduced listening adds to this by deliberately choosing to ignore any suggestions that the sound may convey about its source or symbolic meaning.

In essence, Schaeffer was approaching the idea of the sound object as the basic starting element of electroacoustic composition. This can be compared to the single note as the base compositional unit in scored music – as per Chion, "*Sound object* and *musical note*: insofar as it is a unit of sound, a 'gestalt', which can be made up of several micro-events bound together by a form, the sound object in a classical music cannot precisely match each note on the score: a harp arpeggio on the score is a series of notes; but, to the listener, it is a single sound object."[5]

This is a key concept - the sound object in Schaeffer's terms is equivalent to a note, a building block that can be used in creating an electroacoustic musical phrase – and then the phrase is in itself another sound object, a gestalt made up of the individual sound objects. Therefore we can regard the sound object as a single or gestalt object, which can be used on its own or in conjunction with others to allow an electroacoustic composition to take form. By developing concepts, motifs and ideas in the composition, building anticipation and either fulfilling or not fulfilling it, we create a tension and a flow through the work.

---

[4] Schaeffer, P. (1966). *Traité des Objets Musicaux - Book II Version 1.* Translated by John Dack and Christine North. (2011). Not yet published. p50.

[5] Chion, M. (1983). *Guide des Objets Sonores.* Bibliothèque de recherche musicale. [N.p.]: Buchet Chastel. Part 1 translated by John Dack and Christine North. (2009).. p33.

One instance of how a sound object can be developed in a composition is as gesture – a sound object or gestalt will be highlighted against a background of silence or other sound objects. This development and examination of the foreground object against the background by the composer has resonances in the visual and audiovisual realms, and is one of the areas that will be examined further with reference to examples in the case studies presented in Chapter 5.

The Visual Object

In the same way as the sound object can be separated from the originating source, the visual object can also be captured and reproduced elsewhere. The visual object is not the object itself, it is an object in its own right - a photograph of Marilyn Monroe does not require the presence of Marilyn Monroe herself. The visual object can be a single occurrence or a gestalt of several objects, as per Schaeffer's thoughts on the sound object. In contrast to the sound object, there has been much more consideration of the elements that make a visual object an identifiable object. Gestalt theory is one such approach, which considers both how an individual object can be discerned, and how a group of objects can be perceived to belong to each other and form a recognisable group or gestalt in themselves – which can be compared to the idea of a single sound object, or a combination of sound objects creating a larger gestalt, which is in itself an identifiable object.

As with sound, the visual object can be perceived in terms of its boundaries. The boundary 'line' of a visual object can be observed to have 'exclusive allocation' – the boundary is perceived to be the edge of the foreground object against the background. In some cases, where the background and foreground are not necessarily fixed, the allocation of the boundary line may change as we look at the visual scene, but the boundary can only ever belong to one of the objects at a time, the 'exclusive allocation'.

A further correlation with sound is the concept that a visual object may consist of both a single identifiable object, and a group of objects – a single starling is identifiable as a single visual object, but a flock of starlings also seems to form a coherent visual object, mirroring the idea of 'notes' and 'phrases' as being identifiable sound objects.

Gestalt Theory Comparisons between Audio and Visual Perception

The principles of Gestalt psychology, as originally outlined by Wertheimer[6], were established as a way of defining what gives an object or group of objects this sense of belonging together, although these have been reinterpreted and refined over the years. Gestalt psychology does not address how the human brain physically perceives visual and sound objects, it explains common principles in how human beings tend to combine constituent elements into objects as part of everyday perception.

The classic approach to these principles has been to consider visual examples, as per Bruce et al.[7] for example, but there have been some explorations of auditory equivalents as well. In their essay 'Visual and Auditory Perception – An Essay of Comparison'[8], Julesz and Hirsch spend some time using Gestalt principles to compare and contrast the two senses. It is interesting to note that the psychological approach has generally been to link sight with space, and sound with time – Julesz and Hirsch note at the start of their essay, "… the inherently two and three-dimensional *spatial* patterns in vision are very different from the one-dimensional *temporal* patterns in audition"[9] (my emphasis). This reflects the acceptance in the field that sight is predominantly associated with establishing perception in space while sound, although it has a spatial aspect, is more concerned with perceiving

---

6 Wertheimer, M. (1938). *Laws of Organization in Perceptual Forms*. In Ellis, W.D. (ed.) *A Sourcebook of Gestalt Psychology*. London, UK: K Paul, Trench, Trubue & Co. pp71-88.

7 Bruce, V. Green, P.R. and Georgeson, M.A. (2003). *Visual Perception: Physiology, Psychology and Ecology*. 4th ed. Hove, UK. Psychology Press. pp119-127 and 413.

8 Julesz, B. and Hirsch, I.J. (1972). *Visual and Auditory Perception – An Essay of Comparison*. In David, E.E. and Denes, P.B. (eds.) *Human Communication: A Unified View*. New York, USA: McGraw Hill. pp283-340.

9 *Ibid*. p283.

events over time. However, as with the previous observation concerning Husserl's original work on phenomenology, all phenomenological perception has to take place within a timeframe, and so the temporal element should always be considered.

The core idea behind Gestalt principles is that of *prägnanz*, a German term that loosely translates as conciseness, which is used to summarise the human tendency to organise perception into regularity, simple objects with symmetry. Six of the most common principles are explored below to compare and contrast how both visual and sound objects can be defined in terms of these principles. Julesz and Hirsch employed five main principles, focussing first on the visual field and then examining how sound could also be defined using the same principle – this approach has been adopted below, but use has also been made of the more recent definitions as per Bruce et al. to establish the visual characteristics and compare them to an auditory equivalent.

Gestalt Principle 1: Proximity

In essence, the closer together a group of points or units is perceived to be the more likely the group is taken to be an object or group of objects in its own right. This is illustrated in Figure 1 – the audio equivalent would be a group of sound objects played together, or close together, separated from other similar or dissimilar groups by a time interval. Although the examples here are given in spatial terms for visual objects and in temporal terms for sound objects, it should be noted that audiovisual objects can demonstrate proximity in both dimensions, and across a combination of spatial and temporal dimensions.

Figure 1: Proximity

Gestalt Principle 2: Similarity

When a group of objects is perceived to be similar, that also tends to make us group them together. In Figure 2, although all the circles are the same size and equidistant to each other, the presence of a white centre dot in some of them leads to a perception that they are arranged in distinct groups of similar dots. In auditory terms, a distinction by pitch can be perceived as a distinguishing mark of similarity – if a series of low tones is succeeded by a series of high tones, the low tones will tend to be perceived as making up one group object while the high tones make up another, even though the duration and separation of the tones in each series may be identical. Again, in audiovisual composition, similarity can be displayed across both the spatial and temporal dimensions.

Figure 2: Similarity

Gestalt Principle 3: Common Fate

This principle states that objects that seem to move together in a shared direction are perceived as a group. This is easier to illustrate with a moving image than a still one – a good example is the movement of a flock of starlings in flight, or a school of herrings. Although the group is made up of individual animals, the shared movement gives the overall impression of a larger composite object.

Figure 3: Common Fate [10]

An auditory example would be of pitch movements – if a group of sounds are all moving in one pitch direction, despite differences in volume and timbre, they will be perceived to form a composite group in themselves. In audiovisual terms, a group of audiovisual objects can also be perceived to share a common fate, by exhibiting common movement in both sound and vision, across both spatial and temporal dimensions.

[10] Photograph: David Buimovitch/AFP/Getty Images. Picture source: http://www.theguardian.com/news/2013/jan/25/picture-desk-live-the-best-news-pictures-of-the-day [accessed 18/09/2014]

Gestalt Principle 4: Good Continuation

Following on from common fate, where two objects intersect or interact, the perceptual tendency is to allocate a continuation to the objects – for example, in Figure 3, the two lines tend to be perceived as smooth curves crossing in the middle, rather than two more jagged V shapes touching at the tips. This can be varied by changing the width or colour of a V shape to make it more distinct. Interestingly, this does not map exactly to an auditory equivalent – if two sine waves, one rising and the other falling, are mixed together, the sound is more easily perceived as starting from a high note, falling to a mid pitch, then rising again, against a low note rising and then falling, rather than two distinct movements of pitch, one fully upwards, the other fully downwards. In audiovisual composition, therefore, the principle of good continuation can be used, but it will be linked to the visual component of the audiovisual object rather than the audio component.

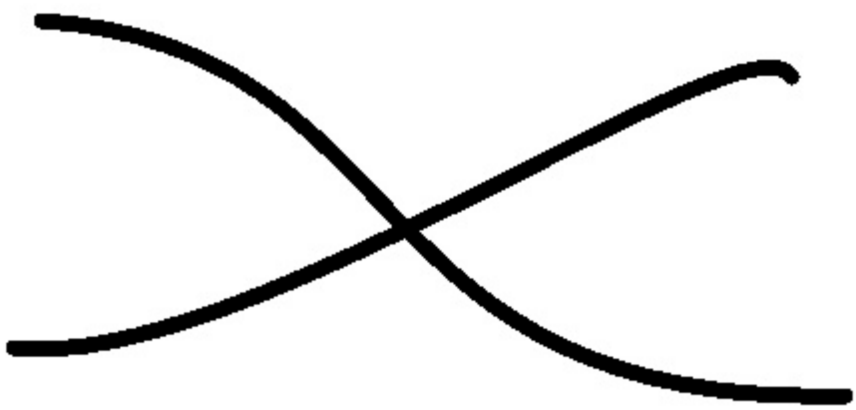


Figure 4: Good Continuation

Gestalt Principle 5: Closure

In visual terms, the eye tends to perceived closed groups of points rather than open ones – in Figure 4; the perception is usually of a square implied by the four smaller squares as corners, rather than of a white centre cross delineated by the sides of the smaller squares. This principle is much harder to establish in sound – due to the way we perceive sounds over time, we cannot enclose the sounds in a form that has a closed boundary. Even if the sound is symmetrical, the ending point happens after the start, and it is impossible to overcome this limitation. For audiovisual composition, closure can again be applied to the visual component, but this is only in a spatial dimension – there may be a feeling that an audiovisual object completes a movement on a temporal basis to close a perceived full object, but this makes use of common fate and can be compared to the musical concept of resolution rather than being a true evidence of closure.

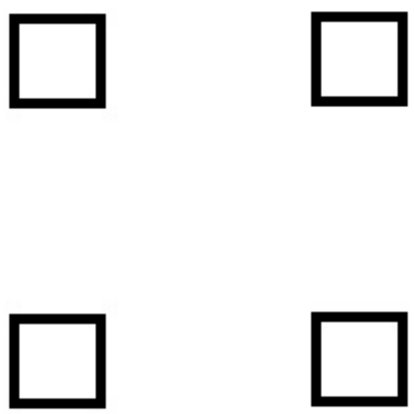


Figure 5: Closure

Gestalt Principle 6: Surroundedness (Belongingness)

If all else is equal, a smaller area will be seen as a figure against a larger area, which will be perceived as the background. This becomes reinforced if the smaller figure is surrounded by the background. In visual terms, this can be seen in any of the illustrations above – the black dots and lines are seen as smaller figures on a white ground. In auditory terms, a sound event will be perceived as an object in its own right if it is contrasted against a more diffuse and/or longer sound in the background. The event can be a short one, such as a percussive strike, or a longer one – as long as it is perceived as being surrounded by the other material, it will create its own sense of 'belonging as an object'.

A specific example of this can be heard in Smalley's 'Pentes' (1974)[11], where the emergence of a Northumbrian smallpipes melody occurs at approximately 6'07". The preceding material is a droning textural background, with some gestural acousmatic sounds. The smallpipes melody emerges from this textural background, an identifiable, specific object in its own right, then fades away and is absorbed into the background.

The point at which the melody becomes identifiable and then when it becomes indistinct can be regarded as the boundaries of the sound – although the background is still there, there is an edge at which the new sound object becomes apparent, which is not part of the surrounding background. The melody is a figure against the ground, surrounded by it but distinct in its own right. This edge is allocated to the object, a principle referred to as exclusive allocation and normally applied to visual objects, but extended to sound objects in Bregman's work on Auditory Scene Analysis.

Bregman observes that visual gestalt principles are echoed in auditory perception - "Parts of a visual field that are similar to one another, close together, whose contours follow a smooth trajectory are likely to be the visible parts of the same physical

[11] Smalley, D. (1974). *Pentes*. CD *Sources/scènes* empreintes DIGITales IMED 0054/2000. Track 5.

object. Similarly in audition, sounds that are close together in time and similar to pitch and other acoustic properties are likely to have arisen from the same acoustic source."[12] In combined audio and visual environments, the human tendency to create associations between what they see and hear also creates this perception that parts of the visual field and audio field, when experienced close together, which share common values of similarity, proximity, etc. will also lead to the impression that there is an audiovisual object to be perceived, composed from the two perceptual streams.

Additional Gestalt Principles

The general Gestalt Principles described above have been re-interpreted over time, and other principles have been added, or suggested as alternatives to the above. For the audiovisual composer, some of these additional principles can also be adopted as useful approaches to creating audiovisual objects within a coherent composed piece. The four main additional principles that are most useful in this capacity are the ideas of emergence, re-ification, multistability and invariance.[13]

Additional Gestalt Principle 1: Emergence

The idea of emergence is that within a field of items within an image, the eye will look for individual identifiable groups of the items that can be resolved into a recognisable object. This is broadly similar to the principle of Surroundedness or Belongingness as described above, but differs in that a discerning selection of constituent elements will coalesce inclusively into a coherent object, but there may be other similar elements which do not contribute to the same object. Similarly, in auditory terms, individual sounds and pitches may combine within audition to form a

---

[12] Bregman, A.S. (1990). *Auditory scene analysis: the perceptual organization of sound.* Cambridge, Mass. ; London: MIT Press. p202.

[13] Lehar, S. (2003). *Gestalt isomorphism and the primacy of subjective conscious experience: A Gestalt Bubble model.* Behavioral and Brain Sciences, 26(04), pp375-408.

chord or coherent sound object, while other individual sound objects may still be discernible, but do not contribute to that larger sound object.

An example of this principle in visual terms can be seen below, an image created by Richard Gregory, which at first glance is merely a grouping of black dots of various sizes against a white background. After a short period, the image of a Dalmatian dog is usually perceived as emerging from this general field. This idea of an object or theme emerging from a diffuse background can be used by the composer – as in the example already mentioned above, in Smalley's *Pentes*, the smallpipes melody emerges from a general, diffuse background – and coexists with this background, but maintains its own identity and integrity, much as the dog below emerges from the visual field and retains its own identity while the other black dots that do not contribute to the dog as an object also continue to exist as separate perceptible visual items within the image frame.

Figure 6: Emergence [14]

[14] Richard Gregory's *Dalmatian Dog* – picture source: http://artandperception.com/2007/05/why-do-images-change-with-viewing.html [accessed 16/09/2014]

Additional Gestalt Principle 2: Re-ification

The concept of re-ification concerns the propensity of the viewer to complete a perceived object when it is not clearly delineated as a complete object in vision – for example, if a car is parked in front of a low fence, the viewer will only see the portions of the fence that are not obscured by the car, but will infer from the parts they can see that there is a continuation of that fence behind the car. Similarly, if an object is inferred to exist through the appearance of other objects, the eye will extrapolate from the inference to perceive an object, even though it may not actually exist in vision – an example of this is shown below. The three circles shown each have a slice 'missing', and the arrangement of these missing slices is such that a white triangle can be inferred to exist, even though there is no actual triangle existing within the image.

Similarly, the ear can extrapolate from the sounds it hears to perceive additional sound objects that do not exist in reality. This has been used for smaller loudspeakers, to create a false impression of a full bass note that the speaker is not physically big enough to create – by creating additional harmonics to the implied bass note, the ear is fooled into recreating the bass note as existing within the music, even though it has not physically been created by the speaker.

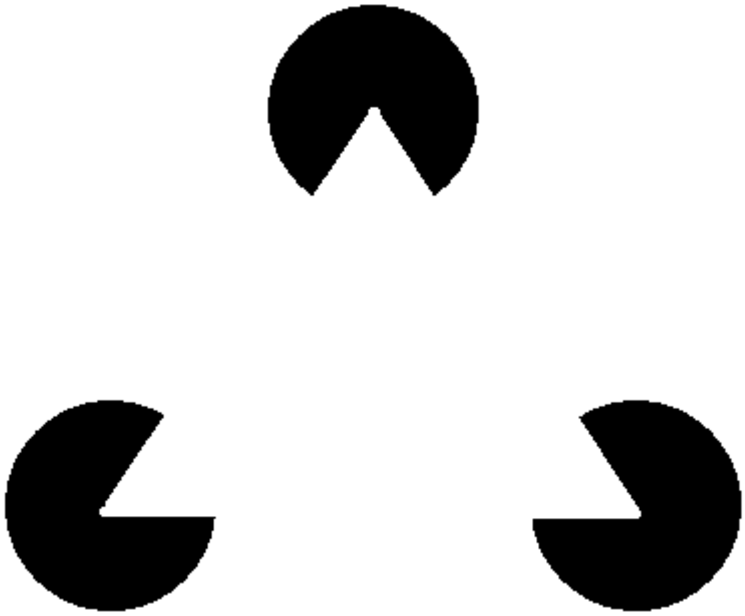


Figure 7: Re-ification

Additional Gestalt Principle 3: Multistability

This principle addresses the phenomenon where an image can be interpreted in more than one way. This has been classically used in optical illusions and even popular design and fashion throughout history. The figure below shows a wireframe of a simple cube, which can be seen in various possible physical positions – the eye will move between these individual possible interpretations. Each is equally plausible and valid, and resolution may only be achieved when the cube is compared against other objects in the background to establish the 'correct' basis for perception.

It should be noted that the perception of the various positions the cube could occupy all exist in separate temporal frames – the viewer cannot hold more than one possible position as a perception at the same time. This ties in with the principle of *exclusive allocation*, examined in more depth below.

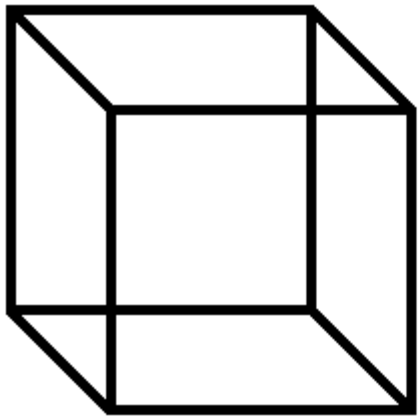


Figure 8: Multistability

Additional Gestalt Principle 4: Invariance

The concept of invariance is that an object is perceived as a whole – so that the viewer will always perceive a specific object as that object no matter how it is viewed, if it is rotated, what position it is in, etc. This principle often combines with the earlier principle of re-ification, so that even if part of the object is concealed, the object is still perceptible in its entirety, even if it is rotated or changed from its normal position.

In auditory terms, this can be likened to a short melody or perceived sound object consisting of a series of pitches – the timbre of the instrument can be changed, the absolute pitch can be raised or lowered, but the listener will still perceive the sound object to be essentially the same whole object.

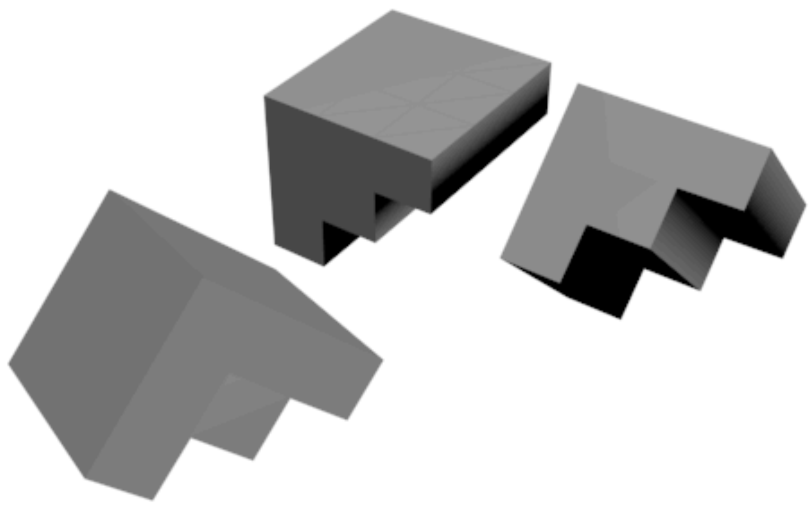


Figure 9: Invariance

Exclusive Allocation

In exclusive allocation, the boundaries of the object are perceived as solely belonging to that object alone, rather than being shared with another. A well-known visual example of this is Rubin’s example from the 1910s of a vase, which could also be perceived as a pair of faces looking at each other (Figure 10). One or other interpretation can be perceived at any one moment – while it is possible to switch between the two, both cannot be perceived at the same time. The boundary between the light and dark areas exhibits exclusive allocation – at any given moment, it either defines the edge of the central vase, or the edges of the faces, but does not exist in the two capacities at the same time either.

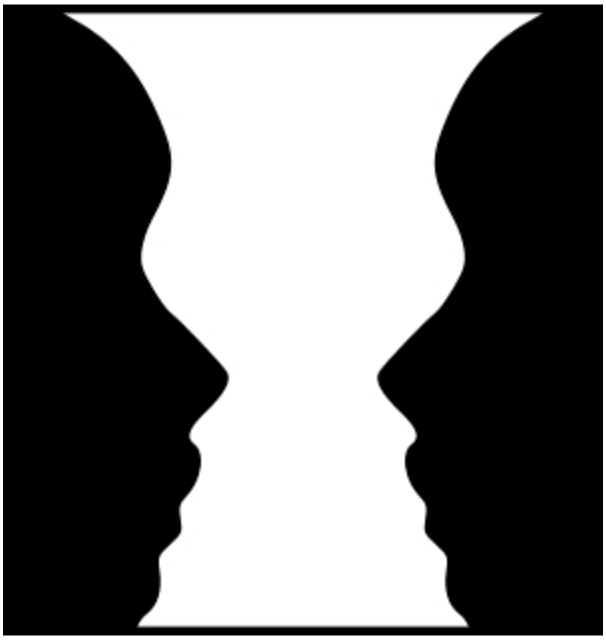


Figure 10: Rubin's Faces-Vase Illusion

Bregman provides examples of how exclusive allocation can also be perceived in auditory information. As this is in a temporal dimension, he refers to the distinct figures as auditory streams, and these form the basis of his analysis of auditory scenes. He observes that, "dividing evidence between distinct perceptual entities (visual objects or auditory streams) is useful because there really are distinct physical objects in the world that we humans inhabit."[15]

In creating the audiovisual object, the composer can make use of exclusive allocation in both the visual and auditory fields to delineate the object against the background of underlying image and sound. This is particularly effective when the object emerges from the background, or a multistable object is being explored, and the audience is encouraged through the manipulation of sound and image to perceive the various forms of the object emerging and changing against the underlying background. By using multiple layers, interacting audiovisual objects, and developing these interactions across the duration of the piece, the composer can

[15] Bregman, A.S. (1990). *Auditory scene analysis : the perceptual organization of sound.* Cambridge, Mass. ; London: MIT Press. p13.

create an immersive audiovisual experience, convincing the audience that the objects created within the composition have an existence and validity within the boundaries of the frame, both visual and sonic.

Gibson and 'Affordances'

There are alternative approaches to identifying how perception works. One, outlined by Gibson[16], postulates that the sensory perceptions we have are based on affordances, which are allocations of value to the percept that we have learnt at an early age – for example, we look at the sea and think of it as a nice place to go swimming; a fish looks at it and sees it as home.

There are some issues with Gibson's approach. Lehar has noted that there is a "distinction between the Gestaltist and the Gibsonian views of perception. That is, the epistemological question of whether the world we see around us is the real world itself or merely an internal perceptual copy of that world generated by neural processes in our brain. In other words, this is the question of *direct realism* (also known as *naïve realism*) as opposed to *indirect realism* (or *representationalism*)."[17]

Gibson's approach involves direct realism, which effectively assumes that the whole world is perceived directly by the human mind without any filtering through sensory input – the process of perception occurs in the real world rather than within your brain. However, there manifestly is processing taking place as we look and listen to the world, which implies that the direct realism approach is inadequate.
The use of Gestalt theories as described above falls into the indirect realism camp – as opposed to Gibson, the assumption here is that we recreate the world in which we live within our own mind, creating our own perceived version of reality. This also has problems, as although it acknowledges the sensory processing our minds carry

16 Gibson, J.J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.
17 Lehar, S. (2003). *Gestalt isomorphism and the primacy of subjective conscious experience: A Gestalt Bubble model.* Behavioral and Brain Sciences, 26(04), pp375-408.

out, there does not seem to be enough neural activity to account for the sheer volume of data processing we would need to undergo in order to perceive and recreate an entire world within our head. Lehar has suggested a containment theory that creates a 'bubble' of perception within which Gestalt theories can be applied. McLoughlin has challenged this approach – he states that "the high-resolution representation is generated only when we pay attention to the input and focus our eyes on the object or texture under inspection. We need not represent even our immediate environment in high resolution unless we need to interact directly with it."[18] In other words, the majority of the world we perceive is not 'captured' in great detail unless we decide, consciously or unconsciously, to pay particular attention to it. A movement happens at the corner of our vision – we perceive it in low definition, then move our eyes to concentrate the image within the fovea, the most perceptive part of the eye, to see it in high definition. Similarly, we hear low-level noise at a cocktail party, but if our name is mentioned, our ears hone in on that particular conversation, isolate it from the background noise, and our mind processes it with specific attention to capture it in high definition.

For the audiovisual composer creating the audiovisual object, we are already working in an environment where the audience's eyes and ears are primed to experience the composed work in high definition. In this environment, the Gestalt principles appear to give the best method of understanding how structure and form in the work is being perceived, and thus how to make use of these principles to emphasize and mould the shape and movement of the audiovisual objects that combine to create the composition. But while Gestalt principles can be used to examine this flow of form and shape, another key visual ingredient is colour, which does not readily conform to the same phenomenological reductions as line and shape do.

[18] McLoughlin, N.P. (2003). *Bursting the bubble: Do we need true Gestalt isomorphism?* Behavioral and Brain Sciences, 26(04), p421.

Colours and Vision

While Gestalt principles establish a method of examining the phenomenon of visual perception with a focus on the boundaries of objects, the lines that define objects in our vision, this leaves aside one of the other major elements of visual perception – the colour or mixture of colours that an object exhibits. The line can be explored as a boundary object, an element of vision that establishes the edges of a shape. Colour is more generally associated with the concept of defining space within the boundaries – the 20th century artist Robert Delaunay observed, "drawing, ... being in effect a stranger to the problem of volume, not creating real volume as is created by color"[19].

A Gestalt approach to definitions of colour in visual perception contains an inherent difficulty – Riley notes that "Amongst the prevailing schools of psychological thought that place color in a corner is Gestalt theory. Its name alone suggest why it might find color secondary. It can be translated as form, figure, or remotely, pattern. It gives priority to figure rather than ground, an outline preference that is antithetical to the usual color association."[20] Therefore, in the incorporation of colour into the audiovisual object as we carry out compositional work, the materiality of colour as we perceive it is a key consideration in the creation of the audiovisual object.

Colour and Human Perception

The physiology of sight can be examined in terms of the receptor cells in our eyes and their responses to the wavelength and intensity of the light that hits them. This biological reception of colour is different from our perception of colour, the materiality of how we perceive it, create it and respond to it. The colours that are normally observed in human vision (where there is no occurrence of colour

---

[19] Cohen, A.A. (ed.) (1978). *The New Art of Colour: The Writings of Robert and Sonia Delaunay.* The Viking Press. New York. p8.

[20] Riley, C.A. (1995). *Color codes : modern theories of color in philosophy, painting and architecture, literature, music, and psychology.* Hanover : University Press of New England. pp299-300.

blindness) consist of two types – *chromatic* colours, which are the ones we would most commonly immediately think of as colours – red, blue, yellow, green, etc. - and *achromatic* colours – white, grey and black[21]. The chromatic colours are effectively those in the rainbow of observable light – with a wavelength of just under 400nm to just over 700nm, ranging from violet through to red. The achromatic colours are effectively a linear scale from white to black. The combination of the chromatic colour, or hue, with achromatic colour, or light/shade, allows more variation in our perception of colour, such as pink or dark red.

The colours that are perceptible by humans lie within a narrow band of the electromagnetic spectrum. Light waves with a wavelength of between about 380-730nm can be detected by the human retina, covering the colours we are accustomed to seeing in a rainbow – the classical Red, Orange, Yellow, Green, Blue, Indigo and Violet described by Isaac Newton.

Figure 11: Colours of the Visible Spectrum

Light, whether received directly from a source or reflected from a surface, enters the eye via the pupil, with the image brought into focus by means of the iris, a diaphragm of muscle adjusting the depth of focus of the eye. The light then reaches the back of the eye, where it falls on the retina. The retina makes up more than half the surface of the rear of the eye, and consists of a very complex array of light-sensitive receptor cells – the ‘rods’ and ‘cones’, named for their relative shapes. There is a small depression at the part of the retina where the focus of our attention would normally lie, the fovea, where the proportion of cone cells to rod cells is much

[21] Kernell, D. (2016). *Colours and Colour Vision: An Introductory Survey*. Cambridge University Press. p4.

higher[22].

The rod cells have a lower threshold for activation than cone cells, meaning that in lower light levels our sight mainly consists of sensations communicated by these cells. Rod cells do not differentiate between the wavelengths of the visible spectrum, and so only return values of light or darkness, which means we are effectively colour blind in circumstances of low light. As a consequence of the higher concentration of cone cells in the fovea, we also have a lower level of light detection within the centre of our vision, and actually find it easier to detect light if looking slightly to the side. Rod cells are also less able to allow us to define sharpness, so our night vision is much less defined and 'sharp' than our vision in full daylight[23].

Our eyes contain three different types of cone cells – L cones are most sensitive for wavelengths in the longer ranges, with a peak sensitivity of between 560-570nm. M cones are most sensitive for the middle wavelength ranges, peaking at about 530-545nm. Finally, S cones are most sensitive for the shorter wavelengths, with a peak at about 420-440nm. However, it is important to note that each type of cone does still respond to a range of wavelengths, and their responses allow the brain to process the signals it receives to place the light it is receiving at the correct point in the visible spectrum. Despite this, due to the different peak sensitivities, the cone cells are also sometimes referred to as red (L), green (M) and blue (S) cones[24].

---

22 *Ibid*. pp103-110.
23 *Ibid*. pp112-114.
24 *Ibid*. pp107-109.

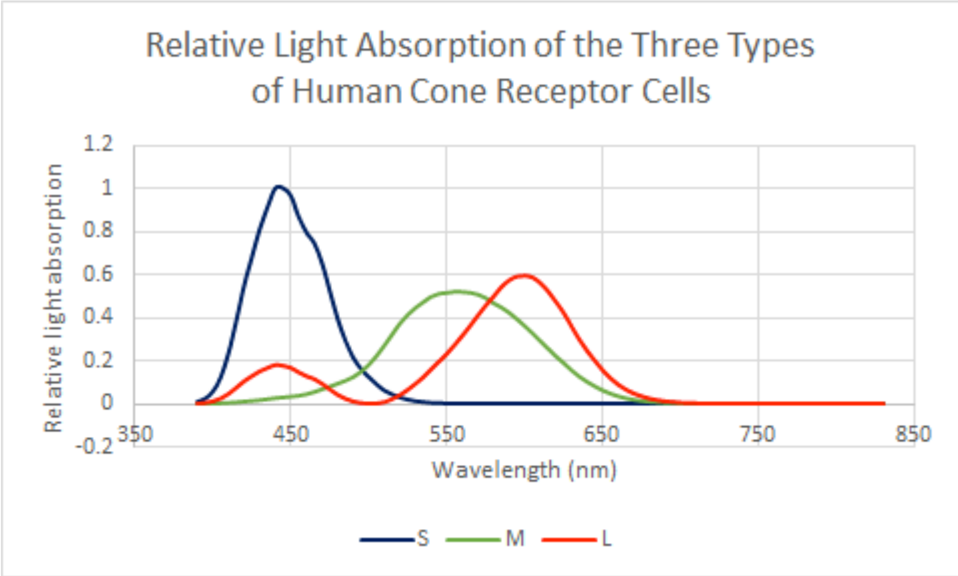


Figure 12: Relative Light Absorption of Human Cone Cells[25]

Colour Definition and Colour Harmony

The definition of colours varies between cultures – the very simplest divisions made in language generally distinguish between light and dark; in languages with a limited range of colour names, the next definition is usually red, while languages which define more colours then usually also include green, yellow and blue[26]. In our current era of computer and television technology, the concept of red, green and blue as primary colours, from which other colours can be derived, has become quite familiar. However, Western artists have generally regarded red, yellow and blue as primary colours, which reflects the properties observed from mixing these colours as physical paint media.

In visual arts, traditional colour theory refers to how we can define colours, how they interact with each other, and how colours mix together to create other colours. Taking

[25] Source data from Colour and Vision Research Laboratory – *Stockman and Sharpe Cone Fundamentals* (2000), accessible at http://www.cvrl.org/ [accessed 12/02/2017].

[26] Arnheim, R. (1974). *Art and Visual Perception: A Psychology of the Creative Eye*. University of California Press. pp331-332

the primary colours to be red, yellow and blue, we can create secondary colours by mixing two primary paint colours together – red and yellow make orange, red and blue make purple, and yellow and blue make green. The secondary colours are often referred to as complementary colours, in that they complement the primary colour that is not part of the mix – therefore orange is the complement of blue, purple of yellow, and green of red. This mixing is a *subtractive* process – in that the colour we perceive when looking at a specific paint is due to the reflection of specific colours to the eye, while other colour frequencies are absorbed by the paint. By mixing paints, the amount of absorbed frequencies increases, affecting the reflected colour frequencies reflected to the eye.

When dealing with colours generated by a direct light source, as occurs when we work with colour images created on a computer monitor or via a projector, the mixing process is *additive* in nature – the colour frequencies are added to when two different colour sources are mixed. The art philosopher Rudolf Arnheim notes, "In *additive* combination, the eye receives the sum of the light energies that gather in one place, e.g. on a projection screen. […] The colors received by the sense of sight are the result of an additive process because the three kinds of color receptors, placed side by side in the central area of the retinal surface, pool the stimuli they receive. *Subtraction* produces color sensations by what is left over after absorption"[27]. As noted above, when we work with an additive process, the primaries are taken to be red, green and blue, with their complementaries as cyan (green plus blue), magenta (blue plus red) and yellow (red plus green).

Starting from these basic points, the ways in which colours appear to work well together or not have been explored in theories of colour harmony. A colour wheel or circle is often created based on the primary and secondary colours, with gradations between the colours as they merge into each other. Starting with the primary colours as the axes of a three-dimensional space, colours can be defined based on the amount of each in the mixture – a full mix of red and blue would result in magenta, while a

[27] *Ibid*. p341.

full mix of red, blue and green would produce a pure white (conversely, where the axes all meet at zero, the result would be black). This can be illustrated as below:

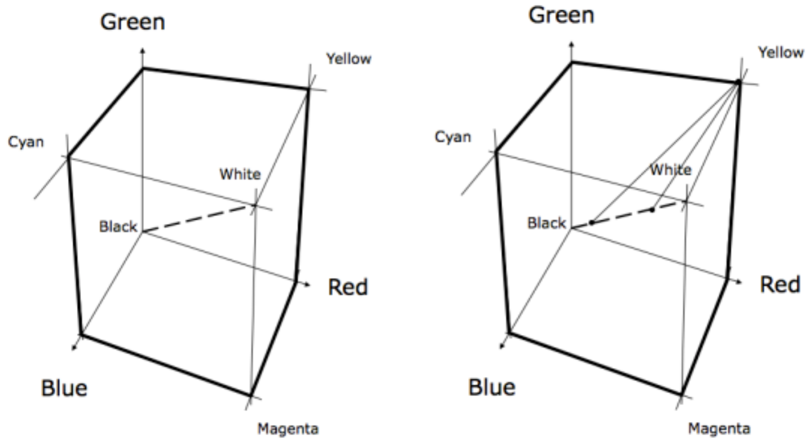


Figure 13: Three dimensional representation of additive colour combination [28]

The 'line' linking black and white can be regarded as a continuum covering the achromatic scale of greys between pure black and absolute white, while individual point positions within the three-dimensional cube can be measured as an admixture of specific amounts of red, blue and/or green.

This provides a method of defining colours – the proportion of each primary colour within the mixture of two primaries can be measured, giving a quantifiable system for defining a fully saturated colour (or *hue*). A further quantifiable measure is how light or dark the colour is, where we can imagine that if white or black is added to our colour, it will become lighter or darker, depending on the amount of white or black – the amount of the added achromatic colour would give a measure of the *saturation* of the hue, while the achromatic colour itself (on a scale from black through grey to white) is defined as the *value*. Our three-dimensional cube from above can be converted via a hexagonal representation through to a circular colour

[28] Coyne, R. (2012). *Nomadology and Colour.* Blog post available at https://richardcoyne.com/2012/06/02/nomadology-and-colour/ [accessed 29/04/2017]

wheel, which can then be used to indicate the values for hue (H), saturation (S) and value (V), respectively based on the 'angle' of the colour, the proportion of achromatic colour mixture, and the point on the black-white greyscale continuum the achromatic colour represents.

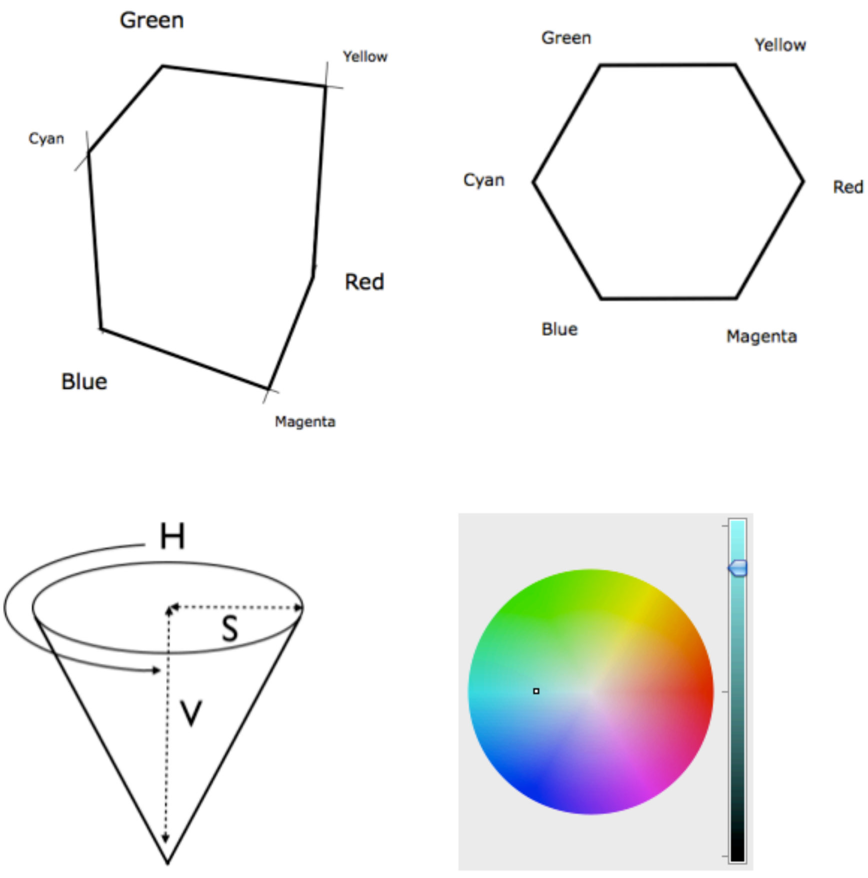


Figure 14: Hue, Saturation and Value as defining measurements for colour [29]

Following this schema, if complementary colours are regarded as those which have the greatest contrast to each other, exact opposites on this colour wheel can be taken to be complementary (e.g. cyan is the complementary of red, yellow of blue, etc.). Complementary colours create a visual tension if placed closely together, while analogous colours are those that are neighbours on the colour wheel, sharing a root colour, and which are considered to work in a stable harmony – such as red and

[29] *Ibid*. [accessed 29/04/2017]

orange. Placing complementary colours, or analogous colours, close to each other, appears to create tensions, beats, or visual relationships that can affect our perception of the visual composition of an image. In the examples below, the complementary colours of blue and yellow would convey a greater contrast, an almost discernible 'pulse' or rhythm, to a viewer than the contrast offered by a combination of yellow and orange – although this does vary from individual to individual.

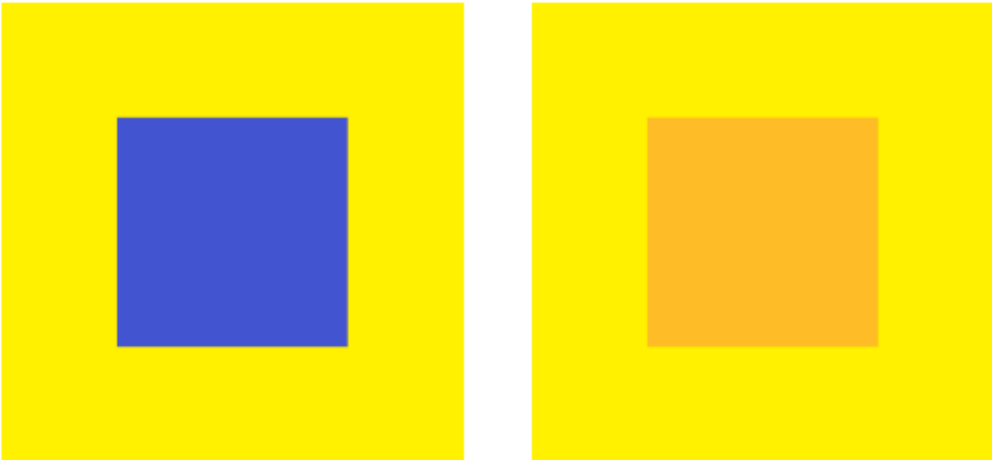


Figure 15: Comparison of complementary and analogous colours

This use of a perceptual tension or beat between strong colours has often been observed by painters and colour theorists – the artist Robert Delaunay noted, "the relationships between the different colors, simultaneous contrast, create a depth, a movement...determined solely by the creator's vision"[30], while Josef Albers also noted, "our concern is the interaction of color; that is, seeing what happens between colors... Colors present themselves in continuous flux, constantly related to changing neighbours and changing conditions"[31]. The development of pure non-representational art in the early 20th century included a strong element of colourism - the works of Mondrian, for example, construct strong images using flat planes of pure colour, achieving a visual balance between primary colours and the form

[30] Cohen, A.A. (ed). (1978). *The New Art of Colour: The Writings of Robert and Sonia Delaunay.* The Viking Press. New York. pp103-104.

[31] Albers, J (1963). *Interaction of Colour.* Yale University Press. p7.

definitions made using black lines and white space.

The early works by visual music composers also made use of strong primary colours - Walther Ruttmann's *Lichtspiel Opus 1* (1921) makes use of simple evolving shapes, generally coloured with a primary hue against a black background. Likewise, Fischinger's works such as *Kreise* (1934) and *Allegretto* (1936) also make use of simple forms and primary colours, using motion and aspects of Gestalt principles such as common fate to create a fusion with the music score. As the artform has matured, however, the use of more subtle, graduated colour palettes has also developed, giving a further ability to achieve a close union between sound and vision by creating associations between colour and sound characteristics and transformations.

Some of the developments in colour transformations and morphing arose with the use of layered transparencies in film - the effect of building up subtle colour graduations using overlaid transparencies had been explored in the early 20th century by colourists such as Delaunay: "Transparency – actually the illusion that a tonal body is built from layers of color chosen according to optical laws of mixture – engenders a sense of depth and was the prime consideration of the Delaunay aesthetic. As colored planes intersect and blend, one seemingly behind the other, a new surface is suggested beyond that of the canvas."[32] The animator Jordan Belson made use of such techniques in his experimental works - in his film *Samadhi* (1967), he superimposed tinted footage of smoke patterns over each other along with both soft and hard masking to achieve an immersive abstract effect.

---

[32] Riley, C.A. (1995). *Color codes : modern theories of color in philosophy, painting and architecture, literature, music, and psychology.* Hanover : University Press of New England. p121.

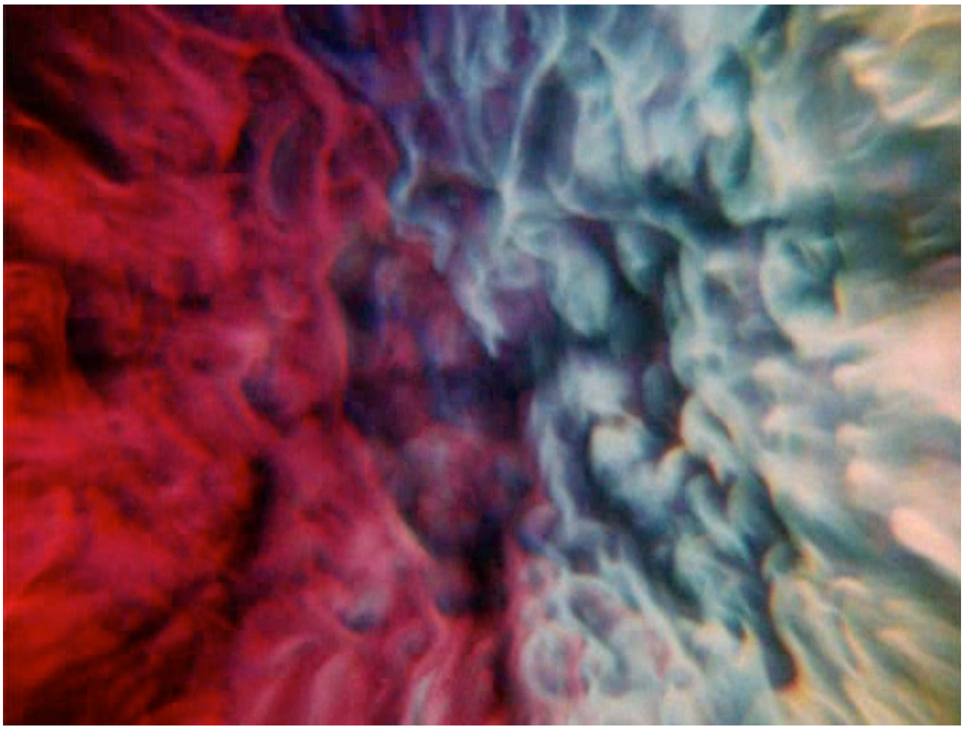


Figure 16: Still from Belson's *Samadhi* (1967) [33]

Using current software such as Adobe Photoshop or After Effects, we can now achieve similar effects, and develop them further, effectively by simulating these layers of film transparencies, with additional methods to define the way in which each layer interacts with others, opening up even further possible opportunities to create interactions between colours. In order to impose some compositional structure on these open opportunities, the composer can adopt principles that have been developed within the design industries, defining a specific limited colour palette to ensure that the colours used in creating audiovisual objects convey the correct impressions of harmony or disunity, according to the composer's wishes.

From the time of Newton onwards, there have been attempts to establish a consistent correlation between colour and pitch, which could then be used to create visual relationships between colours that would correspond to the musical pitches heard in

[33] Image from website http://www.triangulation.jp/2013/04/jordan-belson-5-essential-films.html [accessed 30/04/2017]

compositions. Unlike sound, where each doubling of note frequency gives the same note an octave higher, colour does not readily lend itself to such mathematical mapping. In audiovisual compositional practice, correlations between colours and pitches can normally only be made within the confines of a single composition, with repetition of sounds and notes against the same colours establishing that link, which can then be constantly referred back to throughout the work.

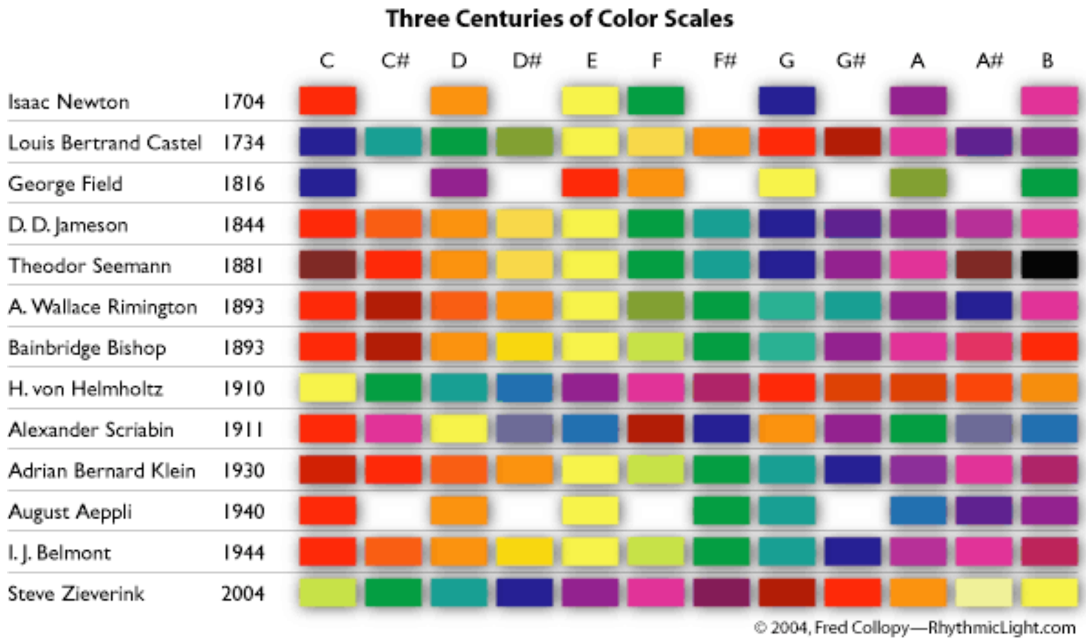


Figure 17: Proposals for colour / pitch correlations in the 18th-20th centuries[34]

For some individuals, however, there can be a consistent correlation between sound pitch and perceived visual colour. Individuals experiencing *chromasthesia* will make a repeatable, consistent link between a specific pitch and a specific colour. This is a form of *synaesthesia*, where sensory perceptions can trigger correlated responses in other sensory streams.

[34] Image from website http://rhythmiclight.com/archives/ideas/colorscales.html [accessed 30/04/2017]

Synaesthesia

Synaesthesia is a measurable psychological condition where there is a crossover between two sensory pathways within the synaesthete's brain. "A synaesthete might describe the colour, shape and flavour of someone's voice, or music, whose sound *looks like* 'shards of glass', a scintillation of jagged, coloured triangles moving in the visual field. Or, seeing the colour red, a synaesthete might detect the 'scent' of red as well."[35]

The most common form of synaesthesia appears to be the correlation of colour with pitch as described above, or *chromesthesia*. For chromesthetes, a specific pitch will invoke a colour or coloured object that is stable and repeatable – for example, the pitch C_3 will always convey a dark red colour, while G_3 may be bright red. However, these colours are not consistent between synaesthetes, and the intensity of the connection can also vary to a large extent. Some synaesthetes will also experience different shapes with changes in pitch and timbre.

The incidence of true synaesthesia in the population is relatively low, estimated at around 1 in 25,000[36]. However, the *idea* of synaesthesia has been discussed for over 300 years, and has provided inspiration for many artists and composers. The Russian composer Alexander Scriabin, who maintained he experienced synaesthesia, wrote compositions for *son et lumière*, where a full orchestrated work also had instructions for specific colours of light to be reproduced in time with the music, to create a multi-sensory experience. This resonates with Wagner's ideas of the *Gesamtkunstwerk*, or total art-form – his opera house at Bayreuth was specifically designed to create an immersive experience for the audience. The Russian artist Kandinsky was also intrigued with correlations between sound and image – he created an image and sound 'opera', *Der Gelbe Klang*, in 1909 (published in *Der Blaue Reiter* almanac in 1912), but sadly never saw it performed in his lifetime. He

---

[35] Cytowic, R.E. (1997). *Synaesthesia: Phenomenology and Neuropsychology: a Review of Common Knowledge*. Chapter 2 in Baron-Cohen, S. & Harrison, J.E. (eds.) (1997). *Synaesthesia: Classic and contemporary readings*. Blackwell. p17.

[36] *Ibid.* p17.

published his thoughts on art and synaesthesia in his work *Concerning the Spiritual in Art* in 1911.

I noted above that Scriabin maintained he was a synaesthete. While this is probably true, many artists and composers in the late 19th and early 20th centuries claimed to be synaesthetes, and many composers published their associations between colours and pitches. Evidence that Scriabin, Rimsky-Korsakov and Liszt actually had developmental synaesthesia is equally hard to find, and the associations made between pitch and colour varied between individuals. For instance, Myers recounts that for Scriabin F# seemed violet, whereas Rimsky-Korsakov believed it to be green."[37] This observation does not necessarily prove that the composers were not synaesthetes, but it does outline the difficulty in forming a common ground between synaesthetic experiences when comparing different individuals.

However, within the individual's own experience of synaesthesia, there is a high degree of consistency. Research carried out investigating the stability of sound-colour synaesthesia concluded, "not only are synaesthetes more consistent over time in their choice of colours, but they also tend to choose very specific colours to represent their experiences."[38] The experience is also generally unelaborate - "while you or I might *imagine* a pastoral landscape while listening to Beethoven, what synaesthetes experience is unelaborated: they see blobs, lines, spirals and lattice shapes"[39]

In my own practice, I experience a mild form of synaesthesia, and relate to these simple forms and colour matches – this is reflected in the work that I produce. This is in line with the approach taken over the 20th and 21st century by practitioners in visual music. The Whitney brothers, working in California in the 1960s and 70s,

[37] Harrison, J.E. & Baron-Cohen, S. (1997). *Synaesthesia: An Introduction.* Chapter 1 from Baron-Cohen, S & Harrison, J.E. (eds.) (1997), *Synaesthesia: Classic and contemporary readings*. Blackwell. p10.

[38] Ward, J. Huckstep, B. & Tsakanikos, E. (2006). *Sound-Colour Synaesthesia: To What Extent Does It Use Cross-Modal Mechanisms Common To Us All?* Cortex, (2006), 42, pp264-280.

[39] Cytowic, R.E. (1997). *Synaesthesia: Phenomenology and Neuropsychology: a Review of Common Knowledge.* Chapter 2 from Baron-Cohen, S & Harrison, J.E. (eds.) (1997), *Synaesthesia: Classic and contemporary readings*. Blackwell. p23.

observed that "sound is image, and image sound, with no fundamental difference."[40]

There is some evidence that the method by which synaesthetes have this combined sensory experience is also accessible by non-synaesthetes. Marks noted that "it does not matter whether one considers perception by synaesthetes or by non-synaesthetes; in both cases the correlations between dimensions of visual and auditory experience are nearly identical. These dimensional, cross-modal correlations express themselves most vividly in people not normally synaesthetic by means of hashish and mescaline intoxication."[41]

There is an indication that "both groups [synaesthetes and non-synaesthetes] use the same cognitive mechanism to map between the auditory and visual domain. This mechanism appears to involve mappings between pitch and lightness."[42] The audiovisual composer can make use of synaesthetic-like connections in their own work in order to evince similar reactions in non-synaesthetes as would be expected in synaesthetes, again leading to the correlation between sound and image we would expect to observe in an audiovisual object. The use of the same cognitive mechanisms by synaesthetes and non-synaesthetes highlights the idea of cross-modal perception, that there "are cross-modal audiovisual areas of the brain that appear to respond more to the combined presence of vision and sound than to either, or the sum of, vision or sound alone."[43]

Cross-modal Perception

While we have been examining correlations between sound and vision by comparing Gestalt properties, there is also evidence in modern neuropsychology that the human

[40] Brougher, K. Mattis, O. Strick, J. Wiseman, A. & Zilczer, J. (2005). *Visual Music: Synaesthesia in Art and Music Since 1900*. Thames and Hudson, London. p125.

[41] Marks, L.E. (1997). *On Colored-hearing Synesthesia: Cross-modal Translations of Sensory Dimensions.* Chapter 5 from Baron-Cohen, S & Harrison, J.E. (eds.) (1997), *Synaesthesia: Classic and contemporary readings.* Blackwell. p78.

[42] Ward, J. Huckstep, B. & Tsakanikos, E. (2006). *Sound-Colour Synaesthesia: To What Extent Does It Use Cross-Modal Mechanisms Common To Us All?* Cortex (2006), 42, pp264-280.

[43] *Ibid.* pp264-280.

brain also relies on combining senses while processing the sensory input it receives. "Behavioral neuroscientists find that integration of information from multiple sensory channels is crucial for attention and perception in humans, monkeys, birds, and insects."[44]

This interaction may be part of the explanation for synaesthesia – the emphasis on cross-modal perception may be higher in a synaesthete's sensory perception processes than a non-synaesthete. It also points to a natural synthesis between sound and vision in human perception, again lending strength to the proposal that audiovisual fusion does take place in artistic works that have a close or designed correlation between sound and vision.

In cross-modal perception, the single perceptual streams of vision and audio can be examined in isolation, and then the possible combinations of the two can also be further explored. By interacting positively, negatively, or in other combinations, the fusion of the two perceptual streams will result in different outcomes.

In simple equivalence or enhancement of the two media, the auditory and visual events are synchronised, the 'Mickey-Mousing' effect in simple animation. Here, the correlation between the two components will either be perceived as simply proceeding as normal, or the elements may reinforce each other, adding an intensity to each other. This is illustrated in Figure 18 (after Spence & Soto-Faraco[45], themselves after Partan & Marler[46]), where the multisensory percepts afforded by combining audio and visual signals are categorised.

---

[44] Partan, S., & Marler, P. (1999). *Communication Goes Multimodal*. Science, 283(5406), p1272.

[45] Spence, C. & Soto-Faraco, S. (2010). *Auditory perception: Interactions with vision*. In D. R. Moore, P. A. Fuchs, C. Plack, A. Rees & A. R. Palmer (eds.), *Oxford Handbook of Auditory Science: Hearing.* (2010), Oxford, UK: Oxford University Press. p272.

[46] Partan, S. & Marler, P. (1999). *Communication Goes Multimodal*. Science, 283(5406), p1272.

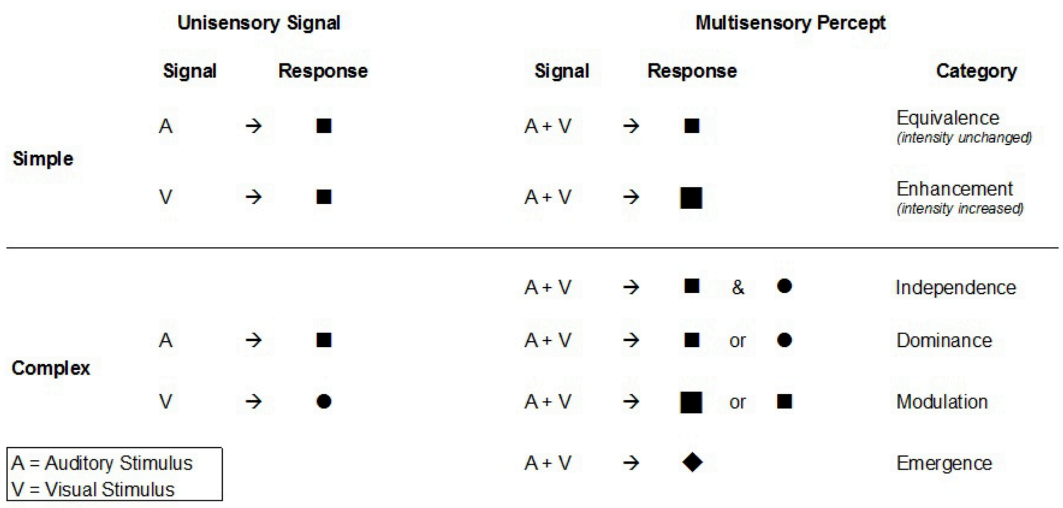


Figure 18: Interaction of auditory and visual stimuli in multisensory perception

In more complex interactions between sound and vision, where the two may be deliberately designed to diverge and feature different emphases at varying points during the work's duration, other interactions may come into play. The two perceptual streams may be experienced as independent, or one may dominate the other – typically, this tends to be the visual component. In some cases, the two perceptual streams interact and modulate each other, which is the phenomenon of synchresis.

The final category given in figure 11 is that of emergence – comparable to the earlier observation of the emergence of an image against a backdrop, this is the phenomenon where the audio and visual signals combine to create a perception in the audience that is more than the sum of its parts, a combined audiovisual object that is distinct from a simple combination of the two parts. In their work on speech recognition, McGurk and MacDonald illustrated this effect of emergence. "… on being shown a film of a young woman's talking head, in which repeated utterances of the syllable [ba] had been dubbed onto lip movements for [ga], normal adults reported hearing [da]".[47]

[47] McGurk, H. & MacDonald, J. (1976). *Hearing lips and seeing voices.* Nature, 264. (23 Dec 1976) pp746-748.

Anton Fuxjäger summarised the connections he observed in visual music into 3 main categories – 'Music Translations', where the visual form is a 'mapping' directly from the auditory form; 'Synthetic Structures', where the audio and visual forms may not necessarily coincide but will form an interacting structure underlying the piece; and "Mutual Disturbance", where there is insufficient correlation between the audio and visual forms, and the audience cannot identify any audiovisual structure.[48] Again, in reference to the observations above, his "Music Translations" may be compared to the idea of 'mickey-mousing', while his 'Mutual Disturbance' could be used as a base point to identify when audiovisual compositions simply do not work, do not gel as a cohesive structure between the audio and visual forms.

For the audiovisual composer, the 'Synthetic Structure' is the most fertile ground for creation, experimentation, and aesthetic satisfaction. The creation of a new perception from non-matching audio and visual elements is an integral part of the composition of audiovisual works. In her experimental work, Cohen has investigated the use of music to influence the perception of visual stimuli, and noted that where the meaning conveyed by the visuals is ambiguous, the choice of music can affect the interpretation made by the audience.[49]

For film sound designers, the use of all sound, not just music, can have a major influence on the final experience the cinema audience will have. The sound designer Gary Rydstrom has noted, "Like with the camera going in for close-up shots sometimes, the sound can be decided like all the visual elements… Think like a director or a DP, staging a scene with equivalent techniques like depth of field, and choose what the audience's ears are focused on."[50]

The interpretation of an emerging audiovisual object can be influenced by pre-cueing

---

[48] Fuxjäger, A. (2012). *Translation, Emphasis, Synthesis, Disturbance: On the function of music in visual music.* Organised Sound, 17(02), pp120-127.

[49] Cohen, A.J. (2005). *How Music Influences the Interpretation of Film and Video: Approaches from Experimental Psychology.* In R. A. Kendall & R. W. Savage (eds.), *Selected Reports in Ethnomusicology: Perspectives in Systematic Musicology*. (2005), Los Angeles, CA: Dept. of Ethnomusicology, University of California. pp15-36.

[50] Sonnenschein, D. (2001). *Sound Design*. Studio City, CA: Michael Wiese Productions. p162.

– in visual experiments, it was found that, "If an observer is given a cue of some kind to where a target is about to appear, a variety of psychophysical detection or discrimination tasks are performed more effectively in this location than in others that have not been cued."[51] By using the principles of common fate, good continuation, and others, the audiovisual composer can lead both the eye and ear towards a specific point in time and space within the composition to create an expectation that an event will occur – and by either satisfying this expectation with the anticipated outcome, or denying it with a silent void or an unexpected outcome, the composer can also add interest and variation into the work.

The Audiovisual Object

From Gestalt theories, we can discern figure and ground in both the visual and audio realms, where the sound object and the visual object can be discerned by the audience. The audiovisual object is formed at the point where the two components intersect in the audience perception, creating a fusion of sound and vision which have at least a few of the Gestalt properties in common, corresponding across the space and duration of the audiovisual experience.

In audiovisual works, there are both spatial and temporal aspects to the work. Normally, objects are portrayed on screen, and therefore have a spatial relationship to each other and the audience. As the objects move across the screen, the temporal aspect comes into play – if a person is walking, a regular beat pattern is seen on screen as their legs move. Similarly, the sound objects have both a temporal and spatial aspect. The temporal aspect is the most immediate – as a melody is heard, the notes follow each other in time, creating a progression of sounds. But the ear also perceives a spatial aspect, where the sound is perceived to originate from. That is an aspect of the sound which is frequently explored and exploited by the electroacoustic composer.

---

[51] Bruce, V. Green, P. R. & Georgeson, M. A. (2003). *Visual Perception: Physiology, Psychology and Ecology*. Hove, UK: Psychology Press. p413.

The simplest form of audiovisual object is that where a very direct synchronisation of sound and vision takes place – an effect often referred to as 'mickey-mousing'. In a short film called 'Dots'52 made by Norman McLaren in 1940, he took advantage of the development of optical sound film negative, drawing shapes on both the visual and sound tracks of the filmstrips. In McLaren's piece, you can see the shapes appear and develop, with a corresponding sound matching their actions frame by frame.

Moving on from simple synchronisation, there is the phenomenon referred to by Michel Chion as *synchresis*, "the spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time"53. If you run a recording of music against any piece of film, there will be moments where the music seems to catch and elevate the action you can see. This synchretic effect is used by sound designers to consciously mould a soundtrack that fits and complements the images caught by the film camera. The beats in music can be used to accentuate movements on screen, the sound of the elements can underscore action sequences, and so on.

Synchresis differs from synchronisation in that the sound and vision can move in separate directions, meeting every so often to emphasize the aspect of the story that is key to that moment. Not everything has to match, and the freedom of movement in both perceptual streams allows a reasonable amount of creativity on the part of the film and sound editors, and the composer.

The Audiovisual Environment

By using a time-based visual medium such as film or animation, we can create a perceptual stream of visual objects which can change in morphology over time, and be processed and changed in similar ways to the electroacoustic composition world –

[52] McLaren, Norman. (1940). *Dots*. Animation, accessible at https://www.youtube.com/watch?v=E3-vsKwQ0Cg [accessed 17/08/2015].
[53] Chion, M. (1994). *Audio-Vision: Sound on Screen.* Columbia University Press. p63.

an electrovideo world. By combining the electroacoustic with the electrovideo, we start to create an audiovisual world. Chion notes there are natural synthesis points, where randomly combined pieces of sound and vision come together and fuse to create a bi-perceptual sensation experienced by the audience. As noted above, he terms this synchresis, a portmanteau of synchronisation and synthesis.

While natural synchresis can be observed by matching soundtracks to video and observing where the sound and visuals are perceived to merge and create an audiovisual fusion, these points of synchresis can also be specifically planned for. For example, sound designers will create sound effects to achieve a specific syncretic point in a film, or a composer for film will time the music beat to intersect with moments of action on screen. A good film/video editor will consciously cut the video element in time to allow for such synchronisation. These points of synchresis are audiovisual objects. As with sound and visual objects, they are separate from the originating sound and video, even if they are a faithful recreation of a live recording. They can be manipulated and planned for, and those that highlight a specific audiovisual moment in the audiovisual experience can be regarded as gestures, corresponding to gesture in electroacoustic composition.

Given the concept of synchresis, there are also the moments in audiovisual work where the sound and vision do not merge. This also can be deliberate, for example as a compositional tool to develop themes in both perceptual streams, which may later combine to form a powerful audiovisual fusion, the impact of which is greater due to the anticipation built up in the separate themes.

Diego Garro has recently put forward a continuum for the relationship between the audio and visual perceptual streams in audiovisual work. His continuum is reproduced below:

Figure 19: Garro's Continuum of Audiovisual Mapping

Garro's continuum can be used to place the relationship of the audiovisual fusion from non-existent at Separation through to totally synchronised at Parametric mapping – separation referring to no discernible connection between the two perceptual streams, while parametric mapping exhibits a very close relationship between sound and vision. As Garro explains, "The continuum presumes, from left to right, a progressively growing strength of the associative links between sounds and visuals. In most circumstances this will correspond to the experience of viewers/listeners: spectromorphological ties in parametrically mapped passages will provide a more convincing gestural integration than synchresis. In turn, the temporal bonds in synchretic gestures will in most cases yield stronger coalescence than a loose, and often subjective, complementarity between the constituents."[54]

This gives the audiovisual composer an approach to considering the structures of the audiovisual interactions that occur in his piece. However, this continuum does not necessarily also map to the effectiveness of the audiovisual interactions. For example, there is a phenomenon known as 'mickey-mousing' where the interaction between the sound and vision appears too artificial, and the aesthetic experience

[54] Garro, D. (2012). *From Sonic Art to Visual Music: Divergences, convergences, intersections.* Organised Sound, 17(02). pp103-113

suffers as a result. This problem has been commented on from the early days of commercial animation, and was particularly associated with the earlier products of the Disney studios, hence the name. Writing in 1946, the animator Chuck Jones observed, "For some reason, many cartoon musicians are more concerned with exact synchronization or 'mickey-mousing' than with the originality of their contribution or the variety of their arrangement. To be sure, many of the cartoons as they reach the musician are something less than inspirational, but most of them, even the best, gain less than they should from his contribution. I have seen a good cartoon ruined by a deadly score."[55]

Garro's continuum may therefore need to be modified slightly if it is also to be taken as a guide to the effectiveness of the audiovisual combinations. The 'dip' in aesthetic satisfaction that occurs with 'mickey-mousing' can be envisaged as occurring towards the parametric mapping end of his continuum. If we regard the 'aesthetic' satisfaction of the audiovisual fusion as gaining strength up to this dip, and then returning for full parametric mapping (live action, or complete synchronicity as can be seen in McLaren's *Dots*), the curve that results bears strong resemblance to the 'uncanny valley' envisaged by Miro in his article, originally published in Japanese but more recently available in English translation[56].

Miro was examining the human response to objects that resembled human beings or parts of human beings. His observation was that as the object starts to resemble the human likeness, it became increasingly familiar or comfortable to look at or interact with. A stuffed animal with anthropomorphic qualities or a humanoid but identifiably mechanical robot were reasonably comfortable for humans to deal with, as was a normal healthy human being. However, where the likeness was almost but not quite human-like, the reaction dipped into a negative, uncomfortable zone, which he called the uncanny valley. His conception of this zone is illustrated below:

[55] Jones, C. (1946). *Music and the Animated Cartoon.* Hollywood Quarterly, Vol. 1, No. 4 (July 1946). University of California Press. pp364-370.

[56] Mori, M. (1970). *The Uncanny Valley*. tr. MacDorman, K.F. & Kageki, N. (2012), Available at http://spectrum.ieee.org/automaton/robotics/humanoids/the-uncanny-valley [accessed 23/08/2015].

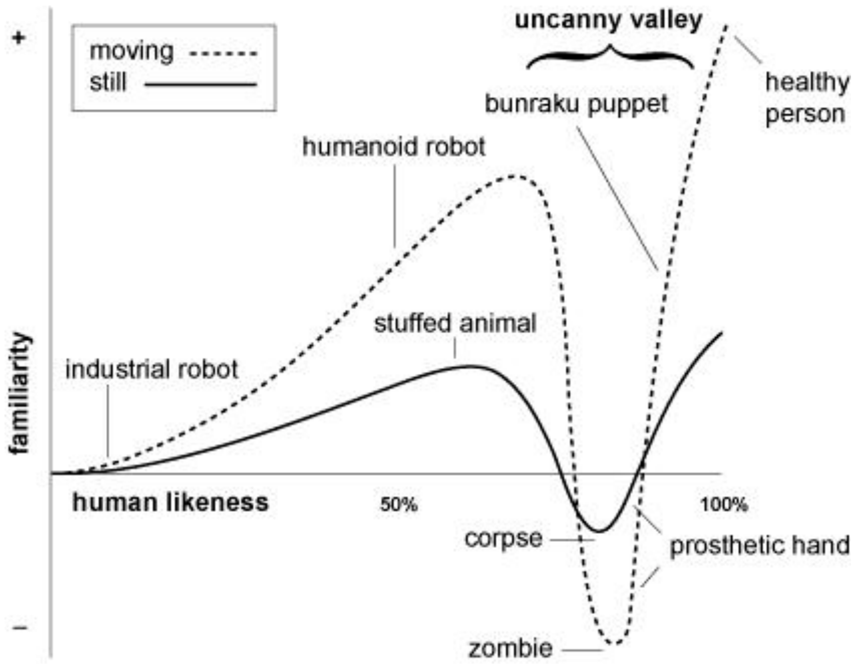


Figure 20: Miro's Uncanny Valley

The 'mickey-mousing' effect has a similar dip in familiarity, "comfortableness", or aesthetic satisfaction for the audience, and can cause the viewer to 'drop out' of the audiovisual experience, or at least cause a disruption in the compositional flow of the piece.

Using Garro's original continuum as a base, and mapping an 'uncanny valley' effect to account for 'mickey-mousing', the following diagram provides a guide to the audiovisual composer when considering how their individual audiovisual gestures will be received by the audience.

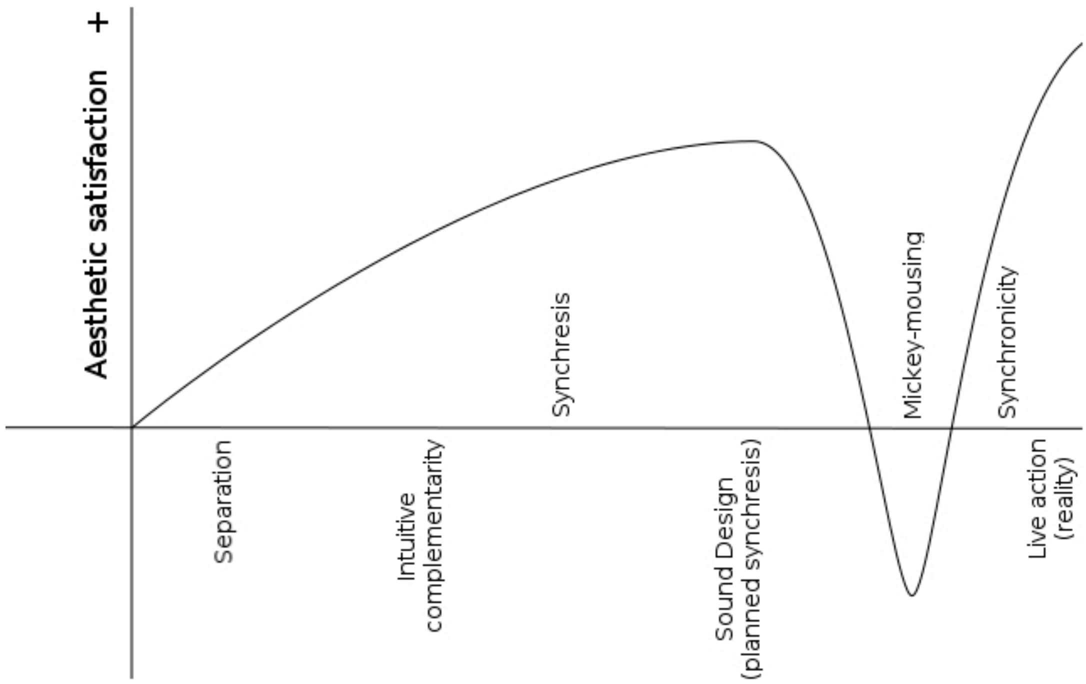


Figure 21: The Uncanny Valley Audiovisual Mapping Continuum

Other than a true reproduction of a live-action recording, the most aesthetically satisfying audiovisual experiences for the audience are spontaneous synchresis, planned synchresis, and synchronicity. It follows that for the audiovisual composer, work exhibiting these three combinations of sound and vision are likely to be the most effective at capturing and retaining the audience's interest, although narrative structure and a coherent visual and sonic palette are also essential in crafting engaging work.

Conclusion

The successful combination of visual and sound objects into a fused audiovisual object provides building blocks that an audiovisual composer can use in their own practice. By using the Gestalt principles outlined above, the composer can not only analyse existing works but also develop their own approaches to creating the objects and themes they wish to use in their own compositions. The audiovisual environment is also important, in that it forms a ‘basal layer’ or background to the objects featured in the foreground, defining the characteristics of the overall piece, and allowing exclusive allocation to occur to highlight the audiovisual objects the composer wishes to foreground.

The consistencies in stable synaesthesia and cross-modal perception also point towards creating a stable, consistent basis for the audio-visual interactions and objects created by the composer. As the evidence points towards the same basic sensory methods for interpreting those interactions in synaesthetes and non-synaesthetes, by maintaining the stability evinced by most synaesthetes, the composition will maintain a coherent internal structure that will both be more accessible and more aesthetically pleasing to the audience.

The creation of audiovisual objects is not limited to a simple synchronisation between sound and image. By making use of synchresis, either achieved through a fortuitous link created by aligning sound and vision together, or by a planned design process, the composer can create audiovisual objects that satisfy the human tendency to seek a fusion between the two perceptual streams. The synchronous and synchretic fusions of audio and visuals can be established within a wider audiovisual environment, where the two perceptual streams can also feature sound objects and visual objects, where these can interact, form counterpoint to, or provide a base form from which audiovisual objects can emerge. By manipulating audio, vision, and audiovision, the composer can create an immersive work that engages the audience’s attention fully.

The approach that I have taken to establishing techniques for audiovisual composition arise from those common in electroacoustic composition, using the sound object as a base material from which a coherent electroacoustic composition can be created. In my own practice, I usually approach the audiovisual composition process from the audio component first, although I build up an imagined visual extension as the work takes shape. For me, this creates a stronger fusion of audio and vision than can be achieved working in the opposite direction. However, the alternative of matching appropriate sound objects to visual elements to create a strong audiovisual fusion is an every day matter for sound designers and film composers, illustrating that the successful audiovisual composer can approach their practice in a variety of ways.

The development of ideas concerning synaesthesia, *Gesamtkunstwerk*, and the Gestalt approach to perception occurred in tandem with each other over the later 19th century and 20th centuries. The development of cinematic technology, abstract animation, and recorded music also occurred across a similar time span, from the start of the 20th century. It is worth examining the development of animation from its earliest origins and comparing it to the development of electroacoustic music, particularly from Schaeffer's thoughts on the sound object onwards, to identify correlations and contrasts. The next chapter looks at the crossovers between the two art-forms more closely.

Chapter 3

The Development of the Audiovisual Object Palette

Introduction

At the heart of the audiovisual object is the ability to fuse sound and visual objects from a vast range of possibilities. This has arisen to a large extent from the freedom allowed in music composition from the expansion of what could be defined as music, with greater acceptance now of both modified recorded material and purely synthesized sounds. Similarly, the freedom afforded in animation allows the audiovisual composer an infinite variety of visual objects to work with, from drawn animation, photographic manipulation, and CGI work amongst many other options. This chapter explores the development of technology and creative techniques in both electroacoustic composition and animation, from the early explorations in their respective technologies through to current contemporary practice. The implications for the audiovisual object and the creation of an immersive audiovisual work are examined in context of the correlations between the two developing creative strands.

Over the 20th century, there was a gradual development of ideas about what could be defined as music, with an acceptance that the purely abstract form of musical notes played by pitched instruments could be complemented with a concrete form of recorded sounds that could be modified and altered by the composer. This allowed a greater freedom for the composer to create sonic work that could include recorded or synthesized material as well as or instead of classical instrumental material. The development of purely synthesized and sampled popular music towards the end of the 20th century indicates how prevalent this freedom has become, although what we would now regard as 'classical' acousmatic music is still a niche market.

In visual terms, that same freedom of expression, the ability to create new material or modify recorded material, is afforded by animation. In the latter half of the 20th century, animation in the Western hemisphere adhered for the most part to principles and techniques established by Walt Disney, to impose his own aesthetic views on the animators in his studio. Given his commercial success, these practices were rapidly adopted by the majority of the other American animation studios, and then by gradual extension, to studios based in Europe and, (to a lesser extent), Asia. It's worth looking at the animation of the 1910s and 1920s prior to Disney domination to see what freedoms were available then and what can be of inspiration to the audiovisual composer today, as well as looking at the auteur animation that has developed in parallel with the popular market, where the animators still made use of the ultimate freedom of the visual animation form.

This chapter examines the development of electroacoustic music from Russolo's proposals for a music based on everyday noises through to the parallel developments of *musique concrète* and *electronischen musik*, both of which greatly expanded the definition of what could be considered as music and therefore increased the palette of sound objects available to the electroacoustic composer. The development of animation is also examined, with particular focus on animation which also allows the widest possible palette of visual objects and transformations for the animator.

Although experimental work in recorded and tape material took place in other centres, such as New York, San Francisco, and the BBC Radiophonic Workshop in London, I will focus on the developments that took place in Paris and Cologne, as these have the most direct relevance and influence on current electroacoustic theories. Similarly, although early animation took place across the world, I will also concentrate on its development in the western hemisphere, particularly in the USA, as this style of animation came to dominate the art-form for many years, and on the development of visual music as an abstract animated art form in Germany in the early 20th century.

The Art of Noises

Although his training was as an artist rather than as a musician, and his politics were very right-wing, the Futurist Luigi Russolo is regarded as a seminal figure in the development of electroacoustic music. His thoughts on what should be accepted as music in the early days of the 20th century have been influential for many composers over the past hundred years, Pierre Schaeffer amongst them. The futurism movement promoted youth, dynamism and creativity, scorning the older established cultural forms. The futurists called for the overthrow of the old, by violent means if necessary, and glorified the 'cleansing effects' of war. While the politics of the movement may have been dubious to modern eyes, the promotion of a new dynamism promoting new ideas in art and culture was carried on in other successive movements, such as surrealism and Dada.

Russolo's innovation was to call for noise to be regarded as equally if not more valid a musical form as the pitched music performed on classical musical instruments. He regarded the increase in technological development as a herald of a new aesthetic in musical composition, and thought that the new everyday sounds of traffic and machinery should be the basis of new futurist music. In his 1913 manifesto, *The Art of Noises*, he contested, "Nowadays musical art aims at the shrillest, strangest and most dissonant amalgams of sound. Thus we are approaching noise-sound. This revolution of music is paralleled by the increasing proliferation of machinery."[57]

Russolo felt that "musical sound is too restricted in the variety and the quality of its tones... We must break at all cost from this restrictive circle of pure sounds and conquer the infinite variety of noise-sounds."[58] In a challenge to the orthodox orchestral instruments, Russolo invented a variety of instruments he dubbed 'intonarumori' to create, amongst other noises, roars, whistles, crashes, screeches and hisses. While Russolo's instruments never gained popularity apart from some use as theatrical sound effects, they were an early analogue version of the sounds and effects we can now achieve using tape machines or digital editing. His thoughts on noise as music inform a growing movement amongst composers across many nations at the time, who as part of the avant-garde were questioning all the perceived wisdom of the previous generations.

---

[57] Russolo, L. (1913). *The Art of Noises*. Available at http://www.artype.de/Sammlung/pdf/russolo_noise.pdf [accessed 5 December 2014].

[58] *Ibid.*

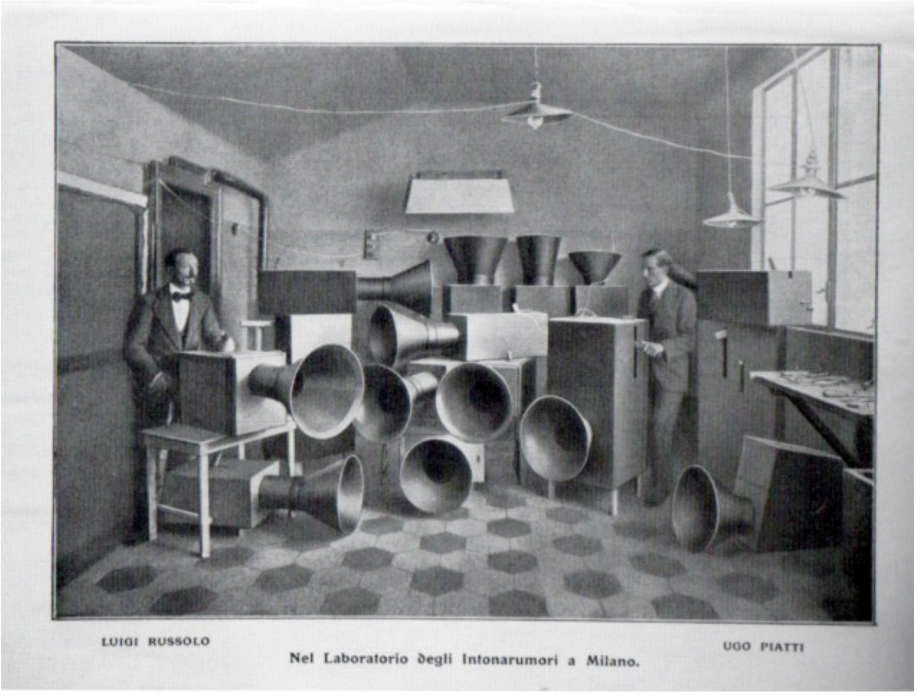


Figure 1: Russolo's Orchestra of Intonarumori for Noise Music[59]

Another avant-garde composer, Edgard Varèse, has been attributed with the concept that music is "organised sound", a concept that is often applied to electroacoustic music[60]. Therefore, if we regard the notes and phrases of classical music as a method of organisation of those specific sounds, then a similar organisation of sound objects (as per Pierre Schaeffer's observations) can also be regarded as music.

Electricity in Music

With the rapid development of electrical science and engineering over the 19th century, musical instrument inventors took notice and explored the potential of the new technology. While the earlier instruments initially did not contribute to electroacoustic composition, which would develop more from recording and

[59] LTM Recordings. (2014). [image online]. Available at http://www.ltmrecordings.com/musica_futurista_the_art_of_noises_ltmcd2401.html [accessed 8 December 2014].

[60] Chagas, P.C. (2014). *Unsayable Music: Six Reflections on Musical Semiotics, Electroacoustic and Digital Music*. Leuven University Press. p90.

playback technology, the principles of synthesizing sound would become an intrinsic part of *elektronischen musik* and the work carried out by Eimert and Stockhausen in Cologne.

One of the first of these electrical instruments was the *Telharmonium*, invented in 1897 by Thaddeus Cahill. This early electronic organ made use of tone wheels for additive synthesis, making use of a base tonewheel to create a pure sine wave tone, with extra tonewheels employed to add harmonics and timbre. The sound output came from paper cones, with the signal fed through on telephone wires. The instrument was very large and heavy, with later versions weighing in at 200 tons. It also consumed an immense amount of power while in use – restrictions which made the instrument unviable for mass production. The *Telharmonium* was followed by a series of new electrical instruments. The *Theremin*, invented in 1928 by Leon Theremin, still has its adherents today, with its pure sine tones and glissando between notes, due to its unique control mechanism for pitch and volume via two antennae, which act as capacitor plates reacting to the proximity of the player’s hands. The *Ondes Martenot*, also invented in 1928 by Maurice Martenot, bears a resemblance to a harpsichord attached to a variety of amplifying speakers. The timbres it produces are similar to the theremin, but more variation of timbre is available, and the construction allows for stepping between pitches rather than the continuous glissando of the theremin.

While the theremin is notable for its unique method of control, most of the new electrical instruments continued to make use of the organ or piano keyboard as a standard control interface. The *Hammond Organ*, invented in 1935 by Laurens Hammond and John M Hanert, also made use of tonewheels, varied by the use of drawbars. But the appearance of the *Clavivox Synthesizer* in 1956, created by the composer Raymond Scott and a young Robert Moog, heralded a move towards the modern digital synthesizer keyboard, which became an important force in popular music in the 1980s and beyond. At the same time, pick-up and amplification technology had been developed for other instruments such as the guitar, leading to the electric guitar and bass guitar, which also came to dominate popular music in the latter half of the 20th century.

In terms of the audiovisual object, the importance of these instruments is in the experimentation with tone and timbre, and eventually in the artificial synthesis of sounds. This expanded the available sounds that could be used in music composition, mirroring the development of theories where composition was beginning to be regarded as the organisation of sounds, which did not necessarily have to be created by traditional pitched instruments. For the audiovisual composer, the expansion of the sounds that can be regarded as musical, or as having the potential to form part of a musical composition, allows for a much greater sound palette to fuse with images to create the audiovisual objects necessary for their compositions.

Recording Technology and Experimentation

The invention of the phonograph and gramophone (1877 by Thomas Edison, and 1887 by Emile Berliner, respectively) heralded a new age for music and sound composition. As the technology developed, it was now possible to listen to your favourite music recordings whenever you wanted to, as opposed to having to play an instrument yourself, or hoping a concert would take place soon with an appealing programme. The gramophone in particular allowed some modification of the recorded sound. Pierre Schaeffer, having set up an experimental radio sound laboratory under the auspices of the French Radiodiffusion Nationale in 1943, started experimenting with record discs in 1948, cutting between the grooves to create basic loops and samples. He used these techniques in one of his earliest works, *Étude aux chemins de fer*, which consisted of manipulated recordings of steam trains.[61]

Schaeffer continued his experimentation in conjunction with the composer Pierre Henri, moving to a new studio in 1951. This studio contained the latest technology – most importantly, tape recording machines. Schaeffer could now carry out much

[61] Holmes, T.B. (1985). *Electronic and Experimental Music*. Charles Scribner & Sons, New York. p120.

more manipulation of the sound, reversing the tape, changing the playback speed, passing the sound through filters, and (most crucially) splicing tape together to create seamless edits of recorded material. Schaeffer made use of these new techniques to create compositions from recorded material within a new compositional form he termed *musique concrète*, which indicated that sounds derived from actuality rather than the abstraction of normal music, realised by instruments following notations. His studio, known as the *Groupe de Recherche de Musique Concrète (GRM)* flourished and is still a centre for electroacoustic composition today.

While Russolo had wanted to create a new musical form made up of the everyday noises around him, but had been limited to the instruments he could manufacture, Schaeffer now had everyday recorded sounds as a source for his creativity. This new form of music rapidly developed new followers and adherents – amongst others, the composers Luc Ferrari, Iannis Xenakis and Edgard Varèse all experimented with *musique concrète*.[62]

Schaeffer's work in this area led to several publications, culminating in his 1966 work, *Traité des Objets Musicaux*. This brought forward several ideas, including reduced listening and the idea of the sound object, as described in Chapter 2. This idea of a sound object that can be used as a building block in composition forms the basis of my proposal that the audiovisual object can also be defined and used in similar ways, as a building block in audiovisual composition. As with the sound object, the audiovisual object is a distinct phenomenon in its own right, formed by a fusion between sound and visual objects. As with the sound object, audiovisual objects can be regarded as the 'building blocks' of an audiovisual composition, with designed 'phrases' of audiovisual objects combining to create a composite audiovisual object within the wider realm of the audiovisual environment of the composition. Chapter 5 outlines several audiovisual pieces as case studies, examining the audiovisual objects in each piece with reference to the Gestalt Principles described in Chapter 2.

---

[62] *Ibid.* p122.

At the same time as Schaeffer started his work in his new tape recording studio (1951), an electronic music studio was also created in Cologne in Germany, under the auspices of *Nordwestdeutscher Rundfunk (NWDR)*. While initially under the guidance of Herbert Eimert, the studio became famous for the work of a newly emergent talent – Karlheinz Stockhausen. The Cologne studio was set up to further the development of electronically synthesized music, rather than the manipulation of recorded sounds as in Paris. The two studios both made use of tape recording equipment and other sound manipulation technology, but to "say there was a rivalry in the early days of these studios would be an understatement"[63].

This pursuit of new sounds and timbres made by electronic synthesis followed on from the theories of serialism put forward by Webern, where the new generation of composers were pushing for ever more control and organisation of the music they were creating. In Schaeffer's work, he concentrated on listening to sound and noise, creating a modulated experience of the concrete world. In the case of the Cologne composers, their aim was to control both the composition of the music and the generation of the sounds used to realise the composition. While the initial compositions, using oscillators and noise generators, were fairly uninspiring, the work that Stockhausen and his contemporaries put into synthesis soon saw a blossoming of the musical form.

In the decades since the two studios were started, both the Parisian school of *musique concrète* and Cologne's electronic synthesis have become stock tools-in-trade of the electroacoustic composer, with the once important lines between the two becoming blurred and merged. Similarly, the audiovisual object can be formed of sound and vision sampled from real life, as per the founding principles of *musique concrète*, or the sounds and vision can be completely synthesized, as per the Cologne School – or any degree in between. Again, this offers an incredibly rich potential palette for the audiovisual composer to draw from.

---

[63] *Ibid.* p64.

Consolidation and further theoretical explorations

While Schaeffer's initial work on reduced listening and the sound object has been at the core of the work of many electroacoustic composers, there have been several further developments in the theory of electroacoustic music over the years. In keeping with the idea of 'organised sound', several of the most influential works in the field are concerned with clarifying how sound is or can be organised. Trevor Wishart, whose practice makes use of a wide vocabulary of human voice recordings and manipulations, both at the point of recording and in later processing, considered that two main traditions have been established in modern electroacoustic composition. The first derives from John Cage's work, where materials are assembled as a montage based on what they represent. Wishart contended that this represents a drive to illustrate the *idea behind* the sounds. The other tradition is much more aligned with the original work carried out at the GRM, and is concerned with the *heard relationships* between the sounds[64]. These observations do not detract from the idea of the sound object itself, but offer the composer an insight into their own way of assembling and composing from recorded and synthesized material – an insight which also maps across to the use of audiovisual objects.

Wishart also proposed that the idea behind the sounds generally evokes an associated image that the audience will imagine, based on a shared cultural heritage.[65] This could be a shared visual image, such as the sound of waves evoking the sea, or rely on associated aural cues, such as additional white noise or hiss altering our perception of an orchestral piece from being a straightforward recording to one that is being reproduced over a radio. This has some resonance with the Ecological approach to perception put forward by Gibson, noted in Chapter 2, where the perception of an object is afforded an interpretation based on a previous association, although in this case Wishart proposes a shared cultural affordance rather than an individual one. The sounds familiar to a UK native will evoke a certain reaction

[64] Wishart, T. (2012). *Sound Composition.* Orpheus the Pantomime. p4.
[65] Wishart, T. (1986). *Sound Symbols and Landscapes.* (Chapter 3 from Emmerson, S. (ed.) (1986), *The Language of Electroacoustic Music.* Macmillan Press Ltd. pp41-59.)

when heard within a composition, a reaction which may not be present in a Japanese listener, or whose reaction may be very different. Simon Emmerson, an influential theorist in the field, has termed this association that the listener has between the musical material and an associated cultural or symbolic link 'mimesis', and notes that it can both be an unconscious connection made within the composition, or deliberately exploited by the composer – though there is the issue of different cultural backgrounds affecting how the mimesis will work, and so planned mimesis may not always be as effective as the composer wishes[66]. It is noticeable also that Emmerson explicitly draws a link between the sounds heard in electroacoustic music and visual or synaesthetic images evoked in the mind of the listener - "the term 'image' may be interpreted as lying somewhere between true synaesthesia with visual image and a more ambiguous complex of auditory, visual and emotional stimuli."[67]

There is a recognition here that the human mind automatically chooses to associate sound with image, a natural synaesthetic or multi-modal reaction to the everyday world of associated sounds and images which the audiovisual composer elevates into an organised sound-image art-form. Current neuroscientific research recognises that "perception is fundamentally a multisensory phenomenon. There can be no doubt that our senses are designed to function in concert and that our brains are organized to use the information they derive from their various sensory channels cooperatively"[68]. By observing the pairing of sound and vision in natural combinations or in aesthetically appealing audiovisual works, the composer can build a library of successful audiovisual objects which encourage audience immersion in the complete composition.

As well as his thoughts on the ideas behind the sounds used in composition, John Cage was influential in other aspects of electroacoustic music theory, particularly

---

66 Emmerson, S. (1986). *The Relation of Language to Materials.* (Chapter 2 from Emmerson, S. (ed.) (1986), *The Language of Electroacoustic Music.* Macmillan Press Ltd. pp18-39.)

67 *Ibid* p18.

68 Calvert, G.A. Spence C. and Stein B.E. (eds.) (2004). *The Handbook of Multisensory Processes*. MIT Press. p19.

where theorists made use of his definitions of the components of sound – frequency (pitch), amplitude (loudness), timbre (tone colour), duration and envelope. Denis Smalley developed the definition of the structure of individual sound objects much further, with his work on the spectromorphology of sound. Taking the spectral typology of the sound as a starting point, he defined basic shapes or morphologies that develop in the sound over time, and the directional tendencies of sounds as perceived by the listener[69]. He then makes use of these spectromorphologies to structure his own compositions. This further investigation and more detailed definitions of the sound object can be extended to the audiovisual object on both the audio and visual level – while we may use the developing amplitude of a sound object to create a sense of increased tension in the listener, we can also develop a visual property such as the colour intensity, light levels or motion of the audiovisual object over time.

Modern Electroacoustic Composition

Modern electroacoustic composition occupies a particular ground, between the traditional, classical Western tradition of composition, and the popular electronic music which appeals across international boundaries. These divisions are, however, slightly artificial, with many ‘classical’ composers making use of digital synthesis and recorded sounds in their work along with traditional orchestral instruments, and composers such as Stockhausen exerting influence on popular music made with electronic instruments (Stockhausen appears on the cover of the Beatles’ *Sergeant Pepper’s Lonely Hearts Club Band* LP, released in 1967).

The earlier divisions between the manipulated real-world recordings of *musique concrète* and the electronic sound synthesis of *elektronischen musik* have fallen away, with the modern electroacoustic composer happy to use whichever sounds are the best fit for his or her creative work, no matter the origin. Despite this aggregation

[69] Smalley, D. (1986). *Spectro-morphology and Structuring Processes.* (Chapter 4 from Emmerson, S. (ed.) (1986), *The Language of Electroacoustic Music.* Macmillan Press Ltd. pp61-93.)

of theoretical backgrounds, the diversity between compositions is as great as ever, reflecting the wide palette of potential sounds the composer can choose from, although it is normal practice to choose some form of limiting the sound palette from the huge universe of sound objects available, as otherwise the work may lack any sense of coherence.

Paulo Chagas has observed, "the differentiation of *music and other artistic forms* increases environmental complexity, constantly redrawing the boundaries between different domains of observation."[70] *(my emphasis).* The development from the use of the sound object in electroacoustic composition to the use of the audiovisual object in audiovisual composition is a prime example of the growing diversity and complexity in the modern world of art and music, assisted by the development of digital technology in the same way as Schaeffer's work was dependent on the technological opportunities offered by sound recording and reproduction.

Freedom in Animation Forms – the Visual Equivalent of Electroacoustic Composition

So far, we have been mainly concerned with the development of electroacoustic composition and the freedom it offers to the composer. By opening up the definition of a valid musical composition to include sound objects derived from manipulated recordings or through synthesis, electroacoustic composition encourages an experimentation with the inherent properties of sounds and the effects experienced by manipulation and juxtaposition of sound elements against each other. The textural aspects of one sound may be used as a continuous background element against which more strictly defined gestural sound objects can be featured in the foreground, with the background texture then modified to become a foreground gesture in itself. The spectromorphological properties of the sound objects become an important element

[70] Chagas, P.C. (2014). *Unsayable Music: Six Reflections on Musical Semiotics, Electroacoustic and Digital Music*. Leuven University Press. pp155, 158.

of the compositional process in themselves.

Within electroacoustic composition, the underlying technology has a greater influence on the composer than in classical or popular composition. The composer may choose to work using existing technology or methods, but it is often the case that they will develop technology further to allow the creation of the sound object they regard as essential for their creative work. Di Scipio referred to this as two contrasting approaches or questions for the electroacoustic composer: “How can I use the available existing task-environment to realize my own idea of composition?- or - How can I design the tools that are necessary to realize my own idea of composition?”[71]. The use and further development of technology allows the composer an extremely high level of control over the composition process. Similarly, animators exercise a similar great deal of control over their animations, and, from the early experimentation with animated forms, such as Fischinger’s engineering of a wax-slicing contraption to create animated backgrounds, to the current investment in developing digital technology for animation by Pixar, Disney and similar companies, there is a similar willingness to embrace and develop technology to achieve specific artistic goals in the creation of animated work.

While we also have the freedom to choose what visual forms can complement the sound in audiovisual composition to fuse and create audiovisual objects, the cinematic format of animation exhibits many of those creative freedoms that are of great use to the audiovisual composer. When looking at the development of animation over the past century, the major difference between its development and that of electroacoustic music is the more commercial aspect – while the first few decades of animation allowed a blossoming of often anarchic visual freedom, the market soon came to exert its influence, and the more rigorous, controlled world of Disney came to dominate, influencing all the mainstream Western animation, even if companies such as the Warner Brothers did on occasion try to break out from this cultural stranglehold.

---

[71] Di Scipio, A. (1995). *Inseparable models of materials and of musical design in electroacoustic and computer music.* Journal of New Music Research 24: 34–50 p36.

The Origins of Animation

How long ago was animation invented? There is in fact a reference in *De Rerum Natura*, written by the Roman philosopher Lucretius (c.99 BC – 55 BC) in the 1st century BC that could be taken as the earliest reference to the basics of animation -

"For, as one image dies, another is born,
in another position, and so we think there is movement;
Of course all this happens at incredible speed."[72]

Lucretius' lines could almost be taken as a reference to persistence of vision, the succession of 'frames' that allow us to fool the eye into thinking that moving images reflect a version of reality.

In more practical terms, there are descriptions from China from around 158AD which refer to a pipe that produces dreams, which was probably an early form of zoetrope or magic lantern – a device which is often pointed to as a precursor of today's animated film. One description notes, "this was called *chao hua chih kuan* (the pipe which makes fantasies appear). Rising currents of hot air were evidently used by Ting Huan (c. 180 AD) who made a 'nine storied hill-censer' on which many strange birds and mysterious animals were attached. All these wonderful creatures moved quite naturally, presumably as soon as the lamp was lit."[73]

However, going back to Roman and early Chinese history for evidence of animation would be comparable to equating the start of electroacoustic music with lithophones. A more appropriate point to start comparing animation with electroacoustic composition would be from the development of the record-player and, more importantly, recording onto tape, allowing the processing and manipulation of sound and the start of electroacoustic composition, leading to Schaeffer's original work.

[72] Lucretius. (1976). *De Rerum Natura: The Poem on Nature*. tr. Sisson, C.H. Routledge, New York. pp124-125.

[73] Needham, J. (1962). *Science and Civilisation in China. Vol. 4: Physics and physical technology.* Cambridge University Press. p123.

This can be compared to the early days of moving images, the period we would consider the start of cinematic animation, which goes hand-in-hand with the development of cinematography. In fact, the earliest cinematographic films made use of many animation techniques, especially using ‘special effects’ for live-action films, which are in essence the techniques used for stop-motion animation, and which lie behind a lot of visual special effects even today.

The Start of Animated Films

The precursors of film animation exploited the phenomenon of persistence of vision – described by Peter Mark Roget in his 1824 publication, *Explanation of an optical deception in the appearance of the spokes of a wheel when seen through vertical apertures*[74] (often erroneously referred to as *Persistence of Vision with Regard to Moving Objects*). He described how the eye could be persuaded that a succession of images should be interpreted as one continuous motion, if shown rapidly enough and with enough illumination. His observations probably contributed to the invention of several devices in the 1800s that made use of this phenomenon – the thaumatrope, the zoetrope, the mutoscope and the praxinoscope. The kineograph, invented in 1868, also referred to as a flipbook, made use of a series of illustrations bound together as a book, which could be flipped through by the viewer to create a sense of motion or animation.

Only four years later, in 1872, the pioneering photographer Eadweard Muybridge was commissioned to capture a sequence of photographs to prove whether or not a horse has all its feet off the ground when trotting (“unsupported transit”). Muybridge managed to take 25 sequential shots per second, which proved that unsupported transit existed. He continued to capture sequential photographs, and made use of them in his own version of the praxinoscope. Muybridge’s early photograph

[74] From the Royal Society website: http://rstl.royalsocietypublishing.org/content/115/131.full.pdf+html [accessed 05/09/2015].

sequences have remained an important reference source for animators.[75]

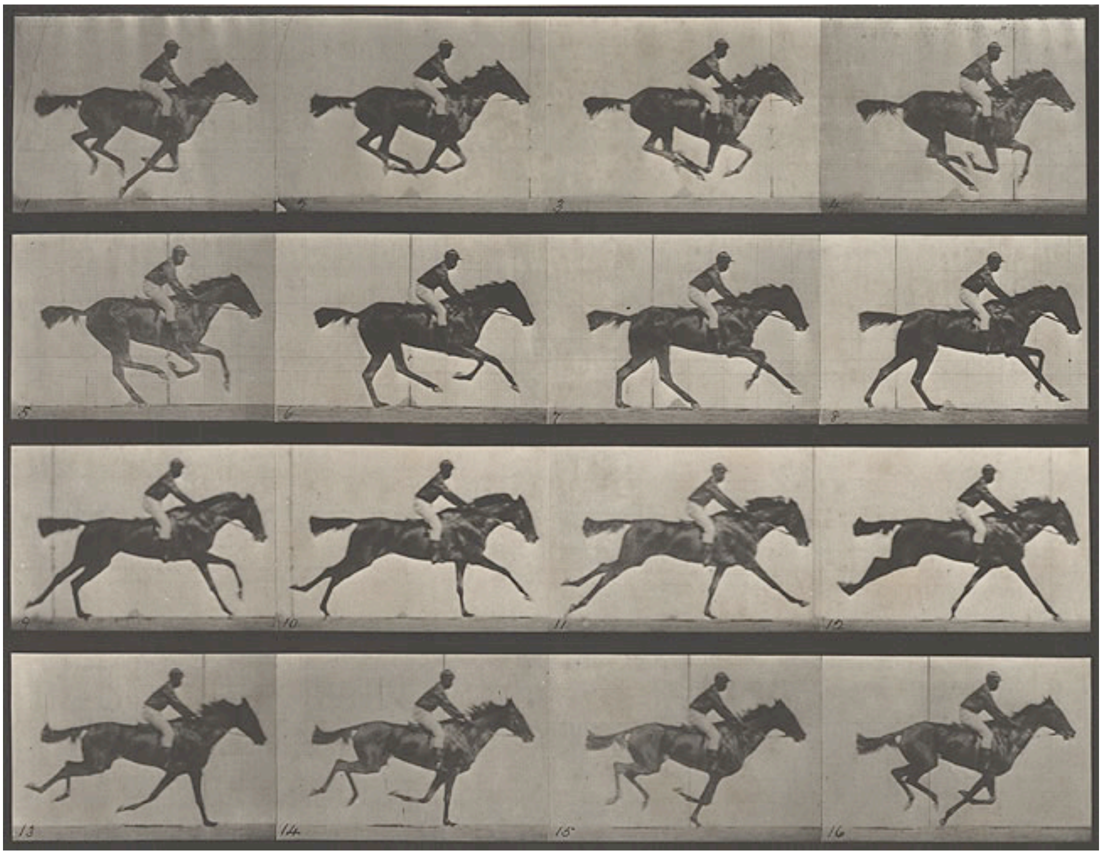


Figure 2: Muybridge's *Horse in Motion*[76]

These early experiments and apparatuses inspired Thomas Edison to create his motion picture projector in the late 1890s. The person who would go on to make the earliest acknowledged animated films, James Stuart Blackton, was at the time a cartoonist for the New York *Evening World*, who was called on to both sketch and interview Edison on his newest invention. The act of sketching the inventor was itself recorded by Edison, giving Blackton his first taste of motion picture creation. Blackton combined his artistic talents with this new medium, and continued to experiment with film. In 1906 he created what is now regarded as probably the first true animated film, *Humorous Phases*, which consisted of a series of animated

[75] Solomon, C. (1989). *Enchanted Drawings: The History of Animation.* Alfred A. Knopf, New York. pp9-10.

[76] Harry Ransom Center (2014).

sketches. But it was his next film, *The Haunted Hotel*[77], that was to establish his name as a key player in the field of animation.

The Haunted Hotel, released in 1907, proved an instant hit, and led to many other film producers trying to exploit animation techniques to achieve the same success. This film, particularly when it was released in Paris, resulted in a massive public reaction. The public were fascinated by the sequences where objects moved and interacted without human interference, and it sparked a reaction amongst French film producers to learn and reproduce the tricks shown on screen. Supposedly, these secrets of animation were carefully guarded. That is, of course, ridiculous. Anyone working in the film industry rapidly became familiar with jump cuts, where objects would appear completely in the 'wrong' place, if continuity was lacking during the filming. But animation made use of these 'errors' in a way that enthralled and intrigued the audience.

Hand-drawn Animation and the Auteur

As the film industry developed in the 1910s and 1920s, those making short films needed inspiration for their work. They specifically needed narratives that were short enough to film, but which could be guaranteed to hook the audience. On investigating other established media, they found that there was already a narrative, visually based genre that filled these requirements, and which had already established itself with the general public – the comic strip.

Strangely, however, the initial adaptations of the comic strip were not drawn animation, as you might have expected from our viewpoint 100 years later. In fact, the first screen versions were live-action, although often with the animated special effects that were used in *The Haunted Hotel*. Why was this? Maybe it's because it was a lot simpler and cheaper to shoot some live action, stop, move a few objects,

[77] Praytaa (2010). *La Maison Ensorcelée* (1907, Segundo de Chomòn). [video online] Available at http://www.youtube.com/watch?v=EqvtRPFnEMo [Accessed 7 December 2014].

and start again, than to physically draw the 3,500 odd frames you'd need for a 5 minute short.

Perhaps fittingly, given the success of *The Haunted Hotel* in Paris in 1907, the first hand-drawn animation was produced by a Parisian caricaturist and cartoonist . Emile Cohl, often referred to now as the Father of Animation (at least in Europe), was born in 1857 in Paris, and was right at the heart of the burgeoning avant-garde artistic movement. Having established his own style as a caricaturist and cartoonist in the 1870s-1900s, Cohl started working for the cinematic company Gaumont in 1906, initially employed to create story outlines[78]. By 1908 however, he had moved on to become a director, and had created his first film, *Fantasmagorie*, an experimental animation consisting of line-drawn characters, and featuring many surreal transformations of animals and objects into other forms. Although it appears to resemble chalk lines drawn on a blackboard, it was in fact drawn as black lines on white paper, then printed in negative, which also involved some further effects to incorporate Cohl's hands, which appear within the film on a couple of occasions[79].

Cohl would continue to work with animated film, experimenting also with stop-motion animation of objects, cut-out animation, and combinations of animation with live action, up until 1923. The First World War had interrupted his film-making career, and although he returned to the industry in 1919 to work on mainly commercial shorts, the aftermath of the war had resulted in a decimation of the French film industry and a new reliance on American imports. Despite his importance in early experimental animation, Cohl was destined to die as a pauper at the age of 81, having suffered from severe burns due to a fire in his apartment, complicated with pneumonia[80].

---

78 Crafton, D. (1990). *Emile Cohl, Caricature, and Film.* Princeton University Press. pp91-93.
79 *Ibid*. p121.
80 *Ibid.* pp214-216.

European Avant-Garde and Experimental Film

At the same time that American animation was developing from an extension of artistic illustration into a more commercial form, the pioneering film-makers in Europe had also started to experiment with the new medium. The end of the 19th century had seen the rise of a new form of art, a movement which would lead to what we now refer to as modern art. Artists such as Monet and Pissarro exploited the new synthetic pigments that were being produced in the latter half of the 19th century, and moved from the normal practice of painting in the studio to working outside in bright sunlight. The resulting artwork was lighter and brighter in tome than the accepted norms of the time, and the new movement became known as Impressionism. At the start, impressionist art was widely criticised by the established art world in France and Western Europe, but by the end of the century, many of the techniques were being adopted across the continent. The rapid spread of photography, and influences ranging from the British Arts and Crafts movement through to imported Japanese art and prints, also had their effect on bringing a new sensibility to the contemporary art scene. Experimental techniques such as pointillism, introduced by Seurat, became an inspiration for the new wave of modern artists[81].

At the start of the 20th century, Paris and Munich became centres of modern art that drew artists from across the world, inspired by the work of the impressionists and post-impressionists. Movements such as Fauvism, Expressionism, Cubism led on into Futurism, Dadaism and Surrealism. These flourishing artistic movements would become known as the avant-garde, a term that would encompass the radical and experimental nature of the work being produced. At the same time as these experimental art movements were gaining traction, the new medium of film was beginning to make its mark as an exciting new visual form, and it was not long before the avant-garde artists started to explore the potential for experimental works using the moving image as well as still images.

[81] Read, H. (1974). *A Concise History of Modern Painting*. Thames and Hudson, London. pp13-29.

The First World War undoubtedly resulted in a slow recovery of commercial film production in France and Germany, but in its aftermath the avant-garde adopted the medium as a fertile avenue within which to experiment. Louis Delluc, who espoused a rhythmic approach to film-making, comparing cinematographic rhythm to musical rhythm, created his first work, *Cinéma et Cie*, in 1919, but could lay claim to establishing the cinematic idea of montage with his 1921 film, *Fièvre*, where all the action takes place in a single room but the story is explored through flashback sequences juxtaposed with the current timeline[82], a technique later famously exploited by Eisenstein in his 1925 classic *Battleship Potemkin*. Delluc also organised a screening of works by Cohl in Paris on 12th June 1920, which would result in the establishment of the *Ciné-Club de France*[83], which rapidly became the inspiration for other film societies across Europe, which offered opportunities for like-minded artistic directors to meet, discuss their ideas and collaborate – which was to prove particularly important in the development of visual music.

The avant-garde movement included work created by artists and film-makers associated with the Dada and Surrealist movements. Man Ray explored both photography and film, with his 1923 work *La Retour à la Raison* showcasing his 'rayographs', silhouetted objects photographed by placing them directly on to the negative before exposure. Perhaps the first Dadaist film, there was no commercial aim behind it, it was made as an individual statement by the artist[84]. Ray later assisted Marcel Duchamp on a more surrealistic film in 1927, *Anemic Cinema*, which features a series of spinning concentric circles offset against each other, providing an animated optical effect, interspersed with slightly obscure puns between the sequences[85]. While these films relied on stop motion or optical effects, they reflected a growing experimental movement in the field of abstract film. At the same time as the Parisian avant-garde were applying their Dadaist and surrealist sensibilities to their cinematic works, a closer union of rhythm in sound and vision was beginning to

82 Dwoskin, S. (1975). *Film Is...* Peter Owen, London. p26.
83 Hagener, M. (2007). *Moving Forward, Looking Back: The European Avant-garde and the Invention of Film Culture, 1919-1939.* Amsterdam University Press. p80.
84 Dwoskin, S. (1975). *Film Is...* Peter Owen, London. p27.
85 *Ibid.* p27.

surface in Germany.

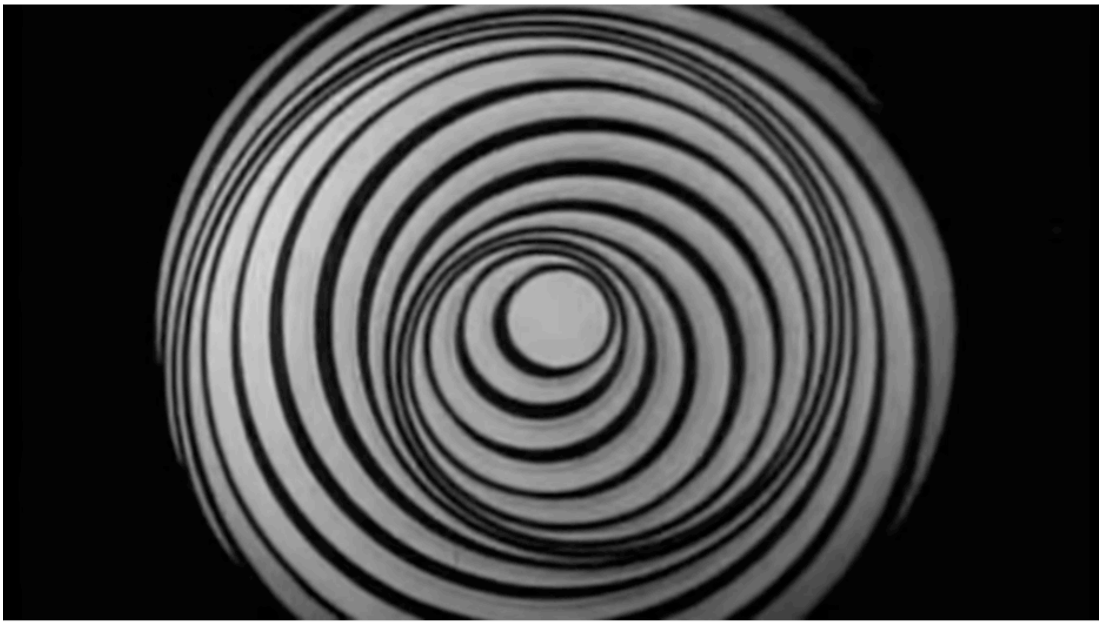


Figure 3: Still from Duchamp's *Anemic Cinema* (1927)[86]

Walter Ruttmann

Walter Ruttmann, a native of Frankfurt am Main, originally studied painting and music at the Academy of Fine Arts in Munich, and started his career as a graphic artist, which may have influenced his decision in 1920 to start a company to make 'drawn films'[87]. His work, particularly his *Lichtspiel Opus* series, featured amorphous blobs and ribbons, brightly coloured using hand-tinting techniques, moving across the screen in rhythmic patterns. The first *Lichtspiel Opus I* film was initially silent, but he would soon add music to the experience, scored to match the visual movements. Film-making techniques at this time were very laborious, involving painting on glass, manoeuvring cut outs, tinting, toning and colouring the film, which also meant that the film could not be quickly copied, but had to

86 Klarecki, C. (2011). Youtube upload of *Anemic Cinema* (1927). Available at https://www.youtube.com/watch?v=AsnhBUp7xxM [accessed 28/04/2017].

87 Keefer, C. & Guldemond, J. (eds.) (2012). *Oskar Fischinger 1900-1967: Experiments in Cinematic Abstraction.* Eye Filmmuseum / Center for Visual Music. p42.

essentially be reassembled each time a new print was made.[88]

As with Delluc, Ruttmann placed a prime importance on the use of rhythm in his work, explicitly making a connection with music. In an undated document, which probably was written in 1919 or 1920, he wrote that his ambition was to create, “a different expression of all known arts, to bring a whole new way of life into artistic form, ‘painting with time’. An art for the eye, which differs from painting in that it takes place in time (like music)... Since this art takes place in time, one of its most important elements is the time-rhythm of the visual event. For this reason, an entirely new art, hitherto only latently available, will be exhibited by artists who occupy a space between painting and music.”[89]

Ruttmann’s work was favourably received at the time, with reviews in leading journals at the time lauding praise on this new combination of music and art – In the Berliner Börsen-Courier, The critic Herbert Jhering remarked, “this attempt by Walter Ruttman as Artist and Max Bunting as Musician is one of the most interesting ones I’ve ever seen in film. It is nothing less than to the implementation of light and colour as music in a visible environment of tones.”[90]. As with the rise of film societies in Paris, members of Frankfurt-am-Main’s Friends of Literature Club took interest in film as art, especially Dr. Bernard Diebold, a theatre critic who espoused Wagner’s ideals of *gesamtkunstwerk*, and saw these new abstract and experimental films as a step towards a totally integrated artwork. Through this connection, Ruttmann’s *Lichtspiel Opus I* was seen by a young Oskar Fischinger at a private preview screening, which would prove a key moment in Fischinger’s own artistic development.

---

88 Moritz, W. (2004). *Optical Poetry: The Life and Works of Oskar Fischinger.* John Libbey Publishing, Eastleigh, UK. pp4-5

89 Goergen, J. (1989). *Walter Ruttmann - Eine Dokumentation.* Freunde der Deutsche Kinematik, Berlin. pp73-74, translated by Andrew Connor.

90 *Ibid.* p99.

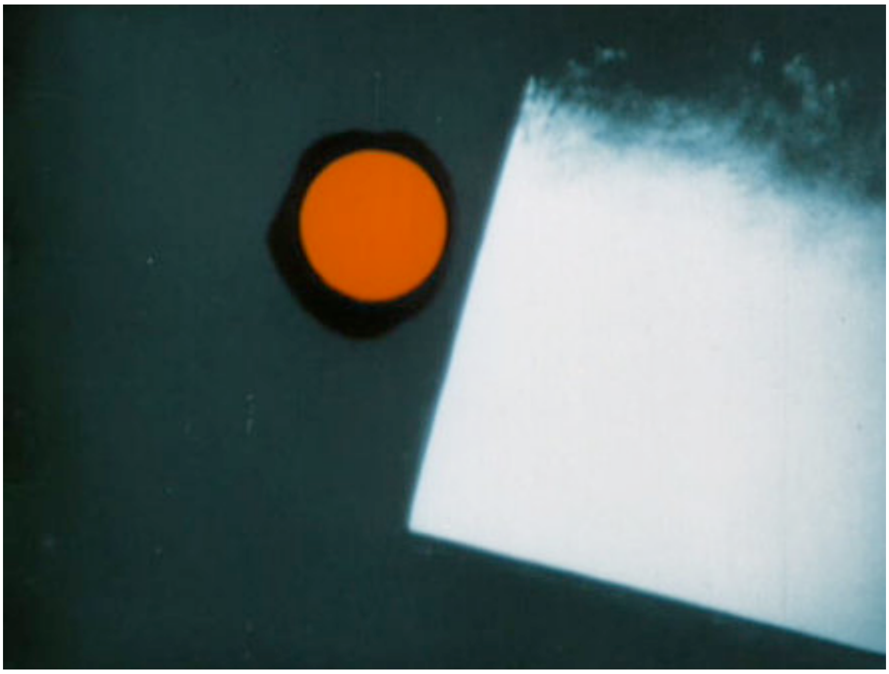


Figure 4: Still from Ruttmann's *Lichtspiel Opus I* (1927)[91]

Ruttmann continued to work on animated film, but is perhaps best remembered for his work *Berlin: Symphony of a Great City* (1927), a montage assembly of footage shot around Berlin, which offers an insight into Germany in the mid-war period, before the rise of the Nazi party. His death in Berlin came as a result of injuries sustained as a photographer on the front line of the war in 1935.

Viking Eggeling and Hans Richter

At the same time as Ruttmann was working on his *Lichtspiel Opus I*, Viking Eggeling was working on his own abstract animations in Brandenburg, near Berlin, as a guest of Hans Richter. Eggeling's initial film, now sadly lost, was *Horizontal-Vertical Orchestra* (1920-22), which, despite the title, was specifically not meant to have a musical accompaniment.[92] Eggeling's only surviving animation, *Diagonal*

[91] Image sourced from http://www.see-this-sound.at/print/work/175 [accessed 29/04/2017]
[92] Vergo, P. (2010). *The Music of Painting.* Phaidon Press, London p271 – contemporary critic Adolf Behne noted that the film was intentionally silent.

Symphony (1921-24), shows a marked attention to intricate detail rather than the simpler shapes of Ruttmann's work. Eggeling seems to have been more interested in the technicalities of animation than the aesthetics of simple movements on screen.

Diagonal Symphony was also intentionally silent, and in comparison to Ruttmann's hand-tinted work, intentionally black and white. Eggeling intended the intricate forms and their development through the animation to be the focus of the audience attention, without distraction by sound or colour. The regularity of his drawn structures with seeming mathematical precision recalls the vertical dimension of music, with the tones and semitones defined in the Western octave, while their logical development recall the regularity of music in its horizontal dimension, the standard rhythm patterns and the development of melody.

Eggeling's early death in 1925 deprived us of the chance to see how his ideas would have developed. His adherence to black and white to concentrate the viewer's attention to the development of forms on screen may have influenced Fischinger's later decision to create his series of *Studies* in black and white.

Figure 5: Still from Eggeling's *Diagonal Symphony* (1921-24)[93]

In a similar fashion to Ruttmann and Fischinger, Hans Richter's childhood fascination with music and art led to his interests in painting, sculpture and film-making. His early work focussed on Cubism, but he was soon to turn to Dadaism, attracted by the opportunity to "blow everything sky high as all the others did"[94]. Both Ruttmann and Richter served in the First World War, which left a traumatic legacy in their artwork – Richter drew several illustrations in 1917 showing mutilated corpses and bodies hanging on barbed wire. The end of the war and his acceptance into the ranks of the Dadaists gave Richter a feeling of liberation, and inspired his search for a positive principle behind his art.[95]

93 Image sourced from http://anthologyfilmarchives.org/collections/reference-library/stills/1218 [accessed 28/04/2017]

94 Grey, C. (ed.) (1971). *Hans Richter by Hans Richter*. Thames and Hudson Limited, London pp20-24.

95 *Ibid.* pp30-37.

Richter met Eggeling in 1918, and struck up a close friendship, with Eggeling staying at Richter's home from 1919 to 1922. They had a similar approach to their artwork, and both experimented with long scrolls of images which would form the inspiration for their films featuring abstract animation. Richter created several animations, although he later developed an abstract style making use of real life footage and montage. His *Rhythmus* series mainly echo some of Ruttmann's experiments with moving form, particularly with the rapidly expanding and contracting squares in *Rhythmus 21* (1921), although his work also includes some quite intricate detail, possibly influenced by his association with Eggeling. Richter also worked in black and white for his early experiments, although he originally intended *Rhythmus 23* to be coloured. He was discouraged from this because of technical difficulties in the colouring process.[96]

Oskar Fischinger and Visual Music

Oskar Fischinger was born in 1900 in Gelnhausen, a small town near Frankfurt in Hessen, Germany. Oskar was the fourth child of six, and while his elder siblings were expected to work within the family pharmacy business, Oskar and his younger brothers were encouraged to develop their interests and schooling, although he was still expected to take on a traditional apprenticeship at the age of 14.

His initial interest was in music, as he learned how to play the violin before taking on his apprenticeship in an organ builders in 1914. However, after the owners of the business were called up in the First World War draft, Fischinger turned to his other great interest, graphic art, securing a place in an architect's office. The Fischinger family had by this time sold their pharmacy and moved into the restaurant business, which had done badly, and they relocated to Frankfurt in 1916. Fischinger attended a local college, and started work in a manufacturing company, gaining skills in technical drawing and engineering, eventually gaining his certificate as an engineer

[96] Vergo, P. (2010). *The Music of Painting.* Phaidon Press, London. p276.

in 1922.[97]

The move to Frankfurt offered Fischinger access to a much more interesting artistic scene. He joined the 'Friends of Literature Club', which took an interest in the recent innovations in the theatre and the new art-forms of photography and cinema. A key contact was Dr. Bernhard Diebold [98], an influential theatre critic and a strong proponent of Wagner's concept of *Gesamtkunstwerk.* Fischinger had a copy of a 1916 article [99] where Diebold called for this fusion of work to develop into *Bildmusikers* (visual musicians) – a likely source for Fischinger's own term for his work, *visual music*. Diebold also took Fischinger to see Walter Ruttmann's *Lichtspiel Opus No. 1* in 1921, and introduced him to Ruttmann at the end of the showing, resulting in a continuing correspondence between the two. Fischinger was also aware of Eggeling's work – in one of his lectures, Moholy-Nagy explicitly noted the influence Eggeling had on Fischinger - "Fischinger's abstracts, partly based on the works of Viking Eggeling […] attempt to formulate their own idiom of movement optically, and to come up with a new dimension of filmic expression."[100] Ruttmann's films were shown along with Viktor Eggeling's in 1921 on at least one occasion,[101] and as Eggeling was a close contemporary of Hans Richter, it is likely that Fischinger, with his interest in the field, was also able to see early performances of Richter's abstract films *Rhythmus 21* (1921), *Rhythmus 23* (1923), and *Rhythmus 25* (1925).

With Ruttmann's encouragement, Fischinger combined his skills in music and graphical arts, and started making his own films. As he felt that Ruttmann was already established in creating drawn and painted film, he decided to adopt a different approach and a technique unique to him.[102] His engineering background led

---

97 Moritz, W. (2004). *Optical Poetry: The Life and Works of Oskar Fischinger.* John Libbey Publishing, Eastleigh, UK. pp1-3

98 *Ibid.* p3

99 *Ibid.* p3 ; original source (unaccessed) Diebold, B: *Expressionismus und Kino.* article*:* "Neue Zuricher Zeitung" 137, Nr. I459, I5. 9. I916

100 Passuth, K. (1985). *Moholy-Nagy.* Thames & Hudson, London. p321; original source (unaccessed) Moholy-Nagy, L. 'Új filmkísérletek' in *Korunk*, 1933. No. 3. pp231-237.

101 Vergo, P. (2010). *The Music of Painting: Music, Modernism and the Visual Arts from the Romantics to John Cage.* Phaidon Press, London. p271. Eggeling's film was likely the now-lost *Horizontal-Vertical Orchestra* - given the date, Ruttmann's film would have been *Opus 1*.

102 *Ibid.* p7

him to make an apparatus that sliced and photographed cross-sections of melted wax cylinders, giving an organic animation of shifting forms. With some initial success, Fischinger wrote to Ruttmann to see if the technology would be of interest to him. Ruttmann provided some initial funding but found the machine too awkward when he tried to use it. Despite the creative films he achieved with the wax machine, Fischinger could not make the machine a long-term success.

In 1922, following his correspondence with Ruttmann, he moved to Munich and decided to establish himself as a film-maker there. One of the key figures in modern art, Kandinsky, had lived in Munich from 1908 to 1914, where he published his work *Concerning the Spiritual in Art* in 1912. Kandinsky drew comparisons between music and art in his writing, and Fischinger would have been introduced to this both in Frankfurt and amongst his peer group in Munich. Fischinger gained some commissions in Munich, and created his own experimental work. Unfortunately his debts mounted steadily, and he eventually decided to escape his creditors and walked to Berlin in 1927.

The move to Berlin proved to be a successful one. The Bauhaus had been encouraging the growth of new artistic endeavours from its formation in 1922.[103] Along with this encouragement of applied arts in Germany, the Hollywood machine had started distributing films, exploiting a fertile market following the disastrous effect the First World War had on the European cinema industry. Counteracting American popularity, the directors making films in the new Weimar Republic sought to establish an intellectually superior cinema - “a highbrow alternative to Hollywood cinemas”[104] where expressionism and style were prime driving forces. In this atmosphere of experimentation, Fischinger gained commissions for special effects in film work, such as rocket sequences and a view of the earth from the moon in Fritz Lang’s *Frau im Mond* (1929), along with critical praise for his experimental animation work. Fischinger continued to create abstract images to accompany music,

[103] Read, H. (1974). *A Concise History of Modern Painting.* Thames and Hudson, London. p173
[104] Kaes, A. Jay, M. & Dimendberg, E. (eds.) (1994). *The Weimar Republic Sourcebook.* University of California Press, Berkeley. p617.

as he felt that "few questioned the abstraction of sound […] Fischinger felt that this could also be achieved in purely visual terms. Shapes and forms need not represent anything at all and would provoke a variety of responses in the viewers who received them".[105]

In 1929, he started a series of black-and-white studies in charcoal on paper, animations made to match short compositions. Although these Studies were non-commercial, Fischinger maintained a good relationship with the Electrola Company, the source of the musical recordings he used. Fischinger advertised the fact at the end of each film, making him an early proponent of the music video as advertisement. Having reached *Studie Nr. 12* (1932), Fischinger became intrigued by the representation of sound that he could see on the optical soundtracks to films. He decided to draw similar 'ornamentations' to see how they would translate into sound and music when played back as a soundtrack. He received a great deal of publicity when he promoted his idea[106], with Moholy-Nagy stating in a 1933 lecture "...we can look forward to even more surprising experiments, such as Fischinger's experiments with sound ornamentation".[107]

Unfortunately, despite the publicity, Fischinger could not secure funding to develop the idea further, although his biographer Moritz notes that Fischinger's conversations with John Cage and Edgard Varèse in the late 1930s about his theories may have influenced their compositions.[108] Norman McLaren's short animation *Dots* (1940)[109] may have also been influenced by these ideas, as the sound is created by painting directly onto the optical soundtrack. Fischinger was to revisit this territory with his Synthetic Sound machine, which he built c. 1948, where he created musical 'scores' - glass plates with drawn shapes – which could be placed at various positions within

---

105 Wells, P. (1998). *Understanding Animation*. Routledge, London. p29
106 Moritz, W. (2004). *Optical Poetry: The Life and Works of Oskar Fischinger.* John Libbey Publishing, Eastleigh, UK. p44.
107 Passuth, K. (1985). *Moholy-Nagy*. Thames & Hudson, London. p323; original source (unaccessed) Moholy-Nagy, L. 'Új filmkísérletek' in *Korunk*, 1933. No. 3. pp231-237.
108 Moritz, W. (2004). *Optical Poetry: The Life and Works of Oskar Fischinger*. John Libbey Publishing, Eastleigh, UK. p44.
109 McLaren, N. (1940). Short film "*Dots*" - http://www.youtube.com/watch?v=E3-vsKwQ0Cg [accessed 17/08/2008]

the machine. The distance between the glass plate and the camera controlled the pitch or frequency of the generated noise, so Fischinger could use the same plate to create the same sound at different frequencies, rather than having to recalculate and redraw the shapes to adapt to a different pitch as he went along. While some recordings exist of sounds and music generated using the machine, a patent was never filed, and it is unknown if Fischinger intended to make a more complex animation using the mechanism[110].

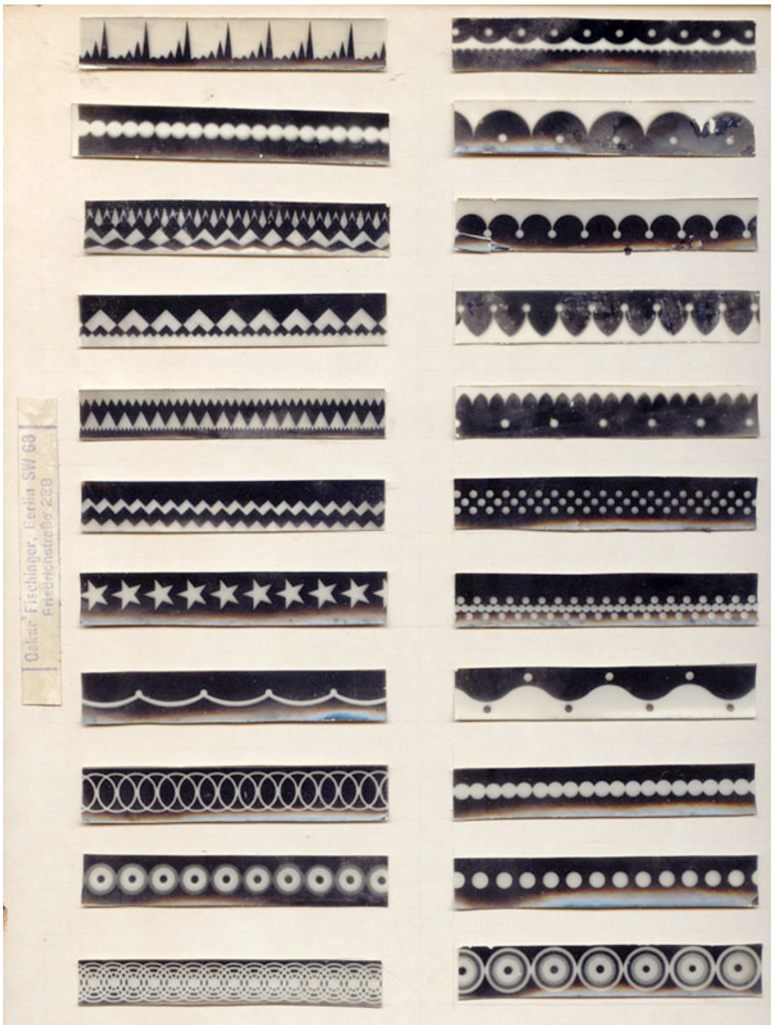


Figure 6: Detail of Demonstration Card with various Ornaments by Fischinger[111]

[110] Keefer, C. & Guldemond, J. (eds.) (2012). *Oskar Fischinger 1900-1967: Experiments in Cinematic Abstraction.* Eye Filmmuseum / Center for Visual Music. pp145-147.

[111] Image sourced from http://www.centerforvisualmusic.org/Fischinger/SoundOrnaments.htm [accessed 28/04/2017]

In 1933, the Nazis took control of the German government, and began to restrict the creation of 'decadent' art. The artists of the 'avant-garde' were increasingly being regarded as subversive and their art subject to governmental censorship. Faced with a lack of distribution due to this censorship, and increasingly finding himself in disapproval with the fascist attitudes of the ruling party, Fischinger took advantage of an offer of work with Paramount in America to emigrate in early 1936.[112] His seven year contract with Paramount was cancelled after 6 months, due to missed deadlines, but one of the contacts Fischinger made at Paramount was with the conductor Leopold Stokowski, with whom he discussed a collaboration, where Stokowski would supply the music and Fischinger the visuals. In 1937, Stokowski would start a collaboration with Disney on an animated version of *The Sorceror's Apprentice*, which later would become part of a larger 'Concert Feature' – *Fantasia*.

Fischinger had already created an abstract animation using Dukas's *The Sorceror's Apprentice* – his *Studie Nr. 8* (1931), and had publicly announced his plans for a feature-length visual music animation at the start of 1935. Fischinger suspected Stokowski of 'selling' his idea to Disney without due acknowledgement of its origin. However, faced with deportation or working for Walt Disney, Fischinger decided to put all his efforts into working on *Fantasia*. Initially, his work went well, with his earlier shorts shown as inspirational examples to the Disney employees during office breaks. However, he found his ideas were simplified, his colours were muted, and Fischinger felt himself marginalised, writing to a friend that "I worked on this film for nine months; then […] I was demoted to an entirely different department". Fischinger left Disney and *Fantasia* at the end of October 1939, requesting that his name be taken off the credits, though his influence is still evident in some of the film sequences.[113]

Despite gaining credit from luminaries such as John Cage, who acknowledged

[112] Moritz, W. (2004). *Optical Poetry: The Life and Works of Oskar Fischinger*. John Libbey Publishing, Eastleigh, UK. pp62-63.

[113] *Ibid.* pp37, 55, 75, 83-87.

Fischinger's influence on his own ideas after they met in 1937 [114], Fischinger struggled to fund his continued work. He managed to raise the finances for his *Motion Painting No. 1* (1947) which matches Bach's *Brandenburg Concerto No.3* with patterns of oil paint on plexiglass, endlessly drawn over, developing with depth and intricacy, but this was the last substantial film Fischinger completed. His principal funding source post-Disney was the Baroness Hilla Rebay, curator of the Solomon R. Guggenheim Foundation and the Museum of Non-Objective Painting. The relationship was very exploitative of Fischinger's work, with the Baroness offering more destructive criticism than encouragement. Her disappointment with *Motion Painting No. 1* brought the relationship to an end.[115]

The sour relationship with Rebay may have contributed to Fischinger's continued residence in Hollywood. He had been urged in 1939 by the abstract film-maker Mary Ellen Bute to move to the East Coast.[116] In New York, Fischinger would have been closer to Mondrian, whose work he admired,[117] and through his acquaintance with John Cage, he might have found support in the 'Artists Club' of avant-garde painters in Lower Manhattan.[118] However, given that Rebay was based in New York, and as the animation and movie industries were gaining primacy in Hollywood, it is understandable that Fischinger decided to stay on the West Coast, especially as he had already invested in settling his family there. Fischinger continued to struggle financially, supported partly by a few television commissions for advertisements, but mainly from the money earned by his wife in costume design and various domestic jobs. His artistic endeavours moved mainly to canvas, which did start to provide some income, and he carried on painting until his death from a heart attack in 1967.

Despite his later lack of success, Fischinger's films proved influential within the genre. The Scottish-born film-maker Norman McLaren, Oscar and BAFTA award winner and a primary influence in Canadian cinema, noted of *Studie Nr. 7* that "[i]t is

---

[114] *Ibid.* p78
[115] *Ibid.* p133.
[116] *Ibid.* pp80-81.
[117] *Ibid.* p93.
[118] Vergo, P. (2010). *The Music of Painting*. Phaidon Press, London p335.

difficult to describe adequately the impact it had on me: I was thrilled and euphoric by the film's fluent kinesthesia, which so potently portrayed the movement and spirit of the music. The experience made an indelible impression on me, excited a yearning in me, and was to have a profound, long-lasting influence on many of my films".[119] Other film-makers inspired by Fischinger included the Whitney brothers who designed the star gate sequence in Kubrick's *2001AD: A Space Odyssey* (1968).

Animation in the USA

Similarly to Cohl's move from cartoonist into animation film-maker, the American Winsor McCay was inspired by the new medium to create an animated version of his popular *Little Nemo in Slumberland* comic strip. In 1911, McCay released a 10.5 minute film, now known as *Little Nemo*[120]. In the earlier part of the film, McCay is shown being ridiculed by his cronies for proposing the idea that his comic strip creations could be animated – after various sequences showing the artist at work in his workshop, we finish up at a sequence, starting at approximately 8'20", showing his famous characters being brought to life. Note that the film dates from 1911, and some of the characterisations here suffer from attitudes to race at the time. Given the popularity of Little Nemo in print the film had an immediate appeal to the public and was immediately successful. McCay followed this animation with other well received works, particularly *Gertie the Dinosaur* (1914) and his animated recreation of *The Sinking of the Lusitania* (1918). This animation is particularly important as it marks the first real use of animation for re-enactment of key historic events, even as a propaganda tool, marking the outrage felt at the sinking of a passenger ship by a German submarine during the First World War.

McCay continued to work in both animation and illustration, but was disappointed with the commercialisation of animation over the next few years. In a dinner for New

[119] *Ibid.* p166
[120] Wangjacup (2006) *Winsor McCay – 1911 – Little Nemo*. [video online] Available at http://www.youtube.com/watch?v=kcSp2ej2S00 [Accessed 7 December 2014].

York's animation community in the 1920's, he admonished the guests - "Animation should be an art, that is how I conceived it. But as I see what you fellows have done with it is make it into a trade... not an art, but a trade... bad luck".[121]

It is tempting to link McCay's dismay at the movement of animation from art to trade to the rise of Disney and his domination of the animation market, but in truth most of the animators in the USA in this period aimed at making money from their work. Given the technology was rapidly changing and there were ever increasing demands for producing regular films to satisfy a voracious public appetite for cartoons, it is understandable that in the United States, the business side of animation came to dominate the art-form, much as American film-making has also been dominated by the profit margin from its early days.

American Animation and the Sweatshop Model

In both Europe and America, the first animations were in general made on an auteur basis by the animator in total – the same person would develop the narrative of the animation, create the characters, draw the individual frames, and then film the entirety to create the final work. While the auteur remained a feature of European film-making into the inter-war period and beyond, as the industry developed in the USA, the production of animation moved from the individual to a studio format, where several animators were employed to work on the animation at the same time. This allowed the works to be created much more quickly, meeting the delivery schedules the studios would agree with the film distribution companies. Within the studio, the animators would be assigned specific tasks, segmenting the production of the animation into separate tasks. In many ways, this was a general movement across industry, following on from Ford's division of labour in his automotive factories.

For most animations then, a lead animator would be assigned, who would have the

[121] Bendazzi, G. (1994). *Cartoons: One hundred years of cinema animation*. John Libbey & Company Ltd. London, UK. p18.

overall creative vision and editorial control. In most cases, this would be the head of the studio, such as Walt Disney. A set of sketches or a storyboard would be produced, as well as a set of reference drawings of the main characters. The story would be split into sections, and a key animator would create the key frames showing the start and end of the main actions in the section. Other animators would then draw the 'in-between' frames leading from the start to the end frame. The drawn frames would then be passed on to colourists to fill in with block colour, again based on a reference chart provided by the lead animator. Finally, the entire work would be photographed and assembled as a full film.

This model of animation is generally referred to as the 'sweatshop' model – this should not be taken as an equivalence to any exploitative sweatshop industries, although there was quite a considerable difference in wages and respect offered to those 'lower' in the pecking order. The colourists were often women, employed at a much lower rate than the generally male animators, but this was not unusual for the period. "Walt Disney would sketch his ideas to his animators who would plan and draw the animation cels, which would then be filled in by female colorists."[122]

Disney was not the only studio head to follow this model – due to the short deadlines imposed by the distributors, all the American animation studios perforce had to adopt this approach. For each studio, a side effect of this approach meant that the ideological or philosophical ideas of the lead animator became central to the entire output. The earlier independent creative vision of the first animators, who could establish their own artistic and directorial styles, were subsumed into a studio vision. Due to the success of Disney's output, his philosophical approach to animation became the standard to which American and then Western animation aspired – in an industry driven by profit, successful box office returns became the key determination of which animation approaches were 'good'. While Disney's philosophical approach and animation techniques are dealt with later in this chapter, one of the successes of early animation, the recurring character, became a key element of his studio's

[122] Tracy, J. F. (1999). *Whistle While You Work: The Disney Company and the Global Division of Labor*. In *Journal of Communication Inquiry*, October 1999 Vol.23 no.4 pp374-389. p375.

dominance.

The Recurring Character

Having hooked the audience with the short film with recognisable characters, the producers saw the benefit in making the next step – a series of short films starring the same character or characters. As the audience already know what to expect from the character, there is no need to set up any back story. They can just settle in and enjoy the story as it is. Some of the first characters established in this way were *Ko-ko the Clown*, who appeared *Out of the Inkwell* in cartoons by the Fleischer Brothers, *Colonel Heeza Liar* by Walter Lantz, later the creator of *Woody Woodpecker*, and the popular *Felix the Cat* – created by Otto Messmer for Pat Sullivan's studios.

Particularly with Ko-ko and Felix, there was an interest in how to manipulate the drawn medium. Ko-ko would draw an object for his own use, while Felix would both draw objects but also use parts of his own body to affect the scenery around him. Ko-ko made his début in 1916, and marked an early use of rotoscoping, using filmed footage of a real actor to provide a basis from which the character's movements could be drawn, although later animations in the series dispensed with the rotoscope[123]. Ko-ko marked an interaction between the free-form world of animation and the more rigid 'real' world – often with the hands of the animator, who exerts a God-like control of the animation when necessary. In these films, there is no limit on how the animated objects on screen can interact. In *Ko-ko the Kop* (1927), his sidekick, Fitz the dog, needs to escape. He lifts up a rock to expose a dark hole behind it, and escapes down the tunnel thus revealed[124]. The physics of the real world are not a hindrance to the animator's imagination.

123 Cavalier, S. (2011). *The World History of Animation*. Aurum Press Ltd. p66.
124 Barrier, J. M. (1999). *Hollywood Cartoons: American Animation in its Golden Age*. Oxford University Press. p27.

Figure 7: Ko-ko the Clown[125]

Felix made his first official appearance in *The Adventures of Felix* in 1919, although precursors to this very recognisable black and white cat had appeared in some earlier productions. While Felix showed some of the expected preoccupations of a cat – often attempting to acquire fish from a local shop, for example – he was anthropomorphised, in the rapidly established fashion of the time. His preoccupied pacing up and down, hands linked behind his back, as he pondered a new scheme became an instant trademark of the character. His world was a surreal development of our own. While children played ball in the street, and shop-owners threw Felix out when he tried to sneak in and steal food from their shops, the laws of physics were often played with for comic and narrative effect. Felix's tail, in particular, would often be detached and changed into whatever item was necessary at the time – an umbrella, a shovel - this world was malleable, flexible and enticing.

Over the course of the next 10 years, Felix would create a ladder out of question marks that appeared over his head as he pondered how to access an attic jail cell (*Felix Saves The Day*, 1922); detach his tail and use it as a walking cane in a Charlie Chaplin impersonation (*Felix in Hollywood*, 1923); and use question marks again to create sledge runners (*Felix Gets The Gun*, 1924). By 1927, Felix was very well

[125] Inkwellimagesink (2014). http://www.inkwellimagesink.com/pages/articles/findingKoKo.shtml [Accessed 7/12/2014].

established, and his surreal interactions with his surroundings were still ongoing. In *Pedigreedy*[126], released in this year, he pulls on his skin like trousers, and when he can't find his tail, a question mark pops up over his head, which he then pulls down to attach to himself as his tail.[127]

In the 1930 film, *Oceantics*[128], Felix wantonly ignores the rules of perspective when he wants to steal a large cheese round from a shop window. He uses his tail to hook the front door of a house in the background, and sticks it onto the front of the shop, opening and reaching in through the tiny door to reach the cheese. He then uses a knife to cut a spiral through the cheese, opening up a long roll with cheese 'holes' through it, which he then feeds into a player piano. The sheer freedom and opportunity that Messmer had to make use of the entire world around Felix is very well illustrated in this sequence.

Figure 8: Screen captures from Felix the Cat *Oceantics* (1930)

Felix had a loyal following, and *Mickey Mouse* still had to make his début in 1928. But for some reason, Pat Sullivan, though just as ambitious as Disney, would not be convinced that sound was important, or that he needed to invest in new film equipment, and Sullivan's contract for Educational Films was not renewed after the 1928-29 season, leaving Felix out in the cold.

---

126 Tvdays (2008). Felix the Cat *Pedigreedy*. [video online] Available at http://www.youtube.com/watch?v=XjjMuqvy7yw [Accessed 7 December 2014].

127 Canemaker, J. (1996). *Felix: The Twisted Tale of the World's Most Famous Cat.* Da Capo Press, New York. p113.

128 Classic Cartoon Matinee. Felix the Cat *Oceantics*. [video online] Available at http://www.youtube.com/watch?v=6jX6Nvs_mlM [Accessed 7 December 2014].

The Domination of Disney

Walt Disney dominated the cartoon world (in the West, at least) from the 1930s onwards. His method of production took him away from front-line animation, but he remained determined to retain the artistic control he had when working as a solitary auteur. He knew the importance of the 'name above the door'. Who now knows that *Steamboat Willie* was mainly the work of Disney's early partner Ub Iwerks? Disney's name was upfront, and that fuelled his success. Disney was a key character in the development of animation and the ongoing promotion of the art-form. His cartoons were popular and provided employment for the burgeoning ranks of young aspiring animators. But his sweatshop model came to dominate the industry methods – even though the Warner Brothers and others carried on with a more anarchic approach to their animation on screen, they had to adopt the sweatshop model in order to compete with Disney.

Disney's own approach to animation was rooted in his conviction that the characters and environment depicted in his works, no matter how fantastical they were in terms of anthropomorphism or setting, still had to operate within the rules of real-world physics. In contrast to the freedom taken by other animators in disregarding gravity, perspective, and cause-and-effect in their works, Disney "insisted on verisimilitude in his characters, contexts and narratives. He wanted animated figures to move like real figures and be informed by a plausible motivation"[129]. This approach to reconstructing the real world within animation was undeniably successful, married as it was to Disney's technological improvements in sound and colour, which brought a regular injection of 'spectacle' into his output, enticing viewers even further into his artistic vision. By the time his first feature film, *Snow White and the Seven Dwarfs* was released in 1937, Disney's model of real-life equivalence had become established as the industry norm - "the animated film had reached maturity, but in doing so had established Disney as synonymous with 'animation'. This has led to animation being understood in a limited way."[130] The dominance of Disney in

[129] Wells, P. (1998). *Understanding Animation.* London: Routledge. p23.
[130] *Ibid.* p24.

mainstream animation meant that the more experimental or abstract approaches of animators based in other traditions became less valued and effectively ignored within the wider industry.

In Europe and Asia, the vicissitudes of the Second World War disrupted their burgeoning film industries, and animators were still working mainly as auteurs. The rise of the Nazi party in Germany led to many of the Dadaist and avant-garde artists making a move to America, with Hans Richter, Max Ernst, Man Ray and Marcel Duchamp all finding a new home in New York[131]. As noted above, Fischinger was employed by Walt Disney to work on *Fantasia* (1940), specifically in the Bach *Toccata and Fugue* sequence, but left when he felt his artistic input was being diminished by editorial decisions.

Credit must go to Disney for creating *Fantasia*, and for recognising the talents of gifted animators like Fischinger, but his insistence on microscopic control kept that talent in a straitjacket. Audience reaction to *Fantasia* was mixed – some thought it was too 'high-brow', and it failed to make a profit. But who knows if it might have been more accessible, not less, if animators such as Fischinger had been given a freer hand. The Bach *Toccata and Fugue in D minor* section from *Fantasia*[132] shows a glimpse of Fischinger's style, with clear correlations to his own animations. Sadly, his original concepts were adapted after his departure from the project, and given to other animators who were more amenable to Disney's methods, so this is only a small hint at what could have been.

'Squash-n-Stretch'

One of the defining elements of the Disney 'style' is his early adoption of 'squash-n-stretch', a term used to describe how cartoon characters are drawn in motion. In this

[131] Grey, C. (ed.) (1971). *Hans Richter by Hans Richter*. Thames and Hudson Limited, London p48.
[132] Thereisonly0neNM (2011). *Fantasia {The Abstract Part}.* [video online] Available at http://www.youtube.com/watch?v=zt9iy5THnbk [Accessed 7 December 2014].

style, the movements may be extreme or deformed, but they still conform to real-world physics – fat bodies wobble in anticipation of movement, limbs stretch and then compress in the same way as limbs do in real life. Walt Disney wanted his animation to be convincing, and the use of this sort of animation grounds the characters in a recognisable world with similar physics to our own. His animations impressed because of the characterisation brought to the animation 'actors' – *Three Little Pigs* (1933) marked a breakthrough in animated characterisation, with each pig having a discernible, separate identity despite superficially all looking alike.

Disney's adoption of the squash-n-stretch principle "necessarily over compresses and elongates character movement to give it an over-determined and often comic style, but it remains that moving figures within the Disney canon correspond more directly to 'realistic' movement than work informed by other approaches."[133] Along with his early adoption of sound and superior marketing efforts, Disney began to dominate the American animation market, and his rivals soon adopted his approach both for the physical creation of the animation and the physics of the animated world.

Alternatives to Disney?

The Warner Brothers set up shop at around the same time as Disney, and continued with the more anarchic animation seen in Felix and his contemporaries. However, in order to compete, they also adopted the sweatshop model, which again limited the creativity and expressiveness of their animators. Their animated worlds also made use of the physics common to Disney, although with a nod towards more freedom – Wile E. Coyote may run off a cliff edge and take some moments to notice what he's done, but gravity always wins in the end. Another studio, UPA (United Productions of America), was an offshoot of disaffected animators who wanted to create more experimental and idiosyncratic work. In some cases, such as *Gerald McBoing-Boing: Jolly Frolics*[134] (1950), the results were very well received, but others were much

[133] Wells, P. (1998). *Understanding Animation.* London: Routledge. p27.
[134] All Classic Video (2012). *Gerald McBoing-Boing: Jolly Frolics* (1950) – Classic Short Film.

less successful.

In Europe, there were several influential animators who established a more experimental approach to animation – as well as Oskar Fischinger, the 1910s to the 1930s saw abstract animation by Walter Ruttman, Hans Richter, Viktor Eggeling and Lottie Reiniger amongst others, mainly based in Germany. The advent of the Second World War and the establishment of the Iron Curtain after the war put a huge block in the development of this sort of animation, and it is only relatively recently that the rich heritage of Eastern Europe animation has become available, and has started to flourish again now that state control of artistic expression has been relaxed. Within this culture, the work by the Czech animator Jan Svankmajer from the 1960s onwards stands out for its experimentalism and surrealism.

Experimental auteur animators still existed in the USA, but their efforts were vastly overshadowed by the commercial animation industry. However, the work of John and James Whitney in the 1960s gained more popular appeal, and inspired the 'star-gate' sequence in Stanley Kubrick's *2001AD: A Space Odyssey* (1968). At about the same time, Jordan Belson was also developing abstract animation, which he staged with specially commissioned music in planetariums across the USA in his series of *Vortex* concerts.

In Japan, animé rose out of the aftermath of the collapse of their studios at the end of the war, mirroring the early inspiration of US animation from comic strips, in this case from manga. Studio Ghibli was particularly successful, and does allow fantastical transformations to take place in its films, although to me, it often feels as if it's halfway between the Disney 'orthodox' standard and a more free flowing experimental style. Interestingly, Walt Disney Studios has international distribution rights to much of Studio Ghibli's earlier output.

The 1990s saw an increase in experimental animation in the UK, arising from an

---

[video online] Available at http://www.youtube.com/watch?v=rDCwg0Xcji4 [Accessed 7 December 2014] initiative run by Channel 4 and the National Lottery. Unfortunately, this initiative has run its course, and Channel 4 is no longer a champion of new award-winning experimental work. Public funds for non-commercial animation are very restricted, which limits the opportunities for new voices to develop.

Animation and its correspondences to Electroacoustic Music

As noted in the introduction to this chapter, there are parallels between the composition of electroacoustic music and the more experimental approaches to animation. Like animation, electroacoustic music has been made possible because of technological advances. And like abstract or art animation, electroacoustic music is perceived as a specialist genre with limited appeal.

The animator has complete control over the visual environment. They can choose where any object appears on screen – how long its appearance lasts – what movements it carries out – what colours are featured – what interactions there are between foreground and background elements. This is far more control over the visual environment than other film genres, such as documentaries, and mirrors the full control we have as composers in this area. This full control over both the audio and visual environment is essential for the audiovisual composer to successfully create audiovisual objects and assemble them into an immersive work.

It is also worth noting that animation has its own categories contained within it, such as the special effects used in live action films – stop motion animation – hand-drawn cel animation – CGI 3D animation. This can be compared to electroacoustic composition, where we have composers who prefer to work with synthesized sounds, with recorded material, a combination of both, or who combine ‘tape music’ with live instruments.

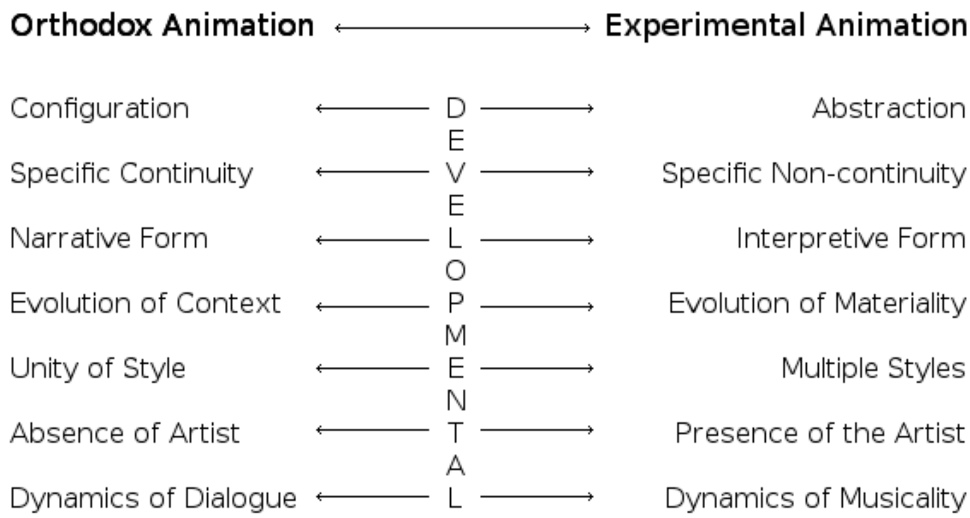


Figure 9: Comparison of Orthodox and Experimental Animation styles, reproduced from *Understanding Animation*[135]

The figure above compares two ‘extremes’ of animation styles, defined as Orthodox and Experimental by Paul Wells. By orthodox animation, Wells is referring to the standard American model, epitomised by Disney. In many ways, this attempts to replicate live-action, even if there are elements that can only happen in animation (even if we’re considering ‘live-action’ films such as *Lord of the Rings*, the fantastical elements are mainly achieved through CGI animation). Even if we’re looking at *Daffy Duck on Mars,* at heart natural physics are still observed, if exaggerated, and the animation is driven by dialogue and narrative. In comparison, in Experimental animation, narrative is not important – it can be there, but may not be the driving factor. Often dialogue is unnecessary. Rather than subsuming the artist into the homogeneity of the studio, the vision of the animator is the driving force.

If we compare this to the approaches taken by the electroacoustic composer, we can see several parallels. While abstraction may be thought of as being inimical to musique concrète, in this case it refers more to the freedom to regard any element as

[135] Wells, P. (1998). *Understanding Animation.* London: Routledge. p36.

valid in importance to the work – just as the electroacoustic composer is not restricted to notes and instruments, the experimental animator is not restricted to characters and narrative, although they can make use of these if they want to. For both the animator and the electroacoustic composer, the constituent elements of their work can range in identifiability, from a recognisable church bell on one hand to a highly processed granular textural sound background, which could derive from any one of a multitude of sources; or an identifiably anthropomorphic cartoon character such as Mickey Mouse through to the swirling textural patterns in Bret Battey's work[136] such as *Autarkeia Aggregatum* (2005) or *Sinus Aestum* (2009), where again the source material is not immediately identifiable and could derive from a wide variety of sources. In creating abstract or experimental work in the audiovisual field, all forms on this scale of identifiability can be used – a cartoon cat can form out of visual chaos and then dissolve into another form – there is an extensive freedom for the animator or audiovisual composer to experiment and realise their own aesthetic using whatever audiovisual objects fit their requirements.

Both electroacoustic composition and animation are time-based but have to consider the 'space' being created within the piece. In his key work on space and the forms it can take in electroacoustic music, *Space-form and the acousmatic image137*, Smalley describes many aspects of space that impact upon the electroacoustic composition. Of particular interest here is his discussion of perspectival space, where "Relative position, size and movement produce a sense of scale in the relations among forms, and a sense of the scale of the articulated space as a whole." During their experience of a sonic composition, the audience form an impression gained over time of a 'landscape' of sound objects, which are positioned in the foreground or background, and move around each other and in relation to the audience. Similarly, perspective is created within the visual field by the animator, with the exploration of the relationships between visual objects over time placing them in perspectival relation to each other – in the foreground, moving towards, behind or away from

[136] Battey, B. Online gallery of works available at http://www.mti.dmu.ac.uk/~bbattey/Gallery/index.html [Accessed 08/10/2015].
[137] Smalley, D. (2007). *Space-form and the acousmatic image.* In *Organised Sound* 12 (01), ed. Landy, L. pp35-38.

each other. By combining the perspective aspects and spatial motions of a sound object and visual object, fusing them into an audiovisual object that has sympathetic spatial movements and characteristics, a much more 'believable' form is created.

I find Wells' definition of Experimental Animation as having the dynamics of musicality rather than dialogue interesting – an acknowledgement by a leading voice in the field that in many experimental animations there is a feeling of a composition, an underlying musicality that drives the piece rather than the narrative dialogue we are familiar with from live action films.

Audiovisual Composition – putting the two together

William Moritz, Oskar Fischinger's biographer, regarded 'true' animation as being only that which was non-linear and/or non-objective i.e. lacking discernible characters or a driving narrative.[138] I would contend that his definition is a little restrictive, given that this would create an extremely small field in which a work could be regarded as a true animation. In addition, given that all animation is time-based, with a discernible start, middle, and end, there is in effect a narrative structure for all animation, as even a transition from one unidentifiable textural state to another similar but recognisably different state requires a development within the work and thus a narrative, However, I would put forward his ideal that the creativity and skills needed for this type of experimental animation are demanding, and in many ways mirror the skills and creativity demanded of the electroacoustic composer. As with electroacoustic composition, the animator has to find the right combination of elements from an infinity of possibilities to add to that starting blank canvas – as with an electroacoustic composer, the animator has to marry their creativity with hard-earned skills.

[138] Moritz, W. (1988). *Some Observations on Non-Objective and Non-linear Animation*. In Canemaker, J. ed. (1998), *Storytelling in Animation: The Art of the Animated Image Vol. 2*; Los Angeles: AFI. Ch2.

Within the field of film sound design, the Hollywood sound designer Randy Thom has been responsible for many outstanding sound tracks, especially in the field of animation (*The Incredibles*, *How to Train Your Dragon*). He takes the view that sound and vision have to be considered in partnership in order to breathe life and vitality into the cinematic experience. Film-making is a collaborative process, where the sound is at least as important as the vision. He has called for the sound designer to be involved at a much earlier stage, and to be viewed as equally important as the director of photography. "I mean new creative ideas, storytelling ideas. New ways to collaborate, new ways to make visual and aural information enrich each other and become indistinguishable as an experience.

"Sound is NOT there to "help the visuals." That's kindergarten film making. Anyone who says that "film is a visual medium" is being foolish and naive. Sound, when given half a chance, is no less important to the audience's experience than the pictures."[139]

Conclusions

As the audiovisual object is created from a fusion of a sound object and a visual object, the audiovisual composer benefits enormously from the almost infinite selection from which he or she can choose. As has been discussed above, there is a rich heritage of both electroacoustic composition and experimental animation built up over the past century and more which the audiovisual composer can draw on for inspiration.

The development of both animation and electroacoustic composition has been driven by developments in technology, particularly in the development of electrical and computing technology. The auteur animator and the electroacoustic composer have both played a role in developing the technology used in their work – both pushing at

[139] FilmSound.org (2009). *What is the basic philosophy of sound design?* [online] Available at http://www.filmsound.org/QA/sound-design-philosophy.htm [Accessed 7 December 2014].

the boundaries of what can be achieved within the platforms that have already been developed for the artform, and developing new technology as necessary to achieve the results they require for their own aesthetic ambitions.

Both electroacoustic composition and experimental, abstract animation exist in parallel with a much more popular commercial form making use of much the same technology and basic creativity – popular electric instrument based music and commercial animation respectively. In the case of electroacoustic composition, it has benefited from its continued existence as a fairly elitist musical form, protected by its position within academia. Experimental animation has been less successful in maintaining a foothold in its field, though art schools and the occasional government funded incentive scheme has allowed it to maintain some viability.

Electroacoustic composition and animation are both experiences that develop over time, but which incorporate a structure and perspective within the space created within the work. This may be a very amorphous structure, or the narrative behind the work as it progresses through time may not be very evident or driving, but the audience will still form an impression of how objects move and interact with each other over the course of their experience. The mechanisms by which the composer or animator can create the perspective of depth or motion in their work can be used to reinforce that structure and motion in the audiovisual objects created by the audiovisual composer.

It is very profitable for the audiovisual composer to study the compositions and animations that have been created over the history of both art-forms. The development of 'organised structure' in electroacoustic music in the way the sound objects are created and combined gives a theoretical basis for how audiovisual objects can also be created and combined, making use of Schaeffer's theories of the sound object and reduced listening as the basis for defining the audiovisual object, as explored in Chapter 2. The freedom of expression and style in experimental animation, particularly as animators were developing their own language of animation, provides inspiration in how the visual component of the audiovisual object need not be restricted to specific guidelines of shape, physics or cultural influence, but can instead be free to make use of any aspect of visual perception the audiovisual composer wishes to explore.

Chapter 4

The Performance Space and Audiovisual Immersion

Introduction

The intent of the audiovisual composer to create an immersive experience for the audience can not be fully realised without considering the environment within which the audiovisual work will be performed. To a large extent, the effectiveness of the compositional techniques used are hostage to the venue in which the work is performed. Ideally, the performance venue should encourage a "willing suspension of disbelief"[140]. While this phrase was originally coined by Coleridge to refer to the writer's interaction with his readers, it applies equally to the relationship between the creator(s) of an audiovisual work and their audience. The physical space in which the performance takes place influences the perception of the audiovisual environment. From my experience, my first audiovisual works were performed in an electroacoustic concert format, which allowed a great deal of influence over the sound reproduction, but was distinctly lacking in immersiveness for the visual component. Later performances have been in cinemas, which again have some advantages and disadvantages. This chapter explores the development of the audiovisual performance space, and the effect each type of space can have on the immersiveness of the performance and the ability of the audience to lose themselves within the composition.

One of the first major attempts to create an immersive audiovisual performance space was Richard Wagner's opera house in Bayreuth, designed to draw the audience into the music and spectacle of his works. His theatrical innovations were to become a standard across the western world, inspiring further design exploration by multi-discipline artists at the Bauhaus. The pioneering work by architects such as Walter Gropius, while not fully realised at the time, point to the later development of curved screens and surround sound in cinemas as well as hemispheric projection and multi-speaker sound in planetariums, both venues in which audiovisual works are often performed. These venues, along with the electroacoustic concert venue format, allow

[140] Coleridge, S.T. (1817). *Biographica Literaria.* Available on Project Gutenberg http://www.gutenberg.org/files/6081/6081-h/6081-h.htm [accessed 24/01/2015]. Chapter XIV Para 2.

a certain amount of audience immersion in the sound and image being played back, but each has its own advantages and limitations.

By exploring the development of performance venues from Wagner's innovative ideas through to the modern IMAX cinematic format, we can identify how each type of venue will affect the audience immersion within the audiovisual environment, and how an audiovisual composition will be received. In addition to these larger public venues, the development of virtual reality, the decrease in technological costs and the rise of domestically available VR headsets also raises the possibility of creating an individually accessible audiovisual experience, also examined here with regard to how the audiovisual composer can approach and utilise this new technology.

Richard Wagner and *Gesamtkunstwerk*

In terms of a theatrical spectacle, the urge to create a space within which the audience could lose itself and feel part of the overall experience predates the cinema. Richard Wagner was driven to create a space within which the audience would be completely caught up in his operas as they were performed. While living in Zurich between 1849 and 1857, his writings often referred to his idea of *Gesamtkunstwerk* or "total work of art" - the "union of all the disciplines: music, poetry, dance, architecture, sculpture and painting".[141] While there are various political overtones in Wagner's work – the phrase could also be translated as a "unified work of art", and Wagner's ideal was a move away from the divisiveness of modern culture back towards a unity with art, spirit and nature – the concept of *gesamtkunstwerk* has continued to inspire artists and musicians to create work that combines several forms of art, or multimedia works, which can be performed in a sympathetic venue to immerse the audience in a vibrant, holistic artistic experience.

For most of its history, the theatre had been a place to socialise as much as to observe

[141] Millington, B. (1992). *Wagner.* Dent, London. p125.

the performance, but Wagner aimed to concentrate the audience's attention on the absorbing artwork he had created. His opera house in Bayreuth, opened to the public in 1876, was a revolution in concert hall architecture – a somewhat ironic triumph of the use of modern technology, given his preference for a simpler experience than the age of mechanical and technological advances he was born into. The orchestra was in a pit, hidden from the audience – a "mystic gulf" from which the music accompaniment would rise and envelope the listeners. A second proscenium was included, forcing the audience to perceive the actors as both larger and more distant, increasing the feeling they were observing legends unfolding before them. Gaslights on the rear of this second proscenium complemented the more standard footlights, and in an innovation that was to become standard practice in theatres and cinemas across the world, the audience lights were dimmed during the performance, directing the audience to pay attention to the stage rather than to each other.[142]

Wagner's opera house was made possible by the use of modern technology, from the architectural design down to the effects he produced on stage, such as his swimming machines for the Rhinemaidens in *Rheingold*. His grand vision of an all-encompassing experience was realised through the machinery available to him. Smith makes the observation that this experience of theatrical spectacle is essentially linked to the adoption of modern technology as available to Wagner - "mechanical production is the *Gesamtkunstwerk's* inassimilable element, its necessary other"[143]. There is a symbiotic relationship between the development of ever more immersive technological platforms and the creation of immersive art. As the technology develops, our expectations as audience members also grows – the early films showing trains approaching the viewer are now quaint to modern eyes, the special effects in 1950s science fiction films are decried by modern viewers as laughable and unconvincing. Yet, at the time, these experiences were convincing for the audience members. Our quest for the ideal venue or technology for reproducing immersive audiovisual work is never-ending, and each new technological advance offers new challenges and opportunities for the artist.

---

[142] Smith, M. W. (2007). *The Total Work of Art: From Bayreuth to Cyberspace*. Routledge. pp30-31.
[143] *Ibid.* p32.

The Bauhaus and *Gesamtwerk*

Towards the end of the 19th century and at the start of the 20th, the German state decided to support a central arts strategy, particularly in the field of applied arts – which were seen as a vital component in increasing Germany's influence overseas, in the same way that Britain's manufacturing might had come to dominate the entire world. Kaiser Wilhelm II enthusiastically backed the formation of the *Deutscher Werkbund*, which fostered the applied arts, but caused some dissension within the ranks as it placed a major emphasis on mass manufacture. The First World War and the subsequent depression for the ensuing Weimar Republic meant that there was very little investment in the arts, but at the heart of the new state, in Weimar itself, a new art school was established which has proved to have lasting influence. This new school, the *Bauhaus*, was initially supervised by the architect Walter Gropius. His vision for the new school was to unify the artist and craftsman, to foster a collaborative approach across the artistic disciplines. Rather than have the artist subservient to industrial technology, he wanted "the artist to learn to make the machine his servant".[144] This vision has immediate resonance with Wagner's *Gesamtkunstwerk* and its reliance on technology.

Gropius was to expand on the potential shown in Wagner's Bayreuth theatre with a proposal he designed for a *Totaltheater* in 1926, which was never realised. This theatre would have developed on Wagner's hidden orchestra and second proscenium with an adjustable proscenium, stage elevators, electric cycloramas – and film projectors fixed throughout the space, aimed not only at the stage but throughout the space, so that scenery and effects could be projected around and above the audience with the aim of totally immersing them in the performance.[145] This panoramic approach to surrounding the audience in the experience may be seen as a precursor to the planetarium.

The Bauhaus attracted modernists from other fields as well as Gropius's own interest

[144] Scheidig, W. (1967). *Crafts of the Weimar Bauhaus*. Studio Vista Ltd. London. p13.
[145] Smith, M. W. (2007). *The Total Work of Art: From Bayreuth to Cyberspace*. Routledge. p53.

in architecture - artist Paul Klee took up a teaching position in 1921, while Kandinsky joined the organisation in 1922, as part of the Masters' Council. At the Bauhaus, Kandinsky followed up his earlier well-received work. *On the Spiritual in Art* (1912) with his *Point and Line to Plane* (1926), both of which outline his thoughts on colour and form in the visual arts. As a synaesthete, Kandinsky's ideas on how music and art influence each other are a great resource for the audiovisual composer.

Another key appointment was the Hungarian Laszlo Moholy-Nagy in 1923, who promoted constructivism, and his belief that it was possible to have a creative vision across all disciplines, not just the artistic. Moholy-Nagy championed the idea of *Gesamtwerk*, which follows the ideal of *Gesamtkunstwerk* in uniting the art-forms, but in Moholy-Nagy's view, this should extend to all human endeavour.[146] Moholy-Nagy promoted the plastic arts, and spoke on the opportunities of cinematic film – he delivered a lecture across various venues in Germany in 1933 which directly referenced the visual music created by Oskar Fischinger, which he lauded as an attempt "to come up with a new dimension of filmic expression".[147]

Cinema may seem the natural home for audiovisual work, and there are several opportunities for experimentation afforded by the different formats currently in use, but many of the current audiovisual composers are rooted in electroacoustic composition. As my approach also has an essentially electroacoustic composition base, it is worth examining the 'traditional' electroacoustic concert venue and its suitability for audiovisual performances.

[146] Passuth, K. (1985) *Moholy-Nagy*. Thames & Hudson, London. p40.
[147] *Ibid.* p323.

Standard Electroacoustic Concert Formats

For the newcomer to an electroacoustic concert, the first impression is of a space very different from that they may have expected based on an orchestral concert. Instead of a group of instrumentalists sitting on a stage at one end of the room, with a raked seating area for the audience, there are a bewildering number of loudspeakers placed around the room, and the audience are normally seated in rows of seats laid out on the flat – and often there are substantially fewer seats than they would expect, grouped around a sound-desk. Even worse, when the concert starts, there's no performance to watch, other than someone moving the faders on the desk up and down.

Despite the apparent profusion of speakers, however, an electroacoustic music concert is in fact laid out according to fairly simple and standard rules. The sound being reproduced is normally from a stereo source, and so each speaker in the room is part of a pair, each mirroring the other in left / right placement. (Some, more modern, arrangements allow for 8- or 16-channel multitrack reproduction, but that is usually an addition to the standard model rather than a normal occurrence.)
The speakers are also laid out with a preponderance at the front or marginally to the sides of the audience. Some are at the rear, but usually only one or two pairs. The speakers at the side and rear allow "a sense of being enveloped in sound; implications of circular motion on the tape can actually be made to circle around the room; and the introduction of sounds behind the listener can still have a startling effect (presumably because our racial memory still responds to aural warnings of potential threats from behind us)."[148]

[148] Harrison, J. (1998). *Sound, space, sculpture: some thoughts on the 'what', 'how' and 'why' of sound diffusion.* Organised Sound, 3(2), pp117-127. p122

Figure 1: Speaker array at a BEAST (Birmingham ElectroAcoustic Sound Theatre) concert[149]

The speaker arrangement and the smaller audience section creates a specific area where the sound will be heard to maximum aesthetic effect – the 'sweet spot'. The placement of the speakers allows for the *diffusion* of the stereo signal throughout the space – by manipulating the faders, the diffuser can materially alter the way the music is perceived. The focus of an electroacoustic concert is all on the sound – often the audience member will sit with their eyes closed, so that they can concentrate on the individual sounds making up the piece they're listening to, hearing the sounds as they swell and move around the space.

It is not unusual in these circumstances to experience a visual equivalent of the sounds, to perceive in the quiet behind our eyelids a sympathetic movement of colour and abstract forms, a synaesthetic interpretation of the sonic input being received. This inspired my own audiovisual compositions, and several other electroacoustic composers have also created audiovisual works in response to their auditory experiences in such concerts. For the composers following this route, it is a natural progression to make use of the familiar loudspeaker arrangement of an electroacoustic concert venue, with the addition of a flat screen in front of the audience on which the visual portion of the work is projected. The diffuser can then still direct the sound around the space, creating the same sonic atmosphere as would

[149] http://www.birmingham.ac.uk/facilities/ea-studios/about/meet-beast.aspx [accessed 27/04/2015].

be experienced in a ‘normal’ electroacoustic music concert.

My experience of audiovisual works in this concert format have generally been disappointing. The screen is usually quite small in relation to the space, and so it is easy for the viewer to be distracted from the visual projection. The forward focus onto the screen also ‘collapses’ the sonic surroundings into a fairly flat wall – as the viewer is concentrating on the image in front of them, the sound now appears to be created in relation to the objects on the screen, particularly given that the underlying sonic source is stereo in nature. When receiving information in both visual and auditory form, the human brain combines the two, and will try to make sense of where both sources originate from by imposing one ‘input source’ of localization onto the other. As Bregman notes, “other research that has synchronized light flashes with sound pulses has shown that the apparent location of events in one modality can be affected by our localization of the events in the other one. […] In a laboratory version of this effect, subjects were asked to point to locations of *both* visual and auditory signals, sequences of light flashes and sound pulses that occurred simultaneously in discrepant positions. Effects in both directions were found, but the effects on auditory localization were much greater.”[150] This reflects the ‘collapsing’ effect of the surrounding sound onto the screen, as the auditory localization is subsumed within the visual localization. Given this effect, the use of the side and rear speakers will normally be largely avoided by the diffuser, as it runs the risk of distracting the viewers from the fusion of sound with the visuals at the front, and disrupting the immersive nature of the piece. A skilled diffuser can make use of these speakers, but will always do so judiciously, to avoid the ‘startling effect’ noted above.

While most electroacoustic composition is stereo in nature, it is increasingly common for composers to work in 8-, 16- or 24- channels, feeding separate signals to specific loudspeakers, usually arranged in a ring or rings around the audience. For a purely sonic concert, this offers a very immersive experience, with the composer able

[150] Bregman, A.S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. The MIT Press. p183.

to select exactly where each sound will appear to originate from. However, with an audiovisual presentation, multi-channel sound suffers from the same dominance of the visual localization created by the flat screen, while the full use of all the speakers in the array will have a greater chance of that 'startling effect' that will break the audience's immersion in the piece.

While the electroacoustic concert format affords the audiovisual composer a venue in which the sonic surroundings and experience of the audience can be highly manipulated to immerse them in the audio perceptual stream of the work, the benefit is outweighed by the disadvantages it offers for the visual perceptual stream. The risk of breaking the audience out of the immersion offered by the complete audiovisual experience is heightened by the relatively small screen size and the distraction of the loudspeaker arrays, as well as the distraction that can be experienced if the rear and side speakers are used.

Cinema and immersive qualities

The cinema may initially seem to be the natural home of the immersive audiovisual experience, following years of development of both the image on screen, from monochrome to colour to widescreen, and the audio accompaniment, from live piano to record player to optical sound reproduction through to the current Dolby surround sound, but it is worth examining how the cinema has developed, and if its current form offers the most effective format for full immersion in the experience of watching audiovisual artwork, as opposed to the latest Hollywood blockbuster.

The very first films were shown unaccompanied by any sound or musical accompaniment – apart from the clattering of the early cinematic projectors. However, even at this point, audiences expected more than just the movements on screen. One early critic felt that street scenes were disturbing without the soundscape one would expect to hear. This led to experimentations with what we would now call 'live foley' effects, with sound effects created in sync with the action in screen. These effects were predominant in the 1900s to 1920s, but gradually were replaced by the louder musical orchestrations provided as specific scores for feature films, and then eventually by sound films themselves.[151]

While musical scores were created for some of the silent films, typically ones with larger budgets, the use of a piano or a small group of musicians to accompany the action on screen also became a regular feature - "Many early films, especially those that toured the vaudeville circuit, came with musical accompaniment on a piano (or sometimes a full orchestra) and often a narrator, who would fill in narrative gaps or otherwise enhance the experience."[152] The use of a narrator was a particular feature of early Japanese films, which developed out of their theatre. The *benshi* would narrate the film from a lectern, and would add commentary, sound effects, and explanations of the plot.[153]

The idea of adding synchronised sound to film was present from the start of the cinema – Thomas Edison in particular experimented with various mechanisms to co-ordinate sound created through his phonographic equipment to the images created on screen. However, the ultimate aim would be to include the soundtrack on the celluloid film itself, to ensure true synchronicity between the screen and loudspeakers. From early experiments by Eugene Lauste in London in 1907[154], several efforts were pursued in the United States and Europe to achieve this ideal.

Various films achieved sound synchronisation before *The Jazz Singer* premièred in October 1927, but this particular film has gone down in history as the one that cemented the arrival of synchronised sound in film, as it featured synchronised dialogue. Warner Brothers led the field with this film and subsequent releases, having established an exclusive relationship with Electrical Research Products Inc. (ERPI), a

[151] Allen, B. (2003). *I Am The Sound Effects Man*. Online Article available at https://web.archive.org/web/20061217184651/http://www.framtid.org/sound-effects-man.pdf [accessed 20/03/2015].
[152] Robb, B.J. (2007). *Silent Cinema.* Kamera Books. p139.
[153] Cousins, M. (2004). *The Story of Film.* Thunder's Mouth Press, New York. pp40-41.
[154] Geduld, H.M. (1975). *The Birth of the Talkies: From Edison to Jolson.* Indiana University Press. pp76-80.

division of AT&T and Western Electric, whose research department, Bell Labs, became the leading operation in developing sound technology for the cinema[155]. The other main film companies at the time had initially signed an agreement to work together using alternative technology, but in the event of Warner Brothers' success, they were also forced to move to ERPI's sound technology.

The development of synchronised sound would continue, with the arrival of RCA/RKO on the market, but the most important change for the cinema industry had already occurred – the introduction of synchronised sound to film proved to be immensely popular, drawing in audiences, lengthening the runs of the new talkies, and producing profits for the film companies. From the view of the audiovisual composer, the true innovation was that now the cinema had become an immersive audiovisual experience - "once a screening began, the viewer became absorbed into the film's story-world and oblivious to the physical environment of the auditorium."[156]

The introduction of synchronised sound to cinematic film also had a major effect on the editing process - "sound did not simply supplement the film image but transformed it perceptually, to powerfully affect editing possibilities."[157] The introduction of sound as an editable component of the film allowed the editor to include sound that continued over the visual cuts, to encourage a sense that action was indeed happening in the same location, to include music to heighten the emotional content of the scene, and to place action off screen, with sound heard from the side of the frame of vision bounded by the cinema screen. The techniques for using sound with film quickly became part of the editor's toolbox in Hollywood, and then spread out across the film-making world. "By the end of the 1930's, editing practice in virtually all of the world's film-producing countries is said to have come to resemble the continuity approach practiced in the United States."[158]

---

[155] Gomery, D. (2005). *The Coming of Sound.* Routledge, New York. pp31-32, 42-43.
[156] O'Brien, C. (2005). *Cinema's Conversion to Sound: Technology and Film Style in France and the U.S.* Indiana University Press. p34.
[157] *Ibid.* pp93-94.
[158] *Ibid.* p84.

Again, this is a development that is of importance to the audiovisual composer, particularly as the homogenisation of the relationship of sound to vision has been distributed across the world due to the dominance of the American film industry. Chion noted, “we classify sounds in relation to what we see in the image, and this classification is constantly subject to revision, depending on changes in what we see.”[159] I would also contend that we interpret images based on the sounds we hear, and the tradition of sound design that has been established by Hollywood has also inculcated a learned response by the audience. We now expect and accept certain sounds as a standard for the action we see on screen – one example is the Wilhelm scream, used (so far) in over two hundred films to indicate an overly dramatic death, usually of a throwaway character. As soon as a cinema-literate audience member hears that scream, they are reassured that the character’s death, while unfortunate, is luckily unlikely to affect the main plot.

This has a direct effect on the audiovisual composer – in creating an audiovisual work, in crafting the audiovisual object, the composer can work equally as well starting with a visual object and combining it with audio, or starting from a sound object and combining it with visuals. It is notable that the first synchronised sound cartoon, *Steamboat Willie* (1928), was created by matching animation to the soundtrack, rather than the other way around. In fact, Walt Disney had to sell his car in order to finance a second orchestral recording, as the first one did not adhere closely enough to the click track he had dictated to the conductor. Luckily for Disney, his investment paid off extremely well for him in the long run.

This approach, of matching the animated action to a pre-existing soundtrack, is commonplace in animation, and allows for creative work on both sides. Disney created a whole series of animations, his *Silly Symphonies*, which took recorded music as the heart of the piece and allowed the animators to experiment with different characters and techniques to match the soundtrack.[160] Douglas Kahn noted,

---

[159] Chion, M. (1994). *Audio-Vision: Sound on Screen*. Columbia University Press, New York. p88.
[160] From The Walt Disney Family Museum website. http://www.waltdisney.org/new-horizons-emergence-walt-disney-studio-1928-1940 [accessed 09/04/2015]

"Sound also became radical once it was tightly tied to cinema in the form of animated cartoons. Not only could sound and image exist in a pronounced one-to-one relationship, but sound came first in the production process instead of being secondary or tertiary to the primacy of visual image and the limited sounds of dialogue."[161]

The use of sound to provide inspiration for the animator is still in practice today. In a 2008 talk given to sound design students at the University of Edinburgh, the Hollywood sound designer Randy Thom described how he was approached to make the sounds of various dragons – roaring, snorting and coughing – which were later passed on to the animators on the film *How To Train Your Dragon* (2010).[162]

In the early synchronised sound cinemas, the reproduction was still monophonic or stereophonic, and so the same limitations are present that occur with the flat screen and essentially stereophonic sound in a standard electroacoustic concert – the sound is flattened and localized to the visual component, restricting the immersive opportunities available to the audiovisual composer. However, even from the early stages of synchronised sound, experiments were being carried out in introducing several channels and speaker positions to heighten the immersive experience. These experiments eventually solidified into a multi-channel form that has become a standard in cinemas, which has allowed the audience to become familiar with a format where they expect sound to come from the side and the rear, and still to remain absorbed in the audiovisual experience.

Surround Sound

The advent of multi-channel sound was relatively early, with experiments in the use of three or more channels starting in the early 1930s. Initial experiments by Bell Labs in 1933 investigated the use of multi-channel sound in the form of three speakers, the

[161] Kahn, D. (2001). *Noise Water Meat: A History of Sound in the Arts*. First MIT Press. p11.
[162] Thom, R. (2008). Talk at Alison House, University of Edinburgh – from my personal notes.

normal stereo pair of front right and front left, with the addition of a central front speaker. “The choice of three unique audio channels here was dictated by aesthetic concerns: while Bell engineers agreed that “an infinite number of front loudspeaker channels was desirable,” the use of left, center and right channels seemed to adequately produce the effect of a “full” stereophonic field without the “hole in the middle” effect common to reproductions using only two (front left and front right) channels.”[163]

While the adoption of a third centre speaker overcame some of the limitations of a stereo system, research continued into the use of speaker arrays arranged around the audience. The leader here was again Warner Brothers, whose Vitasound system introduced speakers that surrounded the audience, first trialled in 1940. The system still had to be refined – this first version had speakers surrounding the audience, but they were only used at specific parts of the film. Dialogue remained with the standard stereo speakers at the front, and only sequences with loud music and effects were switched to the surround speakers to maximise the impact on the audience. Rather than rely on the projectionist to change the settings as the film progressed, the Vitasound system inserted a control track on the film strip itself, between the sprocket holes, which signalled when the switches should be made.[164]

In terms of animation, Disney was also experimenting with the sound for his feature length films. His grand project *Fantasia* (1940) was aimed at bringing his version of visual music to the screen – Disney showed visual music films made by Oskar Fischinger to his animators for inspiration, and at one point Fischinger was attached to the film, but soon left due to artistic differences with Disney. In addition to the close fusion of music and animation that Disney set out to achieve, his ultimate aim was a cinematic form of *gesamtkunstwerk*, where the sound would surround the audience, the action on the screen at the front would be complemented with projectors throwing shadows and moving colours onto the sides and roof of the

[163] Kerins, M. (2011). *Beyond Dolby (Stereo): Cinema in the Digital Sound Age.* Indiana University Press, Bloomington & Indianapolis. p21.
[164] *Ibid.* p23.

auditorium, potentially using a 3-D effect on portions of the film. Possibly most intriguingly, he also planned to spray perfume into the audience during *The Nutcracker Suite*, which was represented on screen by a flower ballet. This would, of course, have had extensive resource implications for performances across the USA and other nations, and so his full ambitions were never realised.[165] In addition to the costs of performing the film as Disney originally envisaged, the release also suffered from the effects of the Second World War – a significant proportion of Disney's income up until this date had come from European distribution, which was completely disrupted, and the demands made on raw materials by the war effort limited the number of prints that could be made. Although *Fantasia* was publicly and critically well-received, it did not actually start making profit from its initial outlay until 1969, after several runs in the cinema.[166]

For *Fantasia*, Disney invested funds and research into his own version of multi-channel sound – Fantasound. The recordings were made in specially created segmented recording spaces, and the sound mix was designed so that individual speakers or groups of speakers were fed with specific audio recordings. As with the Bell Labs experiments, Disney employed three speakers at the front left, front centre and front right. Rather than feed the same signal to all three, as occurred with Vitasound, separate recorded channels were fed to the speakers, making this the first true stereophonic commercial film release. In addition, for selected screenings, more speakers were arranged around the audience – in the case of the film's première in Los Angeles, there were ninety-six additional speakers around the audience in addition to the three larger main front speakers. As with Vitasound, these were mainly employed at specific points during the film to expand and maximise the encompassing sound, such as the massed choir at the end of the film. Various different combinations of channels and speakers were tried during the development of the system - "intriguingly, one of the configurations tried but ultimately abandoned during the three-year process of developing Fantasound was a 3-channel front and 2-channel surround setup similar to the 5.1-channel arrangement DSS

[165] Smith, M.W. (2007). *The Total Work of Art: From Bayreuth to Cyberspace*. Routledge. pp119-120.
[166] Goldmark, D. & Taylor, Y. (eds.) (2002). *The Cartoon Music Book*. A Capella Books, Chicago.

systems would adopt half a century later."[167]

The use of surround speakers for *Fantasia* was not without controversy, as there are anecdotal reports that this first use of Fantasound had that startling effect on the audience that can be a problem in electroacoustic concerts featuring audiovisual work. In the cinema, this is referred to as an *exit door effect* - "Concerns over this [exit door] effect can be traced back to the very first cinematic use of surround sound: supposedly some audience members ran screaming from *Fantasia* when, late in the movie, sounds suddenly appeared in the rear of the theater."[168]

The earliest audiences experiencing surround sound in the cinema may be forgiven for finding the sudden introduction of sounds to the side and to the rear unsettling, as they were still reasonably recent converts to synchronised sound itself. For the audiovisual composer, this can remain a challenge – in the audio composition process, there is always a temptation to make full use of all channels, but it has to be remembered that the visual element compels the audience to concentrate on the scene and sound in front of them – sudden or unexpected sounds to the rear can be too distracting. That does not mean that the side and rear channels cannot be used, merely that a lot of careful planning and consideration has to be applied. A suitable use of the side and rear channels in a surround sound environment is to create an ambience, where the sound aims to gradually fill the space around the audience and create an enveloping atmosphere, which can then be punctuated with gestural sound and images at the front of the space. "Multi-channel ambiences ultimately form the foundation of diegetic immersion […] And because ambient sound so often goes unnoticed, it has the power to affect audiences without them noticing."[169]

While sound configurations in cinemas are still being developed further to create as immersive an experience as possible, while avoiding the 'exit door effect', these developments have in the main now centred around an accepted format, developed

---

[167] Kerins, M. (2011). *Beyond Dolby (Stereo): Cinema in the Digital Sound Age.* Indiana University Press, Bloomington & Indianapolis. pp23-24.

[168] *Ibid.* p158.

[169] *Ibid.* p167.

by Ray Dolby as an expansion of his noise reduction system.

Cinematic 5.1 Sound

Dolby's first true surround sound, used on the film *A Star Is Born* in 1976, was encoded using their Dolby Stereo system, which used a matrix from the left and right channels to create the centre and surround channels. However it was the use of Dolby Stereo on *Star Wars* (1977) and *Close Encounters of the Third Kind* (1977) that made the name Dolby synonymous with fully engaging cinematic sound reproduction.[170] *Star Wars* in particular impressed audiences with its soundtrack - "From sound architecture to spatial awareness, from sound texture to detail, from mixing to editing, from voice characterisation to physical sound, the film introduced a concept of sound that was finally willing to abandon its traditional shyness and move forward to claim a primary role."[171] Over the decade following the release of these two films, Dolby technology was installed in over six thousand cinemas across the world, and has become the standard reference point when discussing cinematic sound.

The richness of the Dolby Stereo sound came from its use of much wider sound strips on the film than had been used up until that point, allowing more tracks and greater definition of individual sounds within the mix. The music and effects were no longer muted in order to let the dialogue dominate – now the sound designer could attribute dominance to each of the three according to their importance in the narrative. The sound environment became much richer and more engaging for the audience. "Only then could noises have a living corporeal identity rather than merely exist as stereotypes."[172]

---

[170] http://www.dolby.com/us/en/about/history.html [accessed 22/04/2015].

[171] Sergi, G. (2004). *The Dolby era: Film sound in contemporary Hollywood*. Manchester University Press. p25.

[172] Chion, M. (1994). *Audio-Vision: Sound on Screen*. Columbia University Press, New York. p147.

Despite the advance in sound reproduction technology afforded by Dolby Stereo and its successors, such as Dolby Digital, the same principles in sound design that became established with the first forays into surround sound still apply. The surround speakers tend to be used solely for ambient sound, or emphasis of music and environmental effects. The majority of spot effects and dialogue remain contained in the front speakers. Sergei refers to this as the "one-wall" narrative principle, an accepted standard in cinematic sound design where all the focussed sound information for the audience is still contained in the sound emanating from the front wall, while the surround sound supports and augments only.[173]

The developments that Dolby have made in cinematic surround sound offer the audiovisual composer some advantages in creating audiovisual work to be performed in a modern cinema. The proliferation of the technology means that there is practically a worldwide network of performance venues with standard screen and sound reproduction facilities, so that the composer is always aware of how the work will be reproduced. This allows the composer to concentrate on how to achieve the best fusion of audiovisual objects to achieve a fully immersive piece, without having concerns about the varying standards of the venue, which can be extremely variable when playing work back in an electroacoustic concert setting. Secondly, the 5.1 standard is now very established, which, when the basic encoding and channel assignment is mastered, allows the composer to 'forget' about the technicalities of how to place the sounds so that they originate in the correct position, and concentrate again on creating closely fused audiovisual objects. Finally, as the majority of the cinema-going public are now very familiar with the Dolby technology and speaker placement, they are more accepting of sound appearing to come from the sides and rear of the auditorium, which gives the composer more freedom in placing the sounds in the composition – although care still has to be exercised, as the accepted cinematic sound design standards still avoid placing much more than ambient background sound and music in the rear left and rear right speakers. However, as

[173] Sergi, G. (2004). *The Dolby era: Film sound in contemporary Hollywood*. Manchester University Press. pp30-31.

Mark Kerins has observed, “at least one thing seems clear: digital surround sound will only grow in its importance to the audio-visual arts for the foreseeable future.”[174]

Multi-screen formats

Moving on from the sound production facilities in the performance space, the visual space can also be improved from a standard flat projection screen, as used in electroacoustic concerts or in smaller cinemas. (Larger cinema screen are generally slightly curved, especially if they are designed for use with 70mm prints or digital films, although the curvature is usually not pronounced enough to be noticeable). In a similar fashion to surround sound, the screen can be altered to achieve ‘surround vision’, where the entirety of the visual field perceptible by the audience member is filled with the visual projection. This can be achieved by projecting on to an overwhelming curved surface, such as occurs in IMAX cinemas or planetariums. Due to budgetary or space constraints, however, another approach is the multi-screen format, most often realised as three screens arranged together. The two side screens are usually angled in relation to the main central screen, at 30º or 45º to the plane of the central screen. This triptych arrangement is familiar from its use from the Middle Ages onwards in Roman Catholic and other Christian churches, often for altarpieces. Typically, the central image conveys the key message from the artist, while the two side panels support it with additional relevant scenes. In a similar fashion, most triptych multi-screen formats also concentrate on the main screen, with supporting material towards the sides. This mirrors the use of the side and rear speakers for mainly ambient soundscapes and music in surround sound cinema.

The use of this format has some advantages for the audiovisual composer. The expanded screen area means that the audiovisual objects forming the composition can be more easily fused together within the sonic and visual space – an object

[174] Kerins, M. (2011). *Beyond Dolby (Stereo): Cinema in the Digital Sound Age.* Indiana University Press, Bloomington & Indianapolis. p327.

moving from one side to the other across the three screens and the sonic space will be perceived as more coherent than it might on a single screen. This counteracts the 'flattening' of the sonic space towards the front of the performance space when a single screen is used. By avoiding this flattening, and by introducing audiovisual elements at the side of vision, the composer is also able to ameliorate the 'exit door' effect, as the audience is less likely to react to an obtrusive noise at the sides if there is a corresponding visual element. The distraction of a sound at the rear of the auditorium will always be problem for audiovisual works, and so only judicious use of the far rear speakers for ambient sound can really be a factor in audiovisual composition.

Although the multi-screen format can offer a greater canvas on which to work, this increase in available screen space does have its own problems, particularly when crafting the visual portion of the composition. Although most animation and video compositing packages are now optimised for use across two monitors, if the composer treats the three eventual screens as one long interlinked panoramic view for the purposes of creating the work, the scaling necessary will cause the vertical aspect to be very constrained, and details within the work difficult to discern. The composer can, of course, switch between this view for an overall impression, and a closer view of the details, or work on one 'screen' at a time, but this can become a tiresome task over time. Perhaps more importantly, the physical separation of the screens can also pose a challenge for the audience. The space between the screens become vertical 'lines' that can break up the connections between the screens too much for the audience to be fully immersed in the work. In the best cases, the lines can be ignored in the same way that the 'lines' separating the windscreen and front side windows of a car become ignored by a driver, but this will always be a risk until the work is realised and performed.

A good example of work that overcomes this screen separation issue is *cMatrix12* by Dr. Bret Battey.[175] In this work, the projection is onto a flat wall, but as three separate

[175] http://www.mti.dmu.ac.uk/~bbattey/Gallery/cmatrix10.html [accessed 27/04/2015].

screens. Interestingly, although the side screens are projected on to the same wall as the central one, they are angled outwards and downwards at the extreme edges, giving the impression that they are placed at an angle to the main screen, in the form of a triptych. This was not specifically planned, but evolved in response to the space in which the performance was staged. The commission was for a space which had a long wall, which Dr. Battey wanted to fill, leading to the three screen format. He notes, "We splayed out the left and right image to create more of a sense of a wraparound configuration - and I liked the introduction of the angles, breaking the monotony of straight three images. And again, this splaying suggests aspects of a standing triptych. It also helps emphasise the focus of activity in the centre-left of the wall and the fact that the image dissolves outward into space. I also was intrigued by the tension setup by the intersection of the geometry of the image with the two vertical columns - which probably makes the piece more interesting than if it were a continuous screen. I didn't start out explicitly with this plan, but it evolved as I responded to the realities of the space."[176]

Figure 2: *cMatrix12* at the Phoenix Square Film and Digital Media Centre, Leicester. November-December 2009

This work was composed by making use of generative software for the visuals, with sound generated by tracking individual particles within the visual fields. The visual focus is on the intersection between the left and centre screens, while the sound fills the space through a quadraphonic speaker array. The composition developed from

[176] From email conversation with Dr. Bret Battey, 27/04/2015.

many iterations of a process that Dr. Battey started from a fortuitous error in some of his visual effects coding, and so the expansion to a multi-screen format was perhaps less of a risk than it might otherwise have been, and a generative process also avoids lengthy positioning of models and effects across a long panoramic visual field. The complexity of the coding has resulted in a very immersive and absorbing work, reflecting Dr. Battey's skill in audiovisual composition.

While works such as *cMatrix12* illustrate how the multi-screen format can be successfully exploited by the audiovisual composer, a commercial version of a continuous screen dominating the human front-facing visual field has also established itself as a standard, and may also be a mechanism for realising the immersion sought for in audiovisual performance.

IMAX

The huge, overwhelming screen found in an IMAX venue is a direct development of the increasing size of cinema screens, a trend that started in the 1950s as sound also developed from a central speaker at the front, through stereo, to 5.1 surround sound. "The clearly delineated segregation of spaces which had characterized previous conditions of motion picture spectatorship gave way to an illusory integration of spaces in which images and sounds from the "fictional" space of the motion picture appeared to enter the "actual" space of the audience; the audience, thus surrounded by images and sounds, felt itself to be a part of the space depicted on the screen."[177]

This 'involvement' of the audience within the space of the film itself is of course the immersion that is sought after in audiovisual performance. The developers of IMAX initially started to develop a multi-projector, multi-screen system, premièred at the *Labyrinth* Pavilion at Expo '67 in Montreal, Canada, but rapidly decided that a single large screen system with a single projector was both technically less complicated and

[177] Altman, R. (ed). (1992). *Sound Theory Sound Practice*. Routledge, London. Chapter 8: Belton, John. *1950s Magnetic Sound: The Frozen Revolution*. p154.

aesthetically more impressive.[178] The first IMAX film *Tiger Child* (1970) appeared at Expo ’70 in Osaka, Japan, and heralded IMAX’s initial dominance by large-spectacle film experiences, typically documentary in nature, exploring the natural world in overwhelming detail.

Figure 3: IMAX audience watching a film about space[179]

In modern cinemas, the sound reproduction is usually through the standard Dolby systems, but the dominance of the huge screen at the front adds to the immersion in sound with immersion in vision. “Being immersed in the image is the defining feature of the IMAX brand, whether in flat-screen theaters, 3-D or Dome screens; “Be There” could easily be the corporation’s trademarked tagline”.[180]

The limiting factors for the audiovisual composer in creating works for an IMAX screen are the high costs involved and the limited access to IMAX performance venues. While most large cities have at least one cinema with an IMAX screen, the investment in the technology dictates that the screen will have to carry popular, high-grossing releases in order to justify its costs, and it is highly unlikely that any shorter, experimental audiovisual work will be shown. Equally, the composer would have to have access to a facility allowing them to work with the high-definition encoding

[178] https://www.imax.com/about/history/ [accessed 27/04/2015].
[179] Image from The Franklin Institute Science Museum. https://www.fi.edu/movies [accessed 27/04/2015].
[180] Griffiths, A. (2008). *Shivers Down Your Spine: Cinema, Museums, and the Immersive View*. Columbia University Press. p94.

used in producing IMAX films, which again incurs a high cost given the level of investment in the technology.

As technology continues to develop rapidly, IMAX cinemas and the format may become as common and as accessible as small-screen cinemas, so offering a potential venue for performance. Until that time arrives, the audiovisual composer will remain restricted to lower-cost and more accessible venues. This is perhaps a regrettable situation, as the immersive nature of the IMAX venue would seem ideal for audiovisual works. A deputy director of the National Air and Space Museum in Washington DC, USA noted "the "psychological impact of this new medium," which for him defied description. An experience "far more than a 'wide screen movie,' IMAX made the "willing suspension of disbelief unbelievably easy," promoting "'on-the-spot' excitement in the viewer.""[181]

Hemispheric Projection

At the same time as the cinema was developing, exploring how best to create an enveloping audio and visual entertainment, another projected image and sound vehicle was being established, primarily as an educational tool for disseminating astronomical images and information. The first recognisable modern versions of the planetarium, with hemispheric projections of the night sky on the interior of a dome, were built in Germany in the 1920s. While the Second World War interrupted the spread of planetariums across Europe and the USA, the post-war environment saw a reasonably rapid uptake in building planetariums, especially as additions to science museums and academic establishments. The majority of these planetariums were, and still are, used in the main for their original purpose, to show information relevant to astronomical phenomena, although the shows now may be more 'produced' than the simpler programmes originally run.

[181] *Ibid.* p224.

In essence, a planetarium is also an expansion of the traditional cinema. In this case, the curvature of the screen has been taken to its logical extreme, a hemisphere facing the audience. The sound is relayed from speakers placed behind this hemispheric screen, usually in an array allowing an effectively ambisonic sound reproduction system – the sounds can be made to appear to emanate from any point on the screen, allowing an intensely close coupling between sound and image. This is all achieved by placing the audience member either directly onto the floor, or in a reclining chair where their head is positioned to make viewing the hemispheric screen above them a reasonably comfortable experience. The entire purpose is to make the audience feel totally immersed in the experience - "the planetarium experience takes place inside a dome where a "virtual" reality is illusionistically constructed; spectators are to imagine that once they take their seat, they are magically transported to outer space and possibly another time, depending on the subject of the planetarium performance"[182]

It should be noted that although the normal expectation for hemispheric projection is now within the planetarium, this need not necessarily be the case. The Director of the Interactive Virtual Environments Centre in Australia (iVEC), Paul Bourke, has developed a single-user version of a planetarium that uses hemispheric projection via a spherical mirror to project onto a dome facing a seated user.[183]

[182] *Ibid.* p116.
[183] http://paulbourke.net/dome/iDome/ [accessed 27/04/2015].

Figure 4: iDome showing user exploring an archaeological site.[184]

Using hemispheric projection to achieve audience immersion in an audiovisual work has already happened, in a series of concerts that took place in the USA from 1957 to 1960. The Vortex Concerts at San Francisco's Morrison Planetarium were among the earliest examples of electronic sound and projected light environments. Conceived for the acoustical properties and colossal scale of the Planetarium's domed ceiling, a vortex of light and sound was created by the painter and filmmaker Jordan Belson and the composer, radio engineer, and programmer Henry Jacobs. The Vortex Concerts broadened conventional cinema from a single projected image to the expansive curve of the dome as a screen for Belson's film projections accompanied by a live mix of amplified, multi-channel sound. His multi-image film projections immersed the audience in an environment of luminous imagery overhead, bathed in tinted film imagery and color washes. A bank of prepared projectors freed the image from its frame, while spatially composed sound by Jacobs (and others), amplified

[184] *Ibid.*

through a circular arrangement of more than three dozen speakers, saturated the audience with multidimensional sounds and images."[185]

Belson's work for the Vortex Concerts included complex layers of strong textural images – waterfalls, waves crashing on beaches, flames roiling across the screen, all coloured, overlaid and interacting with each other. In later works Belson would either create his own sound composition or use classical pieces, but the work by Jacobs complemented his early pieces well and created audiovisual objects within the compositions. In assessing the inspiration behind the work, Robert Riley considered that the prime motivation for Belson and Jacobs was to create a fully immersive, synaesthetic experience of moving image and sound, overwhelming the audience in the two perceptual streams. Riley compared this use of a planetarium with its hemispheric projection and multi-speaker, multi-channel surround sound to the Bauhaus's earlier design for a *Totaltheater,* proposed by Gropius and expanded upon by Moholy-Nagy, following on from their pioneering experiments in using light as an artistic medium.[186]

At the same time as Belson created the Vortex Concerts, the Whitney brothers were developing their experimental approach to film and music – marking this as a particularly productive period for visual music on the West Coast of America.[187] This early use of a planetarium has unfortunately not continued as a regular performance venue for audiovisual concerts, but it has established that this would be a feasible proposition. There have been some commercial developments using hemispheric projection and electroacoustic compositions to accompany moving images, such as an opening exhibition at the visitor attraction *Our Dynamic Earth* in Edinburgh, where the composer Professor Pete Stollery created a soundscape for a gallery exhibiting the movements of magma and lava flows.[188]

---

185 Bernstein, D.W. (ed.) (2008). *The San Francisco Tape Music Center: 1960s Counterculture and the Avant-garde*. University of California Press. Insert: Riley, Robert R. *Liquid to Light*. p21.
186 *Ibid*. pp22-23.
187 Bernstein, D.W. (ed.) (2008). *The San Francisco Tape Music Center: 1960s Counterculture and the Avant-garde*. University of California Press. Chapter: Welsh, Thomas M. *Chronology*. p143.
188 http://www.petestollery.com/ List of Works. [accessed 27/04/2015].

However, these commercial works have generally concentrated on providing a sound accompaniment to an educational display, consistent with the original aim of planetariums to inform, but not building upon the potential of the Vortex Concerts. This would appear to be a fertile ground for the new wave of audiovisual composers who have started to create work in recent years, offering a much more immersive environment than is achievable in standard concert venues, or even in cinemas, given that IMAX is realistically out of reach for most audiovisual composers. Most cities across the UK have at least one planetarium within easy access distance, and although the programmes will be determined in advance, because there is usually a link with a University Astronomical department, they can often be approached via this route. Some Universities and other Astronomical educational bodies also have small scale planetariums which are transportable, and which may be useful for initial project work, although these are often inflatable domes, which then suffer from noise pollution from fans, engines and generators keeping the structure inflated.

Another possibility is the construction of a dome equivalent within a venue, either as a permanent or temporary structure. Plans are available for constructing geodesic domes from a variety of materials, usually to create a lattice of interlinking beams. Speakers can be attached to these beams in a regular array, allowing highly directional sound control, while material is attached to the beams creating a sound permeable screen for projection. The most expensive item for such a structure is a central hemispheric projector, which would argue the case for a permanent installation if purchased. The Australian sound designer and composer Mark Pedersen has developed a 26.2 hemispheric structure for a similar purpose, although in his case he has used individually placed cloth screens and standard projectors to add visuals within the surrounding soundscapes.[189]

[189] Pedersen, M. (2013). *Sound Labyrinth: Exploration Of The Embodied Sublime Through An Immersive Audio/Visual Installation.* Published online at academia.edu, accessible at http://www.academia.edu/8651102/SOUND_LABYRINTH_EXPLORATION_OF_THE_EMBODIED_SUBLIME_THROUGH_AN_IMMERSIVE_AUDIO_VISUAL_INSTALLATION [accessed 27/04/2015].

Apart from the challenge of physically constructing a dome as a performance venue, this then offers the added challenge for the audiovisual composer of creating sound for a 26.2 multi-channel system or similar, plus creating visuals for reproduction through a hemispheric projector, which requires consideration of the final image and encoding to achieve a smooth projection across the dome. The 'sweet spot' for the best experience of the sound is also a limiting factor, making this a venue more suitable for a small audience. However, for an audiovisual composer who wants intense audience immersion in their work, this option would seem to be the most achievable.

Site Specific Sound and Vision

While IMAX and hemispheric projection offer a generic immersive audiovisual experience, allowing works to be performed in a replicable, standardised environment, audiovisual performance can also take the forms of site specific work or installations – often using projection mapping techniques to feature bespoke animation and generative visuals, projected on to a variety of surfaces available in the venue, with audio material likewise adapted and presented in a format that makes use of the specific sonic qualities of the space.

In 1999, the composer Brian Eno, known for his experimental and ambient work, collaborated with the Italian painter and sculptor Mimmo Paladino to create the immersive art work *I Dormienti* housed in the Undercroft of the Roundhouse in Camden, London. The installation featured a collection of sculptures by Paladino – bodies, both whole and in sections, arranged over the central chamber floor in various poses, along with branch corridors with clay crocodiles arranged in a snaking queue. Eno made use of the unique structure of the Undercroft, using ten portable CD players, each set to random play, and containing a unique CD featuring between three and thirty tracks written by Eno. The individual tracks being played back could be heard along the corridors, while they mixed and merged in the central chamber. Given the random nature of the playback mechanism, and the variations on echoes affected by both the architecture and the visitors to the installation, the sound accompanying the visual spectacle was constantly changing, creating new sonic combinations[190]. One reviewer described how the interacted with the space and the sculptures, "the ambient aquarium tones of Brian Eno's music float to you in the darkness from afar… the music stays motion, curls around the figures; it has neither beginning nor end"[191].

With just the minimal illumination offered by the normal Undercroft lighting, the shadowed, chambered space encouraged the visitor to let the experience wash over them, to contemplate the combination of sculpture and structural sound. Writing in *Contemporary Visual Arts*, the critic Hugh Stoddart noted "what one heard possessed an archaic simplicity: a singing voice, and repeated reverberant cadences… I have to report that people seemed utterly held by it. They spent a long time here, peering unselfconsciously at details of the sculpture, listening intently to the music, wandering around in the strange architecture… I found it memorable. The components fitted; it worked."[192]

By taking advantage of the central chamber and side corridor structure, Eno could allow for the sound emanating from the various speakers to overlap, interact and reinforce his musical intentions for the piece, to complement the slightly unearthly visual experience of seeing the recumbent, broken bodies scattered throughout the space. The combination of the sculpture and architecture as visual elements and definition of the background space, interacting with the pervasive surrounding soundscape, created an immersive audiovisual experience which could vary according to the viewer's own position within the space, varying with position and movement. An installation of this nature can be a very affecting and enveloping experience, but generally is decidedly site-specific, and therefore there is an inherently limited potential audience that can experience it.

[190] Putnam, J. Exhibition Notes available at http://www.jamesputnam.org.uk/inv_exhibition_08.html [accessed 30/07/2016].
[191] Peaker, C. (2000). *C Magazine* 64. C Arts Publishing & Production. p.29.
[192] Stoddart, H. (2000). Review for *Contemporary Visual Arts no 26*, available at http://www.hughstoddart.co.uk/pdfs/eno_paladino.pdf [accessed 01/05/2017].

Another example of an installed work that relies on the specific location it occupies was *Pour Your Body Out* (2008), an audiovisual installation by Pipilotti Rist at the Museum of Modern Art (MoMA) in New York. In this installation, three walls of the museum atrium were used as the screens on which a six-channel video projection was cast. The video images were over 7.5 meters high, brightly coloured but designed to blend smoothly across the space and between scenes. The walls were not uniformly flat – "the protrusions in the walls do not disrupt the cohesion of the frame, but rather they add another level of stretching and pouring, as the title suggests"[193]. The accompanying sound also blended slow minor chords to accompany the ever changing visuals, and Rist designed a central, comfortable circular seating area on which the audience could relax and absorb the experience. This installation followed on from several such as *Ever Is Over All* (1997), also at MoMA, which featured two projections at right angles to each other, on the adjacent walls in the corner of a room[194]; and *Sip the Ocean* (2001), which used a mirrored set of projections, on opposite walls. As Rist's artistic journey developed, she expanded her domination of the exhibition space, with each new work occupying more of the surrounding surfaces,

By dominating such a large space, and making a feature of sound, vision and even the surface on which you could stop and absorb the surrounding environment, Rist also ensured that this was a shared, immersive experience. As with the Eno/Paladino installation at the Round house, this installation relied heavily on occupying a space with specific properties. Unlike *I Dormienti*, however, *Pour Your Body Out* is potentially transferrable to other venues, although finding a space which offers the same scale as MoMA's Marron atrium would be a challenge. Her recent *Mercy Garden Retour Skin* (2014)[195] conquered the final fourth wall, encompassing the audience from all sides with her colourful, languid audiovisual work. The slow,

---

193 Kane, C. (2011). *The synthetic color sense of Pipilotti Rist, or, Deleuzian color theory for electronic modern art*. In *Visual Communication*. Vol 10, Issue 4 pp475-497.

194 Rist, P. (1997). *Ever is Over All*. Catalogue Listing at Museum of Modern Art, New York. Available at https://www.moma.org/collection/works/81191 [accessed 01/05/2017]

195 Rist, P. (2014). *Mercy Garden Retour Skin*. Installation at Museum of Contemporary Art, Sydney as part of 29th Biennale of Sydney 2014. Catalogue listing available at http://www.biennaleofsydney.com.au/19bos/artists/rist/ [accessed 01/05/2017].

developing nature of her multimedia artwork encourages the audience to relax into it and suspend their preoccupations with daily routine and minutiae, with the domination of her artwork in every direction you look, fusing with the music by Heinz Rohrer also pervading the entire space. This domination of the visual and audio senses is another method of achieving an immersive audiovisual experience.

Inter-disciplinary Audiovisual Collaborations

By combining scientific exploration with artistic aesthetics, the danceroom Spectroscopy (dS) project utilises techniques for simulating motion dynamics, based on data gathered from depth sensors within an immersive environment. The data gathered is used to generate both sound and video, based on transformation algorithms, such as a fast Fourier transform to calculate a spectral decomposition[196]. While the combined hardware and software that has been created by the project team lends itself to a variety of purposes, one of the early applications has been in the artistic field, using the motion of dancers to drive a surrounding soundscape and generative video projections. The dS system was used to create the performance piece *Hidden Fields*, first performed in Bristol at the Arnolfini in 2012[197]. The performance featured a group of dancers, tracked by depth sensors, with the data feeding interactive graphics and soundscapes. The audiovisual responses to the dancers' movements are themselves responded to by the dancers, resulting in a very interconnected experience.

The complex relationships achieved between dancers, graphics and sounds driven by depth sensors achieves a very immersive performance, but this is the result of a highly technical, very complex collaboration between scientists, artists and composers with a wide set of skills. Achieving this without such a support structure would be virtually impossible for most audiovisual composers. However, the

[196] Glowacki, D.R., O'Connor, M., Calabró, G., Price, J., Tew, P., Mitchell, T., Hyde, J., Tew, D.P., Coughtrie, D.J., McIntosh-Smith, S. (2014). *A GPU-accelerated immersive audio-visual framework for interaction with molecular dynamics using consumer depth sensors*. In *Faraday Discussions 2014*, 169, 63-87. p66.

[197] http://www.danceroom-spec.com/project/dance-2/ [accessed 01/05/2017].

combination of audiovisual compositions and dance also occurs elsewhere, where there is less technical interaction between the dancer and a responsive environment, but where there is collaboration from the start between the participants to achieve the desired immersive artistic performance.

Nikhel Mahajan, a sound designer based in New Delhi, became interested in electronic music as a teenager, learning as much as he could working in various studios while supporting himself as a graphic artist. His music and art have also combined with his interest in ayurvedic practices and meditation, which has inspired his Sattyananda projects[198]. Exploring electronic music and sound design to add to a meditative and holistic experience, he also creates audiovisual fusions to further immerse the audience within a contemplative environment. In more recent experimental work, he has collaborated with the dancer Lalita Shivani to introduce her kinetic motion to add to the overall experience.

---

[198] Conversation with Nikhel Mahajan, May 2016, also http://sattyananda.com/about/ [accessed 01/05/2017].

Figure 5: Sattyananda performance featuring Lalita Shivani, Moscow, April 2017[199]

The combination of audiovisual composition with other creative artists opens up additional methods to add to the overall experience of experiencing an audiovisual experience. In the case of the Sattyananda performances with Lalita Shivani, the dancer evokes traditional and contemporary dance movements reflecting Indian culture, which combines with the audiovisual composition by Nikhel Mahajan to reinforce his underlying concepts of spirituality and meditation. In comparison to Eno and Paladino's *I Dormienti* with its static sculptures, this combination of dance and audiovisual composition also achieves a spiritual, meditative immersive effect, but with a more immediately engaging, kinetic and enticing inter-disciplinary approach.

[199] Reproduced by kind permission of Nikhel Mahajan

Virtual Reality – Bespoke Immersion

In all of the potential performance spaces discussed so far, a common element has been that the experience of watching and listening to the audiovisual works being presented is in a shared environment, where several people are present in the auditorium. However, this shared public experience is not the only performance space that can be used to show audiovisual works. The advent of television was seen as a potential threat to the cinema – while this has not proved to be the case, it has placed audiovisual reproduction technology in virtually every Western home. More recent developments in computer and videogame technology, as well as the introduction of "home cinema", has also brought high definition visuals and surround sound to the individual at home.

The television and/or computer screen with built-in or separate speakers may be seen as a direct competitor to the cinema screen and speakers, but the true competitor for a fully immersive audiovisual experience is virtual reality, especially when experienced through a dedicated headset including headphones. This "new area of research emerged in the 1980s: the development of *virtual realities*. The objective is to make the entire *experience* of a simulation of reality seem completely real. If the objective of much of today's work in graphics is to create *images* that appear to have a quality indistinguishable from a photograph of the real world, the objective of work in virtual reality is to create *experiences* that have a quality indistinguishable from an experience of the real world. Virtual reality is the ultimate simulation."[200]

Although the first steps in virtual reality were taken in the 1980s, it has only been in the last decade that computer technology has become advanced enough to realise a reasonably convincing level of detail in both sound and moving image to allow the experience to become truly convincing and immersive. The limitations of earlier screen resolutions and cathode ray tube monitors meant that the first attempts at virtual realities and game environments had an inherent lack of immersiveness -

[200] Holtzman, S.A. (1994). *Digital Mantras: The Languages of Abstract and Virtual Worlds*. Massachusetts Institute of Technology. p197.

"no-one in their right mind would argue that staring at a low-resolution image on your computer screen that you can navigate around in any way approximates the experience of actually being in that very same space or even the visceral thrill of standing in the middle of a 360-degree panorama."[201] While there were some attempts at creating wearable hardware to create an immersive environment, the technical limitations meant that these early headsets were physically restrictive and uncomfortable to use for lengthy periods. "Early models of the head-mounted display were awkward and heavy, uncomfortable to wear for more than a few minutes. The tracking of the wearer's movements was crude, the resolution and colour of the screens poor, and the rate at which pictures were displayed slow and out of step with the wearer's movements."[202]

The development of photorealistic graphics and high-speed graphic capabilities along with light, LCD and OLED flatscreen technology has now allowed the construction of tailored headsets incorporating headphones and high resolution screens, effectively creating that 360-degree panorama for the user, an immersive, convincing alternative world. The virtual reality environment presents the audiovisual composer with a great opportunity to create a truly immersive experience, but increases the complexity of creating audiovisual work immensely. For this environment, the tools and techniques of the gaming industry become an asset to be developed and employed by the composer.

A popular virtual reality platform is *Second Life*, developed by Linden Labs in 2003. This is a virtual world wherein participants interact with their surroundings and other users through avatars, 3D CGI models which may or may not mimic their real-life appearance. *Second Life* bears a similarity to many online role-playing games in its construction, but is not primarily concerned with a game environment. It has been used for online meetings, a variety of online businesses, and more controversially, for gambling and pornographic activities, although these have been curtailed greatly by changes in Linden Labs policies over the years. Of particular interest is the use of

---

201 *Ibid.* p80.
202 Woolley, B. (1993). *Virtual Worlds*. Penguin Books. p238.

Second Life for music concerts, where the music is created either through a direct online "chat" facility where the performer's voice is broadcast live, through a streaming service, or samples triggered by actions within the virtual world. There are specific areas or locations in *Second Life* that specialise in specific music genres, such as country and folk, jazz and blues, experimental music, etc.[203] However, while *Second Life* offers an online concert venue accessible to participants spread across the world, there are inbuilt limitations that make it an unattractive proposition for the audiovisual composer, as the visual elements are restricted to the game engine standard designs provided by Linden Labs, severely limiting the flexibility in creating fully integrated audiovisual objects.

The gaming environment does however offer more opportunities for audiovisual work. While individual game engines have recognisable visual styles, there are enough engines available for the composer to choose one sympathetic to their own visual style to work with, allowing more flexibility for visual composition. In contrast to the requirements of the online role-playing game standards, the audiovisual composer does not need to allow for true interactivity in their work, where the 'in-world avatar' has to be able to walk around and interact with objects in the virtual world – this would actually detract from the primary concern of full immersion. The virtual world can be treated as the equivalent of a concert hall experience – you enter the world, occupy a single fixed space, equivalent to a seat in a concert hall, and then experience the surrounding sound and vision as if you were really present in the created environment. The audience member can look all around, above and behind them to experience the full surrounding audiovisual composition, but is not able to directly influence the work as they would in a normal game environment. This allows the composer to concentrate solely on the experience they wish to craft, rather than be diverted with game mechanics and storylines.

This particular method of performing audiovisual work is very immersive, as the headphones and dedicated screens block out the real world and replace it with the

[203] *Second Life* music destinations are listed at http://secondlife.com/destinations/music [accessed 14/06/2015]

one designed by the composer. There also has to be a sympathy for the user in placing them into this new world, allowing a transition from reality to virtual reality, an ‘entrance hall’ through to the main auditorium to allow the user to become settled into place in this new environment. One of the earlier artistic works using virtual reality presented through a dedicated headset and bodysuit was Char Davies’ project Osmose (*1995)*, which completely replaced the normal world with a new reality. “The equipment for *Osmose* is a stereoscopic 3-D head-mounted display and a motion-sensitive bodysuit with a breathing and balance sensor, all of which is worn by the “immersant”. […] For the immersant herself, however, the experience is solitary and all-consuming: once the head-mounted display and the bodysuit are on, the immersant has very little sensory access to the exterior world.”[204] While the act of donning a heavy headset (as it would have been in 1995) and a full bodysuit would provide that transition period for the user, modern VR headsets are much more immediately usable, taking less effort to put on than a motorcycling helmet, and so the composer must be aware of the potential to confuse the user at the start of the process and provide a suitable introduction to the experience.

There are several advantages to working in a virtual reality environment. When the soundscape is designed to be heard through high-quality headphones, the composer can work comfortably with binaural recordings and processes, placing the audience member at the precise centre of the sonic world. The user is aware that they are in a safe virtual environment, so the perceived ‘exit-door’ threat is minimised, allowing sound to appear to come from all directions, especially as it will encourage the user to move their head and look at all the aspects of the audiovisual world they are experiencing. The visual world is likewise made more accessible, with the visual portion of audiovisual objects being present even if the sound originates from an area normally ‘off-screen’, but the composer can still introduce objects from a distance or up from the ground, allowing audiovisual themes or melodies to emerge and be recognised.

---

[204] Smith, M.W. (2007). *The Total Work of Art: From Bayreuth to Cyberspace*. Routledge. p160.

VR technology is also becoming more affordable, with the imminent release of the *Oculus Rift* headset, which though still expensive is still in the price range affordable to most early adopters of new technology. If this becomes a standard for VR delivery, as 5.1 surround sound has in the cinema, it also provides a universal standard for the audiovisual composer working in virtual reality. This allows their work to be easily distributable across the world, with no need to have a specific location or time for a performance. The virtual reality audiovisual artwork may become the most accessible format possible in which to distribute composed works.

Figure 6: Oculus Rift and game controllers in action[205]

However, there are also some disadvantages. The *Oculus Rift* technology is proprietary, which may limit the access composers have to the software and programming tools needed to create and encode the composed works. The quality of the equipment can vary as well, as the headsets allow the user to substitute their own headphones, which can result in great differences in response and quality from the standard set distributed with the headset. One particular sonic issue is the restriction

[205] Image from the Oculus Rift website (blog post dated 11th June 2015). https://www.oculus.com/en-us/blog/the-oculus-rift-oculus-touch-and-vr-games-at-e3/ [accessed 15/06/2015].

of true bass when using headphones, which reduces the impact of the ground-shuddering bass that can be experienced in cinemas and concert halls. The control of the composer's Intellectual Property Rights may also become an issue – does that rest solely with the composer, or will the creators of the hardware and software platforms also have a potential interest? A final issue is the flipside of creating a tailored individual experience – if you're not part of a large audience, there is no sense of occasion, that shared experience that can be talked about and discussed afterwards.

There is no doubt that the development of virtual reality will result in new artforms and experiences for the audience. The introduction of affordable virtual reality systems in the form of the *Oculus Rift* and similar headsets offers a potential new venue for realising immersive audiovisual compositions. However, as with any new technology, the creative collaborations still have to be established, and an open environment fostered where experimental art can be created within the new technology. This field may initially be more attractive to audiovisual composers with familiarity with games technology and audio composition for the games business, as it is likely that the initial encoding and software tools that can be used to create work for the virtual reality format will be based on existing games software.

Conclusions

Wagner's initial ideas of *gesamtkunstwerk*, realised in the closest match he could achieve by way of his Bayreuth theatre innovations, are now much more achievable. The creation of immersive audiovisual work is achieved by a fusion of creativity with modern technology, and as the ability of technology to provide an immersive platform increases, so too does the immersiveness and believability of the audiovisual experience, where believability measures the accomplishment of the environment as a coherent virtual world, rather than implying it has to be able to potentially exist in the real world.

The successors to Wagner continued to explore the limits of the achievable in terms of submersing the audience into the visual and audio experience they wanted to create. The Bauhaus expanded on Wagner's pioneering work to explore the concept of *gesamtwerk* – where Moholy-Nagy contended that all human endeavour contributes towards a creative vision. Gropius's vision of a *Totaltheater*, where the audience would be surrounded on all sides by light projections and sounds, presaged later developments in audiovisual performance spaces such as the IMAX cinema and planetariums.

For the audiovisual composer, the performance space in which their work will be performed has to be considered from the start, not only for technical considerations such as the number of audio channels being used, but also from a creative viewpoint. When an audiovisual composition is performed in an electroacoustic concert venue, there is a great deal of control of the audio environment, but the small flat screen at the front limits both the visual realm and collapses the sound to a front wall, limiting the opportunities to develop the sonic space. A standard cinema allows a more immersive visual spectacle, but also limits the sonic environment, especially where the composer is unable to directly control the audio playback.

For the most immersive concert experiences, open to more than a single audience member at a time, the most effective performance space appears to be the planetarium, as evidenced by Belson's successful Vortex Concert series in the late 1950's in San Francisco. The visual element, an entire hemisphere of projected image, more than fills the perceptive field of the audience members, while the use of multi-speaker arrays allows for much greater control of the apparent origination of the sound, giving the audiovisual composer an opportunity to create very closely fused audiovisual objects and a truly immersive audiovisual experience for their audience. However, despite Belson's early successes, the planetarium remains a venue that concentrates on educational audiovisual work, and is not generally an approachable venue for most audiovisual composers.

Inter-disciplinary collaborations with other creative practitioners also provides an opportunity for the audiovisual composer to create an immersive experience, which can explore the performance space or the relationship with other art practices. This can encourage further investigation and development of the composer's own practice, but does have the risk that the impact of the audiovisual work is overshadowed or downplayed within the overall experience.

In terms of creating a highly immersive experience for the individual, the imminent arrival on the technological market of Virtual Reality headsets such as *Oculus Rift* and *Project Morpheus* may also open up a new performance space for audiovisual compositions. These versions of VR allow a completely immersive world, where the individual is exposed to a completely controlled environment in both sound and vision. Initial audiovisual compositions in this field may concentrate on establishing a fixed medium, where the individual enters the experience and stays in a fixed position relative to the work being performed, but there may be a further development in this field in creating a truly interactive environment, where the individual user can move in, around, and interact with the audiovisual objects.

At present, the majority of my work has been composed for and performed in electroacoustic concert spaces and cinemas. However, I would argue that for creative development in the field, audiovisual composers should be pushing at the boundaries. It has already been established that audiovisual work is successful in planetariums, and it should be possible for audiovisual composers to work in a group to create a trial programme of works that can be performed in a planetarium to establish if modern techniques are just as convincing as Belson's film-based works. The educational nature of the planetarium programming may allow some experimental works to be included, while the more commercial priorities of IMAX cinemas, and the much higher production costs for the cinematic material, will probably make the IMAX venue prohibitively expensive for any experimental audiovisual work. The emerging VR technology may also become prohibitively expensive in terms of gaining rights to use proprietary hardware and software, but there is an early opportunity as the technology becomes established to gain entry into the field, and to work with the content creators for these platforms to allow some audiovisual experimentation to take place. This field in itself may prove to be the most rewarding and innovative for the audiovisual composer.

Chapter 5

Case Studies

Introduction

Michel Chion opens his seminal work on the correlation of image and sound in cinema, *Audio-Vision*, with an examination of the opening sequence of Ingmar Bergman's *Persona* (1966)[206]. He invites the reader to watch the sequence without sound, and suggests the scenes now lack unity and rhythm. He then suggests listening to the soundtrack of Jacques Tati's *Monsieur Hulot's Holiday* (1953) without the images, and notes that the implied space is very different than that actually shown visually. He refers to the *added value* that sound adds to image, how it enriches what we imagine we're actually seeing, but notes that this also sometimes make us underestimate the value of sound in the films we watch.

In terms of the audiovisual object, both the image and the sound add value to each other, reinforcing each other within the world the film is creating, captivating the audience's attention. This reinforces Chion's original exercises of watching films without sound, and, crucially, listening to films without image. In both cases, the missing element means the film loses the audience's interest and empathy – and even if we are experiencing a film that we haven't seen with both sound and vision before, we can still identify that there is a missing element. Films designed to include sound, as opposed to earlier 'silent' films, are decidedly incomplete without that sound. Nicholas Cook notes there is a "tide of film-critical and other narrative-based approaches which reduce the role of music to that of a mere supplement"[207]. This serves as a reminder when approaching an analysis or critique of audiovisual compositions that both perceptual streams are equally as important as the other, and that all discussions about the effectiveness or otherwise of any particular gesture, background or movement within the piece has to be considered with equal regard to the sound and visual components.

[206] Chion, M. (1994). *Audio-Vision: Sound on Screen*. Columbia University Press, New York. pp3-5.
[207] Cook, N. (1998). *Analysing Musical Multimedia.* Clarendon Press, Oxford. p vi.

In this chapter, I will examine three audiovisual compositions, summarising the audiovisual composition and elements in each, and providing an analysis based on the principles established earlier in the thesis. In selecting these case study compositions, the first and the last were chosen for very specific reasons. The first is one of my own compositions, which had been constructed towards the end of my PhD studies. This work incorporates many of the techniques and correspondences between audio and vision that I have discussed in the main body of the thesis. Given that the work was in part designed to test a great many of the principles of audiovisual composition I had explored in my research, it is particularly useful to critique the work with reference to those principles, to establish their success when put into practice. *Canto* (2015) illustrates several of the Gestalt principles used in creating audiovisual objects, as discussed in Chapter 2, as well as examples of both synchronisation and synchresis at work, allowing a comparison and analysis of their respective effectiveness to be made. The work is divided into three sections, each of which illustrate different approaches to the compositional process.

The second piece is *Autarkeia Aggregatum* (2005) by Dr. Bret Battey, based at De Montfort University. Dr. Battey's work has a rich texture in both the visual and audio fields, while his construction techniques involve audio and video generation using algorithms and plug-ins, which is a very different approach from mine. The video is driven from real images, and recognisable sounds appear to emerge over time, which allow for a discussion of Emergence as a compositional technique. This case study also illustrates that the principles discussed earlier in the thesis can also be applied when analysing work created by others. In this case, the work offers a good illustration of the Gestalt principles of Common Fate and Emergence.

When selecting the second case study for inclusion, I considered a wide field of existing compositions created by many audiovisual composers currently working in the field. Alternatives to *Autarkeia Aggregatum* that were considered included *Patah* (2010) by Dr. Diego Garro, based at Keele University; *White Noise* (2007) by Dr. Dennis Miller, based at the Northeastern University in Boston, USA; and *Breath of the Compassionate* (2009) by Bill Alves, a composer, video artist and writer who teaches at Harvey Mudd College in South Carolina. Each of these works combine audiovisual objects in complex, evolving relationships, but the general approach taken to audiovisual composition is similar to my own – the sound is composed first, generally using electroacoustic composition techniques, with the visuals then created to complete an audiovisual fusion between sound and image. As Dr. Battey's working methods involve generative techniques for both sound and vision, I judged that his very different approach from mine would offer more opportunity for exploring the principles of audiovisual composition.

The final piece is *See Hear (Outdoors)* (2015) by Amy Bruning, Tyler Carrigan, Pamela Gray, and Alice Pearse. This last piece is of interest as it was constructed by 1st year undergraduate Animation students following the principles on audiovisual composition I have outlined in earlier chapters. Although the students followed my normal compositional process, creating the soundtrack first and then animating it, their strong foundation in animation skills is evident in the final composition, with the students finding several very inventive combinations of audio and image to create an accomplished piece. The students' work outlines how my compositional principles and techniques can be used by others, with a predominantly visual approach compared to my background in sound composition, to create an immersive audiovisual work.

I also gave consideration to including other work made by audiovisual composers working within the research community in the UK at the current moment – having organised joint concerts with Dr. Andrew Hill (University of Greenwich), Dr. Andy Willy (Keele University), Dr. Andy Dolphin (University of Portsmouth) and Dr. Ben Ramsay (Bath Spa University), I am familiar with their work and compositional practices. However, as with the choice of *Autarkeia Aggregatum*, many of the compositional techniques exhibited similarities to either my own approach or that of Dr. Battey's, and I therefore decided to limit my case studies to these three works, allowing me to examine each to a reasonable extent.

Canto (2015)

Construction

Canto consists of three sections, loosely based on Dante Alighieri's *Divine Comedy*. The *Divine Comedy* is made up of 100 sections or cantos (a term for subsections found in epic poetry), which are divided equally between *Inferno*, *Purgatorio* and *Paradiso*. The title for the piece came from this term for the poem's sections. The inspiration for the piece came from personal experience of depression, where at its worst, I have felt completely despondent and in my own personal hell. The idea of moving from this state, through a gradual movement towards a happier, lighter state, recalls Dante's work to me, as reflecting a journey from a classical hell through to a perfect heaven, although that state is likely to always remain an ideal rather than an actual. Following Dante's schema, the 11 minute long piece is divided into 3 sections, *Inferno*, *Purgatory* and *Paradise*, with transitions between them. The audiovisual element proper consists of 10 minutes and 10 seconds, with the remaining time taken up by titles and credits. The duration of the audiovisual work of just over 10 minutes is towards the longest period I feel an abstract audiovisual work retains an effective immersive quality, based on my observations of the Soundings International Festival of Sonic Art at the University of Edinburgh. I curated two concerts of audiovisual compositions, and in both I observed that the audience grew restless in longer works, of approximately 15 minutes duration or more. From discussions with audience members after the concerts, their feedback was that the longer pieces became too onerous to follow with both senses of perception, and most closed their eyes to concentrate on the sounds instead.

My normal practice for all audiovisual composition is to create the sound element first. This is partly due to my connection of the audiovisual object to Schaeffer's ideas about the sound object, but also to counteract the normal human propensity to attach more importance to the visuals. This is noted by Spence and Soto-Franco, "the results of many studies have now shown that visual stimuli often (but not always) dominate (or strongly modulate) people's perceptions of multisensory objects or events"[208]. In my experience, I have found that creating the sound first allows a more equal balance between the two perceptual streams, achieving a more immersive quality to the work. However, even as I create the soundtrack, I also have the visuals forming in my head as a counterpart to the sound I'm hearing, so the process is less divided between sound and visual composition than it may sound - in many ways the animation stage is just completing the overall composition.

While the majority of the electroacoustic music I have composed in the past has been acousmatic, in that the source material has all been recorded from the real world and then manipulated to a greater or lesser extent, I had started to experiment with more synthetic material towards the end of the degree. For this piece, both *Inferno* and *Paradise* feature a large quantity of synthesized sounds along with manipulated real world recordings, while the *Purgatory* soundtrack is mainly acousmatic, with a lesser use of synthesis.

The sounds for *Inferno* rely initially quite heavily on synthesis. I had an image in my mind of a burning sphere, one that would erupt at the start, and continue to burn until guttering out, with various movements within and around the flames and surrounding space. This reflects the generally accepted Christian view of Hell as a world of punishment, a furnace in which the wicked are eternally burning[209]. As a result, I knew that I wanted to create a striking initial noise, such as a thunderclap, to give the impression of a large explosion, followed by an ongoing rumble to convey the continuous burning flames. I made use of various plug-ins available for the Cubase digital audio workstation software produced by Steinberg to create the initial explosion and continuing rumble, adapting preset values for thunder effects. I also added in additional phased low level rumbling throughout the section, as well as electrical hums and a recurrent choppy noise, reminiscent of a helicopter. These synthesized sounds are accompanied by two acousmatic sounds – the first being two

[208] Spence, C. & Soto-Faraco, S. (2010). *Auditory perception: Interactions with vision*. In D. R. Moore, P. A. Fuchs, C. Plack, A. Rees & A. R. Palmer (eds.), *Oxford Handbook of Auditory Science: Hearing.* (2010), Oxford, UK: Oxford University Press. p273.

[209] Catechism of the Roman Catholic Church, Article 12 sections 1033-1037. Available at http://www.vatican.va/archive/ccc_css/archive/catechism/p123a12.htm [accessed 14/07/2015].

groups of ‘screams’ formed by convoluting the sound of air escaping from a balloon’s tightly stretched neck with a recording of my own screams, while the second sound forms the climax of the first section. This is a manipulated recoding of a train passing by at a railway station platform, where the commuter conversations and the oncoming trains have been made metallic. The gradual increase in volume peaks as the train passes by, and then trails off. Additional points of sonic interest are created by the ‘sparking’ sounds where the train wheels scrape against the tracks.

Visually, I wanted the images in this section to reflect the flames, electrical hums, phased sounds and general unease conveyed in the soundtrack. Dante envisages Hell as a series of circles, leading down to Satan at the very centre. The individual circles hold punishments for specific sins – in the second circle, lustful sinners are souls blown about in a violent storm, in the sixth, heretics are trapped in tombs, forever burning. While many sins are punished with specific tortures, such as being encased in mud or faeces, one of the most prevalent forms of torture is fire, from the river *Phlegethon* which is made of boiling blood and fire, to those guilty of simony being inserted head first into holes in a rock with their feet set on fire. However, the very lowest levels of hell are frozen, with Mordred, Judas Iscariot and Satan himself encased in ice. I therefore decided that the central image would be a sphere, cracked across the surface to reflect the mud and other encasing materials in hell, which would initially be on fire but would gradually gutter out to an icy, dark bleakness, in keeping with the final layers of the Inferno. This sphere hangs in a void held within a greater cavern, also cracked and lifeless. I also wanted to reflect the lost souls buffeted in the storms, and the electrical storms and sounds included in the soundtrack. However, I also wanted to keep the overall visual form quite abstract, to hint at the idea of hell, but without a concrete representation of a classical underworld. Therefore the souls were conveyed as wisps of white vapour circling the central sphere, while the phases of electrical sounds were envisaged as waving plasma-like tentacles growing from the sphere. As the piece continues, I wanted the surface fire to die down, but the underlying heat and threat to continue, and reflected this by creating a dull red/orange glow in the cracks in the sphere and surrounding cavern. Finally, as the flames die off completely and the cracks show only a sullen glow, there is one last rallying flicker of the underlying lava-like glow as the deformed train noise rushes past. The section finishes with the screen dark and empty of all form. This reflects some earlier work I had created using some of the colour principles proposed by Wassily Kandinsky in his work *Concerning the Spiritual in Art*, specifically his observations that black "is something burnt out, like the ashes of a funeral pyre, something motionless like a corpse"[210]. Black is the void, empty of all possibilities and signifying a complete lack of creation – a suitable colour state to associate with the end of Hell, and also recalling the idea of a 'Stygian gloom', again referencing Dante's journey across the Styx and its marshes surrounding the walls of Dis, the city at the centre of the Underworld.

The majority of the *Inferno* visual sequence was created using Maya, a 3D CGI software package. The central sphere, surrounding cavern and specific effects such as the circling vaporous screams, the growing electrical tendrils, the three lightning bolts, the initial explosion and the continuing flames, were all created using this package, with the mud-like cracked textures imported and applied to the surfaces on screen. The glowing effects for the lava-like cracks as well as the phased pale flames corresponding to phased movements in the soundtrack were all created in AfterEffects.

The idea of Purgatory to me recalls a state of slow progression, a slow movement towards Paradise as the soul is purified. This is in contrast to the Presbyterian faith in which I was brought up (there is no concept of Purgatory in Presbyterianism)[211]. The Roman Catholic church regards Purgatory as a place of cleansing fire, where the prayers of the faithful will allow the impure to be purified of their sins and allow them to enter Paradise[212]. This does not necessarily imply movement of any sort, but within Dante's *Purgatorio*, as Dante ascends the terraces of Purgatory, his sins in each category are washed from him. This implies a progress based on movement,

---

210 Kandinsky, W. (1977). *Concerning the Spiritual in Art*. Dover Publications Inc. New York. p39.

211 Presbyterian Mission Agency, discussion on what Presbyterians do not believe. Available at http://www.presbyterianmission.org/ministries/today/dont-believe/ [accessed 14/07/2015].

212 Catechism of the Roman Catholic Church, Article 12 sections 1030-1032. Available at http://www.vatican.va/archive/ccc_css/archive/catechism/p123a12.htm [accessed 14/07/2015].

while Dante's *Inferno* has sinners held permanently in the appropriate hell for their sins. Therefore, I decided to include sounds in this section which recall movement and travel forward, a journey from despair to the promise of something better. As I grew up in Northern Ireland, the journeys that held excitement and promise to me always involved travelling across water - to Scotland, England, the Isle of Man - so that felt to be the right medium for that idea of travelling onwards to something new. Dante's *Purgatorio* is located on an island, so again the idea of water on all sides seemed appropriate.

The sounds in this section are initially synthesized, starting with a high pitched sustained vibrato note signifying the transition from *Inferno* to *Purgatory*. The section then continues with a growing phased set of synthesized sounds, providing an underlying, pulsing sense of movement, but still with a slightly discordant, minor chord pattern to indicate that the environment is still uncomfortable. The continuing phased sounds here recalled the ocean to me, the swell of long rolling waves, which is the ideal of movement and motion I wanted to achieve. Against that, I wanted to include another sense of imbalance, of background threat, which has been incorporated in the form of a processed recording of a seabird colony, where the many seagull and other bird calls run into each other and are slightly distorted. This keeps the feeling of the sea and ocean from the associated wildlife, but the processes applied hide the source material enough to also suggest that the sounds might be a crowd of human conversations, twisted and slightly menacing. Against these mainly background noises, I decided to include a slightly more harmonious foreground sound, which I sourced from a recording of attendants singing in the highly resonant Baptistery of St. John, in the Piazza dei Miracoli in Pisa, Italy. This recording was reversed and several transformations applied to remove some of the background sounds of the tourist audience present, and to disguise the more recognisably human attributes of the source. The use of this material allows a nod to the religious source of the ideas behind the composition, but also provides a foreground sound which is more appealing and attractive to the listener, as a counterpoint to the more disturbing background sounds. The final sound included in this section is the sound of a small stream trickling over stones as it runs down the beach at Gullane Links into the sea.

This retained the overall sense of water and motion I wanted to keep in the section, while the trickling sounds established that the sound is set where water passes over land, indicating that a destination will be reached imminently.

Visually, the idea of water and the sea were always the most important to convey in this section. As the previous section ended in a dull black screen, this section also starts off in darker colours, mainly blues and greens to convey the idea of being underwater, far out to sea. The moving water-like shifts on screen were created using an ink particle effect in Maya. Three versions of this effect were created to last over a long frame sequence, each requiring over 3 days of processing to render fully. These were then overlaid and reversed in AfterEffects, giving the more saturated, deeper colours at the start of the section, but eventually culminating in the initial ink 'drop' coalescing and disappearing up and off the screen towards the end of the section. One of the sequences was duplicated and a layer effect applied, resulting in a dappled effect, which for me reflected and reinforced the processed seabird colony sounds. A cyclic colour transition was also applied to two of the ink drop sequences, matching the phasing in the underlying background sounds, cycling the dominant colours from blue to green. I also wanted to include an element that would look very organic and still a bit unsettling, so experimented with the lightning effects in AfterEffects to create a set of interlinked strands, resembling seaweed. I varied the intensity and thickness of the strands to correlate to the manipulated sounds of singing in the Pisa Baptistery. I also adjusted the parameters for these strands so that they would move slowly, occasionally jerk around, and change colour in line with background changes in the moving ink-clouds. The reversed ink-cloud sequences were carefully calculated so that the disappearance of the ink material up and off the top of the screen would coincide with the climax of the rising stream trickling sound, leaving the screen white and empty – this also corresponds with Kandinsky's view that the colour white "is not a dead silence, but one pregnant with possibilities. White has the appeal of the nothingness that is before birth, of the world in the ice age"[213]. The use of white here as the transition colour reflects this idea that it is a positive space, a

[213] Kandinsky, W. (1977). *Concerning the Spiritual in Art*. Dover Publications Inc. New York. p39.

blank page that can be filled in with the creativity of Paradise. Robert Rauschenberg tapped into this idea of the purity of white, the essence of creativity in his series of White Paintings, where his aim "was to create a painting that looked untouched by human hands, as though it had simply arrived in the world fully formed and absolutely pure"[214].

The final section representing Paradise was more rooted in reality than the others, in that I felt it would be almost impossible to create a virtual paradise that can better an earthly one. In this, I have almost completely deviated from Dante's vision of *Paradiso*. For me, particularly in comparison with the bleak moods induced by depression, Paradise is a feeling of calm, a lack of stress, and an ability to enjoy the world around me. Therefore this section starts with one very simple, restful recording, with only minimal processing to remove wind noise. This recording is of woodland beside the Scottish seaside Loch Nevis, on the Knoydart peninsula. This particular area is only accessible by sea, as there are no major roads in the area, and none that connect with the main road network. The recording features a recurrent call by a willow warbler, with the gentle sounds of trees and a slight hint of the sea in the background, recalling the water sounds prevalent in *Purgatory*. In addition, at one fortuitous moment a bee flew around and past the microphone, adding a further foreground point of interest to the sound. While this recording places the *Paradise* section in a restful scene of reality, I also added in a gradually expanding set of glassy synthesized chords, recalling the music often used in film to signify a motion into heaven or a heavenly realm. This was built around the song of the warbler, and initially started in a gentle minor key to recall the initial discomforts of *Inferno* and *Purgatory*. While the synthesized chords start off as a very low, gentle addition, I used the humming of the bee as a marker to introduce more defined textures and tones in the sequence, leading gradually to a climax and move to a major chord, resolving the small underlying tension in the piece and establishing a final sense of completeness and stability. At this point, the recorded sound of the forest and

[214] San Francisco Museum of Modern Art, notes on their Rauschenberg Collection. Available at http://www.sfmoma.org/explore/collection/artwork/25855 [accessed 25/08/2015].

birdsong also fades out, leaving just the glassy chords to gradually reduce down to a single sustained, slightly phased note, that then fades out to silence.

Visually, I decided to make use of a real location as an ideal of a personal Paradise rather than create an abstract set of images. With my current adopted home being in the UNESCO World Heritage Site of Edinburgh, I had a wide variety of locations that I could use, although the soundtrack of gentle wind, tree noises and birdsong predisposed me to find a location offering a very green, open or forested location. I decided that I would use a timelapse sequence for the entire section, starting with a blue sky with clouds passing across it, then a gradual tilt down to reveal an ‘earthly paradise’, ideally with a view of trees and Edinburgh itself. The choice of location was suggested by a colleague working in film production, and turned out to be ideal for the timelapse photography I had in mind. On the day I carried out the photography, the weather also co-operated, with just the right amount of cloud cover to give visual interest, but still allowing a good blue sky to dominate. This initial cloud sequence recalls the idea that heaven is in the skies, while also offering a bridge from the abstract shapes of purgatory into a real world setting. The later part of the sequence revealing Salisbury Crags and Arthur’s Seat on the right bed the scene in ancient natural history, the Castle on the left gives a background of human antiquity, while the tree waving in the foreground is a living history, a connection with the idea that nature itself can be seen as a paradise, a stability in the continuance of growth and the seasons.

The timelapse sequence posed several challenges, particularly in achieving a tilt from the sky to gradually reveal the central tree and wider landscape. Ideally, I would have used a motorised tilt head to create a smooth, even transition, but the cost and logistical problems in providing power to this sort of tilt head in an outside location made this option untenable. I had determined that I wanted to start the tilt at the point where the sound of the bee appeared in the soundtrack, and to finish at approximately 60 seconds of screen time later. As I was working on an interval duration of 1 second between capturing frames, I calculated the point at which I wanted to start the tilt, allowing a few extra seconds for the transition between *Purgatory* and *Paradise*, then also the end point of the tilt sequence. This meant that I would need to spend over 1 hour and ten minutes to capture the timelapse sequence, with the tilt starting at 25 minutes and ending at 42.5 minutes. Given the lack of a motorised tilt facility, I tilted the camera down incrementally by a few degrees every 100 seconds, and 'stitched' the tilt together using AfterEffects. The resulting tilt includes a few noticeable jerks as the camera moves from the last frame of one sequence with the head at a specific angle to the starting frame of the next sequence, but these are acceptable within the composition and the slightly jerky motions on screen as are typical of timelapse photography.

The initial cloud movements worked well as a transition from the relatively organic abstract shapes used in the purgatory sequence, moving from a created world to the real world. The gradual revelation of the tree and background through the tilt sequence was serendipitously completed by a passing aeroplane contrail passing over the tree just as the tilt finished. A gradual increase in green and yellow saturation as the tree becomes dominant also reinforces the hyper-reality of the timelapse scene and makes the foliage more visually appealing. The scene ends as the music coalesces to the quieter closing chords, with the tree and background fading into white through an edge-detect effect which briefly outlines the objects in vision.

As with earlier uses of white and black in line with Kandinsky's theories of their significance, the title and credits also reflect the idea of black as a dead state, while white is the potential for creation. This is of course based in the Western associations of colour with specific meanings – in Asian traditions, while white can be associated with purity (brides in Japan wear white with other colours, predominantly red for luck and prosperity), it is also associated with mourning in many parts of China. The opening title is predominantly black, with only the name of the piece showing in white, while the end credit sequences are black text on white backgrounds. This establishes the start as being from a dead, hopeless state, hell or inferno, to the final state of creative potentiality, or paradise.

Analysis with reference to Gestalt principles

In Chapter 2, the Gestalt principles of perception are described, with brief examples of how they are usually applied to visual and auditory streams of perception. *Canto* offers examples of audiovisual objects which illustrate how these Gestalt and associated principles can be applied to audiovisual objects. The following analysis refers to specific audiovisual objects and sequences in *Canto* that correspond to audiovisual perception. Note that some of the Gestalt principles described in Chapter 2, such as Good Continuation, Closure, Multi-Stability and Invariance, are not illustrated with examples from *Canto*, as there are no good clear examples of these principles shown in the piece. This illustrates the point that it is not necessary to try and make use of all the principles while constructing a piece – it is only those principles which support and develop the composition as a whole aesthetically satisfying work that need to be employed by the composer – different works will require and illustrate different principles and compositional techniques.

Proximity

In the classical approach to Gestalt principles in phenomenology, the principle of proximity refers to groups of objects belonging together when they can be perceived as distinct groups. Similarly, in audio terms, sound objects in close proximity, such as repeated patterns, are perceived as a group belonging together. For audiovisual objects, there is an initial principle of proximity in terms of close synchronisation – when audio and visual objects are perceived as occurring at the same time, they are linked together in our perceptions, a strong link which is at the core of constructing the audiovisual object.

The opening sequence of *Canto* illustrates the concept of proximity in terms of close synchronisation of sound and vision. The central sphere erupts into a flaming explosion at 0:20, which rapidly diminishes into a central fiery ball, at the same time as a large thunderclap bursts, diminishing into an ongoing thunderous rumble. As an abstract image, there is no definite sense of scale on screen to set the size of the central sphere, but the impression created by the thunderclap and ongoing rumble is of great size, an immensity hanging over the viewer. The audiovisual composer, Robert Robertson, on watching this sequence observed, “At the beginning I had the impression of a sphere on a very large scale, which was also expressed in your music. But you have sound being heard at the same time as the explosion at the beginning, and also later the pulsating cracks in the sphere appear to be heard simultaneously. If you had a delay between the visual event and the sound it produces, would one’s sense of scale be augmented, or would it just sound as if it’s a mistake?”[215] In reply to Robert, I noted that I had experimented with a slight delay between the visual and audio in the past, but that it always seemed to diminish the overall effect, and to disturb that immersive quality and synchronisation of audio and visual objects together forming a cohesive audiovisual object. This may reflect the cinematic ‘training’ we have received where explosions are seen and heard together to heighten dramatic tension, even though in reality there would normally be a discernible delay between vision and sound unless very close to the source.

Robert also pointed out the close synchronisation between the pulsing light in the cracks on the sphere and a pulsing low rhythmical sound (2:20 – 3:25). This also illustrates proximity in terms of a repeated pattern of audiovisual objects – the recurring visual glow with the recurring phased sound cements the perception that there is an identifiable phenomenon taking place within the cracks, a cyclical surge of magma with associated rumble.

[215] Personal email conversation with Robert Robertson, 06/07/2015. Composer’s website available at http://ocatilloaudiovisual.weebly.com/ [accessed 22/07/2015].

Similarity

The principle of similarity states that we perceive objects as groups where there is a specific similarity between particular objects, even though all other attributes may mirror other objects in a continuous perceptual stream. In terms of *Canto*, there are two very similar sound object sequences during the *Inferno* section where there are low phased rumbles (2:20 – 4:14). When paired with visual objects to create audiovisual object sequences, however, the perceptible differences in the visual stream create specific similarities that allow the two *audiovisual* objects to be perceived separately as the recurring glowing cracks on the sphere surface, and the recurring glowing cracks in the surrounding cavern or 'shell'.

There is another linkage of colour shifts with phased sounds for both the flames in *Inferno* (1:52 – 2:35), where they fade from a normal red/yellow colour to a faded white to match a low 'sweeping' sound; and in *Purgatory* (4:37 – 5:45), where the dominant hues of the shifting 'wave' patterns phase from blue to cyan and green, matching the shifting patterns in the mid-tone sounds. However, these shifts are less a realisation of similarity than another example of proximity – there are no similar objects from which these audiovisual patterns need to be differentiated, so the boundaries of the sound and colour shifts here only delineate the objects themselves, rather than highlighting their difference from another similar but distinct audiovisual object.

Common Fate

In Chapter 2 I noted that the principle of common fate refers to objects moving together, forming an impression that there is a common fate binding them together, illustrated visually by a flock of starlings, and in auditory terms by a group of sound objects changing pitch in the same direction and speed. In audiovisual terms, the principle of Common Fate can be observed in the *Purgatory* section of *Canto*, in terms of the movement of the ink drop sequences and particularly at the end, where the individual sequences combine visually to consolidate, retract into themselves and disappear up and off the top of the screen at the same time as the underlying sound sequence also starts to rise in pitch, with this pitch rise gradually increasing in pace until it also rises beyond normal human auditory perception into ultrasonic ranges (7:27 – 7:44). The two ‘rising’ phenomena combine to create an audiovisual object encompassing common fate, as the shared perceptual ‘rise’ ties the two perceptual streams together.

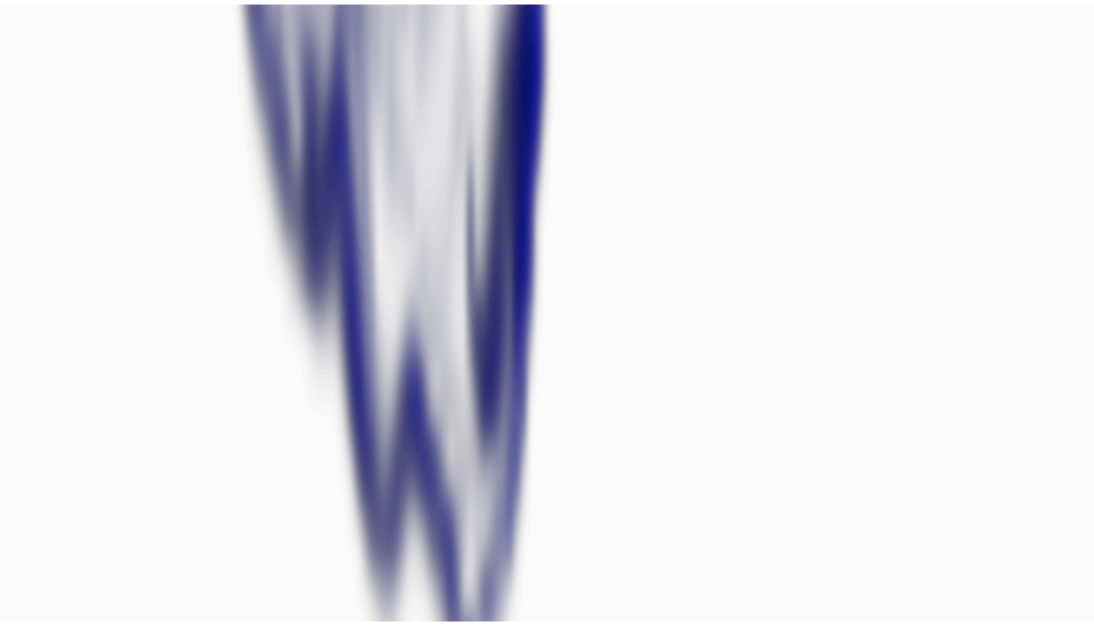


Figure 1: Ink-drops rising to disappear off the top of the screen in the *Purgatory* section of *Canto*

To a lesser extent, the closing sequence of *Inferno* also exhibits characteristics of Common Fate, where the manipulated recording of a train passing is synchronised with a swell of glowing red in the central sphere cracks, rising to an oversaturated red flash as the train noise reaches its climax with the train passing by (4:17 – 4:20). In this case, the similarity between the sound and vision is not reflected in motion on screen, rather than by the increasing colour intensity followed by a rapid decline as the train noise dissipates. This illustrates that common fate need not be restricted to pitch or visual movements, an implied common motion, growth or decrease can all contribute to the cohesion of the constructed audiovisual object.

Surroundedness / Belongingness (figure against background)

The principle of surroundedness states that, if all else is equal, a smaller area will be seen as a figure against a larger area, which will be perceived as the background. This occurs visually throughout *Canto* – in *Inferno*, the central sphere is a coherent object that 'hangs' in a surrounding cavernous 'shell', in *Purgatory*, the waving tendrils that appear as the piece develops are a foreground object against the dappled background and moving 'waves'. In *Paradise*, the use of real world objects also establishes this foreground / background difference, between the tree in the foreground and the cityscape behind it. The concept of exclusive allocation is in operation here – the boundary between the foreground object and its background is perceived to define and belong to the foreground, highlighting the object. Similarly, in the audio perceptual stream, there are sound objects that appear as foreground objects against the background of longer, more diffuse background sound objects.

Within *Inferno*, the combination of foreground visual and sound objects fuse to create the audiovisual objects of the 'screaming ghosts', wisps of vapour that move around, in front and behind the central sphere, combined with the sound object of recurring high-pitched screams (0:25 – 0:53, 1:55 – 2:23). Both the audio and visual aspects of these objects exhibit this property of 'surroundedness', as the vapour trails are evidently separate from the background, and by moving around the central sphere, establish not only are they separate to other objects in vision, they also emphasize the sphere itself as a separate foreground object. The screams are sonically also in the foreground, as they are markedly different from the lower pitched continuous rumbles that correlate to the flames and other audiovisual objects.

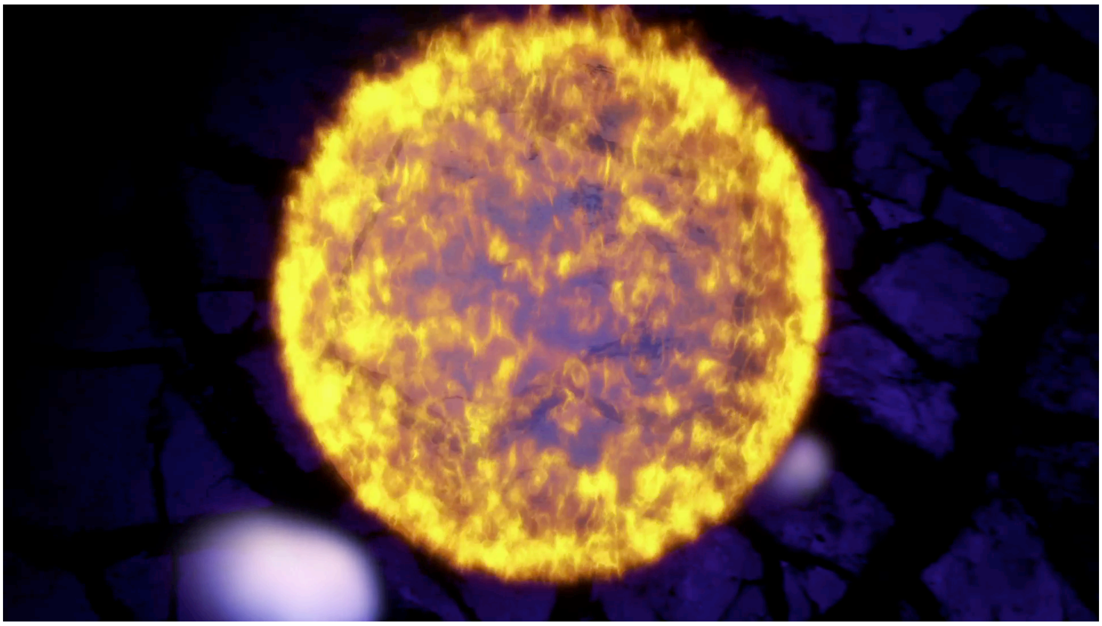


Figure 2: “Ghost” vapour wisps moving around the central sphere in the *Inferno* section of *Canto*

Another example of a foreground audiovisual object is the waving tendrils in the *Purgatory* sequence, which are definitely at the forefront of the visual layers, with the dappled background behind. The sound corresponding to the tendrils is the long, processed singing voices from the Pisa baptistery, which again are in the foreground over the long, phased wave and seagull sounds.

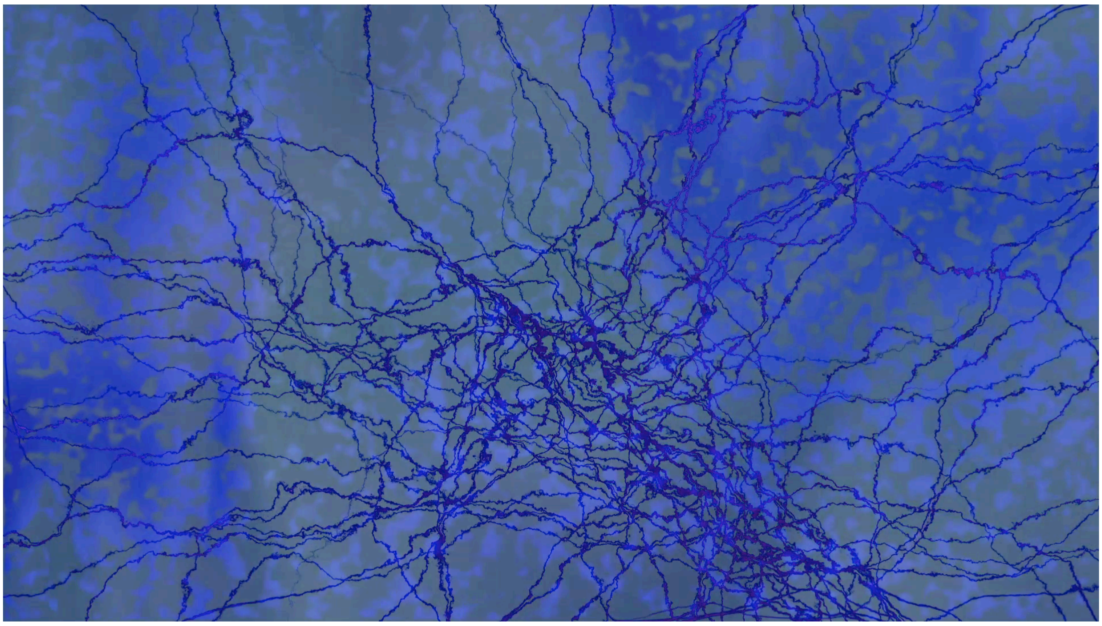


Figure 3: Tendrils in the foreground against a dappled background in the *Purgatory* section of *Canto*

Emergence

One of the additional Gestalt principles, emergence, is similar to surroundedness in that a foreground object is discernible against a background. In this case, however, the object coalesces from items in the perceptual field, rather than being immediately obvious. In audiovisual terms, the idea of emergence is most easily seen when a foreground object develops from a background, gradually becoming obvious by sound and visual objects becoming identifiable from an underlying ground. The majority of audiovisual objects in *Canto* are well-defined from the start, and although some objects fade in and out, such as the ghost vapour objects, they are superimposed on the background rather than developing naturally from within it. The only audiovisual object that really shows an element of emergence is the extrusion of waving hairs that recall the flickering tendrils of plasma globes, which are fused with an electrical hum, in the *Inferno* sequence (0:53 – 1:39). Following closely on from the first set of vaporous ghosts, this audiovisual object emerges initially in the audio perceptual stream as a growing electrical hum. The visual portion emerges more gradually, with anchor points slowly becoming established on the surface of the sphere, obscured to a certain extent by the continuous flickering flames covering the sphere surface. As the audiovisual object develops, 'hairs' extrude from these anchor points, growing out from the central sphere and glowing in a neon blue/purple colour, reminiscent of the glowing strands that form when a hand is pressed against a plasma globe. As this object develops, the plasma hairs begin to retract and return to the anchor points, while the electrical hum persists – it continues for another 20 seconds or so past the point where the hairs have disappeared, marking a long disappearance to emphasize the emergence and disappearance of the visual aspect of the object.

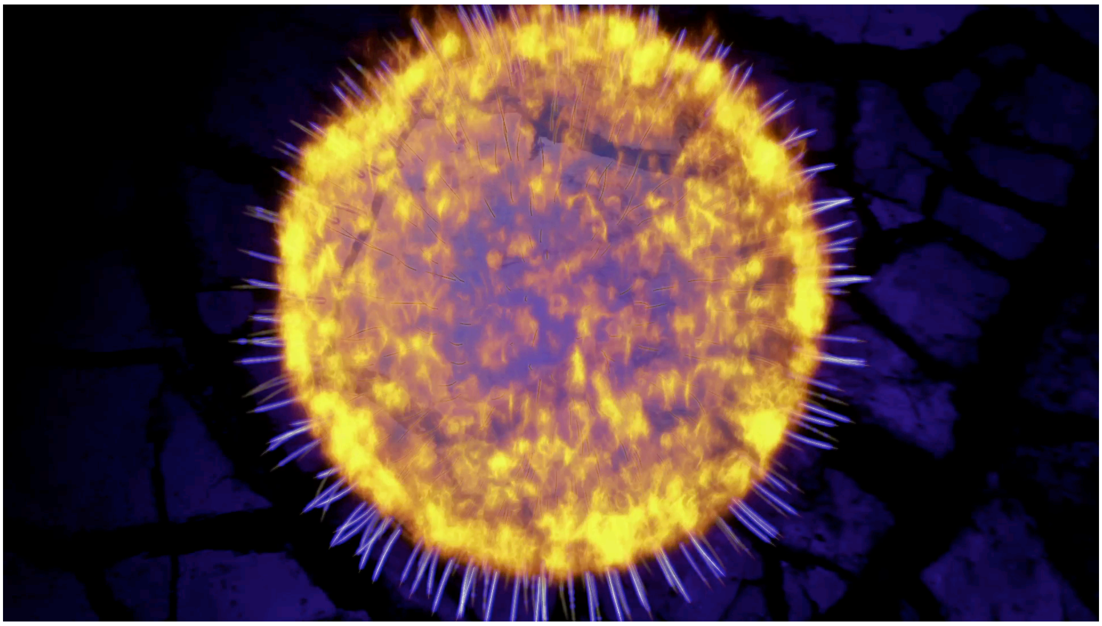


Figure 4: 'Plasma hairs' emerging from the central sphere in the *Inferno* section of *Canto*

Re-ification (inference from elements in view)

The principle of re-ification allows for the tendency in human perception to infer the whole structure of an object based only on those elements in view – such as recognising that a fence continues and exists behind a parked car, even though only the potions that are not obscured by the car are in view. In *Canto*, the position of the central sphere in the *Inferno* section is immediately established, and the entirety of the visual hemisphere we would expect to see is in view, with shading and lighting effects that convey that we are looking at a curved spherical object rather than just a flat circular plane. This is subsequently reinforced by the movement of the ghost vapours around and behind the sphere, disappearing behind its rim but appearing again almost immediately, which reinforces the impression that the visible circle in the centre of the screen is definitely a sphere. This re-ification of the vapour trails as they move behind and out from the back of the central sphere illustrates how this principle of re-ification can be used to establish the 3 dimensional properties of the objects on view in the composition.

Synchronisation and Synchresis

In Chapter 2, I describe the phenomenon of *synchresis*, as originally conceived by Michel Chion, a "spontaneous and irresistible weld produced between a particular auditory phenomenon and visual phenomenon when they occur at the same time"[216]. In practice, if you place any soundtrack against any set of moving visuals, the human mind will attempt to create linkages between the two. This can be successful to a greater or lesser extent, but in many cases is an effect actively sought for by the sound designer when creating soundtracks to go with edited film footage. In contrast, simple synchronisation occurs when a sound is placed exactly against a visual object, forming a distinct perceptual bond between the two perceptual streams. In excess, this can create a very artificial atmosphere, often described in animation terms as 'mickey-mousing'. Audiovisual compositions consist of many synchronisations and synchreses, with skill, art and creativity dictating which technique is most appropriate at every stage of composition.

In *Canto*, I deliberately designed the piece to make use of both synchronisation and

[216] Chion, M. (1994). *Audio-Vision: Sound on Screen*. Columbia University Press, New York. p63.

synchresis, but with an emphasis on synchronisation at the start, moving on to almost pure synchresis at the end. The opening section of *Inferno* consists of audiovisual objects that are mainly created through synchronisation of sound and image, such as the initial explosion, the ghost vapours, and the glowing cracks. There is relatively little use made of synchresis, apart from audiovisual objects such as the flames covering the sphere, lasting almost the complete duration of the section. However, the *Purgatory* section makes use of both synchronisation and synchresis. There is a direct synchronisation of image and sound from the start in terms of the phased sound and the colour changes on screen for the wave patterns created in the ink drop sequences, with the lighter green colours corresponding to the higher pitches in the phased sound. The sounds of the seabird colony, however, are not directly synchronised with the dappled background effect, the correlation here is much less defined. Similarly, although the waving tendrils fade in and out of view with the increase and decrease in volume of the Baptistery vocals, there is no deliberate synchronisation of the tendril movement or snapping with movements in the soundtrack. The final *Paradise* section is almost completely reliant on synchresis – there are no deliberate points of synchronisation, although some of the prominent passages in the soundtrack are used to mark the start and end of the tilt sequence.

When watching *Canto*, the move from dominant synchronisation to dominant synchresis is gradual across the three sections, and so does not become immediately apparent. However, my own impression is that the middle section achieves the best balance between visuals and audio in creating an immersive audiovisual environment, with fewer but more aesthetically appealing audiovisual objects. The first section contains more audiovisual material and variation, but seems to be less immediately immersive, while the final section, although achieving the goal of a paradisical setting, appears to me to convey a historical moment in time rather than an immersive environment. As an experiment in trialling different levels of synchresis and synchronisation in audiovisual work, I regard the outcome as supportive to my personal compositional view that synchresis affords the most immersive bonds between audio and vision, but an element of synchronisation is also necessary to provide a base against which the strength of the synchresis can be measured.

Autarkeia Aggregatum (2005)

Construction

Autarkeia Aggregatum was composed by Dr. Bret Battey in 2005, making use of music synthesis programming languages and bespoke packages in combination with visuals created using the visual programming language *Processing* and specific plug-in designed for the video effects package *Apple Motion 2*[217]. The audio and video packages were designed by Dr. Battey specifically to achieve the audio and visual effects he intended to use in this composition. The full composition can be viewed on Dr. Battey's website[218].

Dr. Battey's approach to audiovisual composition is markedly different to mine in that his audio and visuals are both created through synthesis, as opposed to the mainly processed sound and visuals I use in my compositional process. This emphasizes that there is no correct or singular approach to achieving immersive audiovisual work, although the principles observed throughout this thesis can also be applied in analysing Dr. Battey's work. By its very nature of composition, the work is highly synchronised throughout, as the raw material feeding the synthesis of both the audio and the visuals is the same – the audio signals produced by the sound synthesis processes are used to drive the synthesis of visual points, which change and move around each other in almost Brownian motion, but with a very definite relationship driven by the soundtrack.

The soundtrack to *Autarkeia Aggregatum* is rich and complex, immediately establishing a distinct audio environment. Despite being synthesized, the sounds are tantalisingly recognisable – the sound of a wind instrument, the faint chanting of

[217] http://www.mti.dmu.ac.uk/~bbattey/Gallery/autark.html [accessed 26/07/2015].
[218] *Ibid*.

voices, a zither-like sound, which may be coincidental features, but which influence the audience to add an interpretation that the sounds may emanate from a middle-eastern religious setting. The audio is continuous and constantly developing, with sounds arising as foreground objects, panning across the stereo channels, and disappearing into the rich textural background. At the same time, the visuals consist of swirling, complex animations, which occasionally form patterns reminiscent of mandala forms or the patterns seen in middle-eastern art and textiles. There are almost but not quite discernible bases to the visuals, with underlying shapes and colours behind the ever present swirling points and lines. The motion of the dots is extremely complex in any given part of the work, but there are group motions and colour changes which correlate to pitch and panning movements in the audio stream. The close relationship between sound and vision is particularly evident in the speed of particle movement, which seems to be keyed to the audio amplitude – at the few points where the soundtrack dwindles to almost silence, the swirling motions slow down, almost stopping, before starting up again and building speed in line with the increased volume.

Gestalt principles

Common Fate

This work exhibits several Gestalt principles, but the two that are most prevalent are Common Fate and Emergence. As noted above, the principle of Common Fate refers to objects moving together, forming an impression that there is a common fate binding them together. In *Autarkeia Aggregatum*, there are many instances of Common Fate. The intricate swirling motion of the particles move in sympathy with the panned movement of foreground sound objects in the soundtrack, creating an instantaneous impression of common group movement of the particle points, or of the changes in colour they exhibit, in sympathy with the audio movements. While this occurs throughout the work, there is a particularly rich sequence from 2:43 – 3:25, where metallic surges in the soundtrack are mirrored by wide circular sweeps of the particles across the screen, and the movement of pale coloured segments sweeping through the particles, again corresponding to foreground audio objects.

There are also correlations drawn between the timbre of the foreground sound object and the correlating visual complement – the brassier, horn-like sounds tend to be associated with red, brown and golden colours sweeping through the particles, while more metallic, glassy surges are pale blue and white. This is not set in stone, but is a prevalent trend throughout the work. This accentuates the common fate of the particles, as the colour movements create tightly knit groups, again reminiscent of sweeping flocks of birds or clouds of insects.

Emergence

Autarkeia Aggregatum also exhibits a constant continual emergence of audiovisual objects, as the audio changes drive the motion and movement of the swirling particles. The individual pitches and timbres within the soundtrack naturally emerge, increase in amplitude and dominance, then die away within the surrounding background noises, supplanted by other elements rising to the foreground in their turn. The section running from 4:36 to just after 6:00 illustrates this emergence well. After a period of panning glassy foreground notes leading up to the section, the music has diminished to a quieter mood, with a corresponding slowing and dimming of the swirling particles. At 4:36, the horn-like tone established earlier in the piece returns, and is matched by interlacing cables of particles moving across the screen. The main horn tone phrases are interspersed with a metallic hum, at which point the cables break down slightly into wider swirling patterns, which are then tightened up again into the cable structures as the horn re-establishes its dominance. This pattern of emergence and dissipation continues up to 6:00, where the horn tone is superseded by a rapid crescendo into a louder lower held note and higher fluttering tones, accompanied by a sweeping wash of lighter blue particles across the screen.

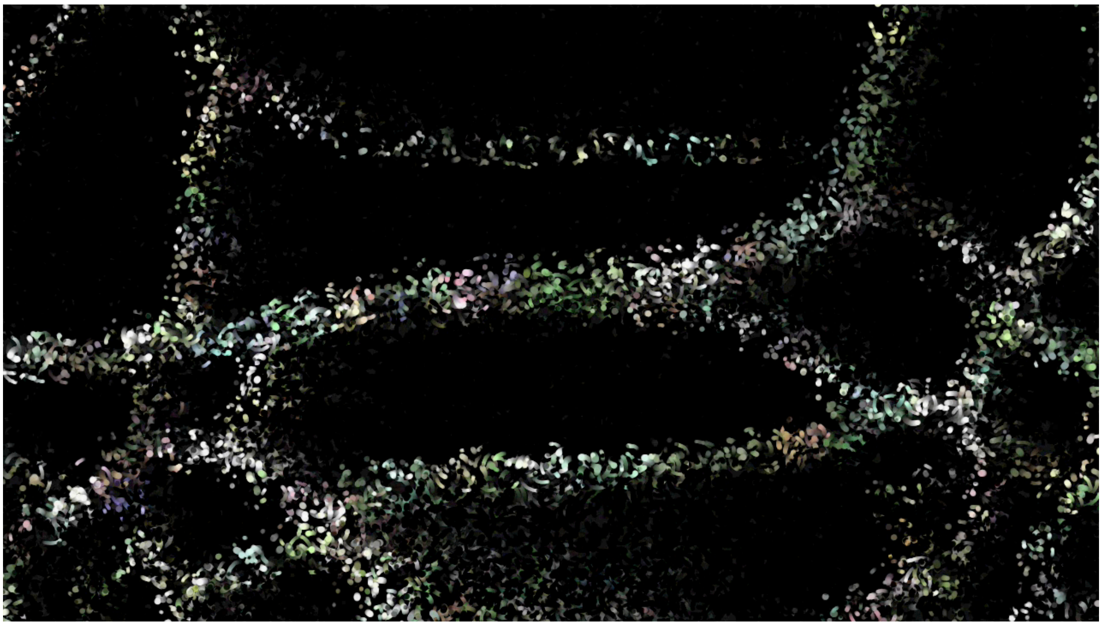


Figure 5: Cable-like confluences of particles matching horn tones at 5:01 in *Autarkeia Aggregatum*[219]

There is a further particular example of effective emergence featured in the work, but it is particularly effective due to the expectation that has been built up in the piece leading to a climax, with a specific outcome that the preceding passages have been building towards.

Expectation

Throughout *Autarkeia Aggregatum*, the soundtrack has a richness in tone, with a recurrent unusual horn tone and zither-like sounds, recalling Middle East instruments and a tone associated with religious rituals. This is added to with the glassy, metallic shifts in tone, similar to the use of glassy and 'heavenly choir' timbres in the *Paradise* section of *Canto*. In a similar fashion, the shifting particles seem to overlay or slightly obscure movements and colours behind them, ranging from underwater scenes to intricate colour weaves, again reminiscent of middle eastern carpets and religious patterns.

[219] Reproduced with permission from Dr. Bret Battey

Then at 8’07”, the sounds shift from abstract, strongly defined synthetic combinations, to a recorded, actual soundscape, with identifiable chanting and birdsong. At the same time, the intricate, swirling visuals twist and settle into a subtly shifting scene, reminiscent of worshippers gathering in an early morning square, or light shining across a pavement.

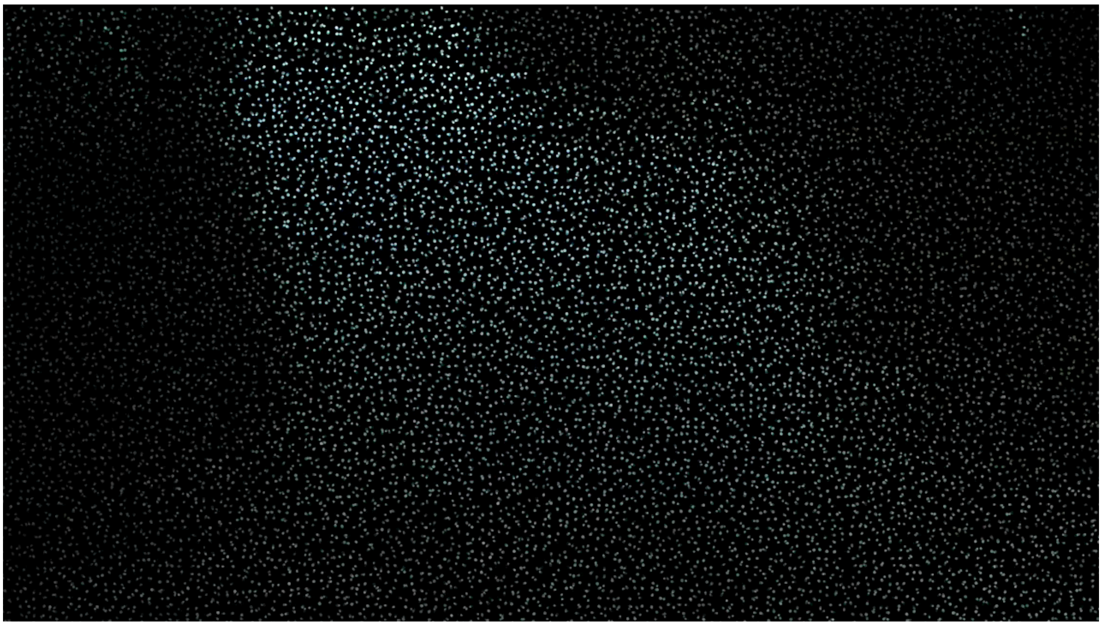


Figure 6: Climax of piece, recalling a holy space and associated sounds in *Autarkeia Aggregatum*[220]

This revelation of a concrete audiovisual landscape, an identifiable experience, if mainly defined in sound rather than image, resolves all the expectations that have been built up, if only sub-consciously, throughout the piece up until this point. The quasi-religious tones, the rich interweaving colours, all are resolved into this moment of calm contemplation and religious observance. A persistent held note continues throughout this section, maintaining a link to the earlier pitched tones, but also adding another mystical tone to the setting. At 8:37, the screen particles start to break up and swirl again, as the pitched and glassy tones return. Having created the expectation and deliverance of this one climactic audiovisual scene, there is little left to accomplish in the piece, and it quickly resolves the pitched tones into a final

[220] Reproduced with permission from Dr. Bret Battey

melodic closing chord, fading to silence as the particles lose their energy and also fade to black, closing completely to a blank screen at 9:12, followed by the credits sequence.

Autarkeia Aggregatum is an ideal piece to study to see how subtle uses of pitch, tone, recurrent audio themes can be combined with underlying visual motifs and recurring colour patterns to build up this sense of expectation, particularly effective in this case as the initial reception of the work is just to become absorbed in the intricacies of the particle movements and the close correlation of shifting particle movements and colour changes with changes in pitch and timbre of the soundtrack. It is only when the climax is reached, and the underlying themes that have been recurring and being built upon in the piece are finally revealed, that the expectation is rewarded, and the audience realises how much of the material they have seen and heard has led up to this moment of revelation.

See Hear (Outside) **(2015)**

During my final year of research, I took part in teaching a module for the 1st year undergraduate Animation students, designed to expand their knowledge of sound and how it works with animated images. In previous years, the students were given a stock soundtrack containing location sounds and sound effects, and were asked to create animated sequences as a visual interpretation of those sounds.

With my involvement, we completely redesigned the module, making use of the theories of electroacoustic composition and audiovisual objects outlined earlier in this thesis. I outlined the ways in which sound and image can be combined to create audiovisual objects, and held screenings of relevant visual music pieces to illustrate the theories. The students were introduced to the idea of the sound, visual and audiovisual object, as well as receiving training in how to actively listen to sounds around them, and in close microphone recording techniques. The eight students were split into two groups of four, and tasked with recording a variety of interesting sounds from either “outside” or “inside”. These recordings were then used as the raw material for electroacoustic compositions by each group, of two minutes each. The groups were then tasked with creating abstract animations based on these compositions, exploiting synchresis and synchronisation to realise an absorbing audiovisual environment. The module continued with the students creating further animations based on electroacoustic compositions provided to them, but the two films that resulted from the first stage of the course proved to be of high quality, and were featured in the Edinburgh College of Art Animation Degree Show screening at the Edinburgh Filmhouse, receiving a very favourable reaction from the audience.

Although both films were of high quality, the composition based on the “outside” recordings is of particular interest, as the students combined the sounds and images in very creative, unexpected ways, exploring the theories and techniques discussed with them at the start of the module. The audio composition process was planned carefully, making creative use of rhythmic patterns discovered in the original recordings. The visuals employed in the animation process were particularly inventive, creating unexpected audiovisual bonds between the sound and images that initially seem unrelated, but which work particularly well.

Construction

The recordings made by the students were used as the raw material for their composition, created using Adobe Audition. As Audition is part of the Adobe Creative Suite, and the students were accustomed to using other Adobe products for their animation and compositing work, this package offered a familiar editing environment within which to work. The brief issued to the students was to create a short two minute composition, ideally using acousmatic compositional techniques to obscure the original source of the sounds – although both groups included identifiable source material as foreground highlights in their work. The students were given a workshop in audio editing and processing skills, and were encouraged to experiment as much as possible while creating their soundtrack, while also having a strict deadline to meet. I was present during most of their compositional sessions, but only to offer advice when requested, rather than actively guiding their work. In the event, my assistance was only necessary when the students required specific effects to be applied to their audio objects, and were unsure of the correct processes to follow. The resulting two minute compositions were well-conceived, and on their own, without visuals, were coherent and aesthetically satisfying.

The students assumed that the compositions they were creating would be given to the other group for the animation process, although in fact they were asked to animate to their own composition as I judged that their familiarity with the material would lead to a much more effective creation of audiovisual objects. Within each group, each student was tasked with creating thirty seconds of the animated material to accompany the soundtrack, while also ensuring the transition from their work to the next student's was seamless. The group dealing with the "inside" material had retained more of the recognisable sounds of mechanisms they had recorded (a kettle, an elevator, a fan and a microwave) and used identifiable line drawings of these elements as the transition points, then altering and transmuting the images to accompany the changes in the soundtrack. The "outside" group took a different approach, using some drawn animation in parts of the composition, but in the main applying recorded visual footage of ink in water, moving lights and other footage to match up with their sound objects, making effective use of synchresis in realising close audiovisual fusion.

The compositions were composited and edited using Adobe Premier Pro, allowing direct data transfer with Audition, which simplified the overall compositing process.

The students provided feedback on the process, describing how they decided which images to create or select to use against their compositions. Tyler Carrigan, one of the students who worked on the "outside" composition, had contributed a lot of the rhythmic material in the audio composition process, and noted that he used the inherent motion of the sound to inform the visuals that he felt fit against that sound movement - "to get that movement, you kind of have to imagine, you have to really listen to the sound and see what kind of shape it makes, or where it goes"[221]. The students had different philosophical approaches to the process of audiovisual composition. Some felt that sound and vision were two distinct perceptual streams, but with strong influences on each other – one felt that "I perceive them [sound and vision] as two separate perceptual streams. I think sound can be a strong emotional influence towards one's perception towards a visual"[222], while another thought "I see [sound and vision] as integrated, I find myself more detached from the animation when you're making it and have no idea what it will sound like"[223]. In the critique session examining their audiovisual work, all the students agreed that approaching sound and vision as entwined parts of an integrated audiovisual perceptual stream made the bonds between the soundtrack and animation much tighter, and felt that they had become more appreciative of sound as just as important as vision in creating high quality audiovisual work.

See Hear (Outside) - Critique

Although the students each animated a thirty second segment, there was a lot of discussion at every stage of the visual composition stage in order to ensure a consistent approach across the entire work, while still allowing each animator to create work they felt comfortable with in their segment. This continues the gestalt practice across my own approach to audiovisual composition, with the individual animators having to create a continuity across their combined work, exploiting similarities, repetition and re-occurrences of themes to tie the material together.

The opening thirty seconds make good use of an underlying 'cricket'-like background sound texture, with an initial brief visual of white smoke against a dull black background, switching to dust particles again in a dark dull setting. The sound then begins to feature repeated looped material, forming a distinct rhythm, which is

[221] Recorded conversation with Animation students, 10/06/2015.
[222] Email conversation with Alice Pearse, See Hear "outside" project, 04/04/2015.
[223] Email conversation with Zoe Hutber, See Hear "inside" project, 03/04/2015.

matched to an ongoing repeated visual clip of loosely grouped, dusty ‘flashes’ across the screen. The repeated sound loops continue and gain in complexity, while the visuals remain the same – this does not detract from the audiovisual fusion, rather the audience seeks to find closer correlations between the sound and the visuals. Every so often, the two appear slightly out of sync, but that actually makes the synchresis more effective, as there is a ‘satisfaction’ when the slight discontinuity is resolved and the sound and visuals return to their ‘proper’ synchronisation. The repeated sound and vision is interrupted three times by a glassy ‘chime’, panning from right to left, which is mirrored on screen by a small circular blue light also passing from right to left, a close synchronisation which provides a counterpoint to the less co-ordinated synchresis of the main repeated theme. The section draws to a close with a muffled, falling-pitch echoing sound object, which is matched on screen by falling loops of white and pale blue ink drops, against a background of pale blue dust particles, recalling the initial starting images, and rounding the section off, leading to a black screen and the start of the second section.

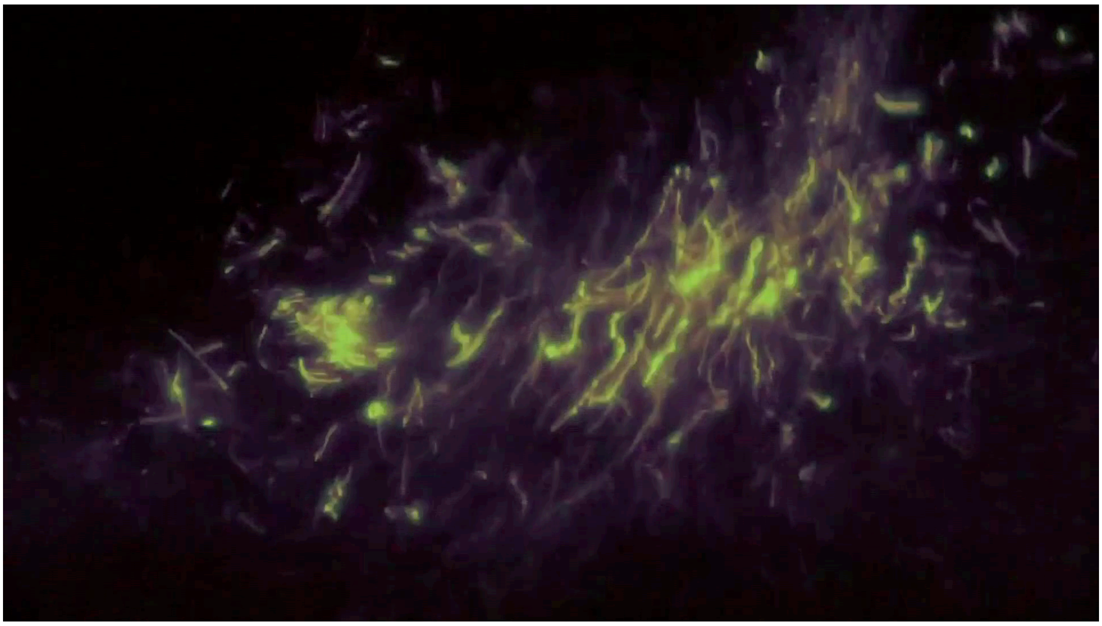


Figure 7: Recurrent visual theme in first section of *See Hear (Outside)*

The second section starts with an identifiable foreground sound object, a recorded voice message on a telephone line. The message consists initially of repeated numbers, interpreted literally on screen with a series of numbers in different font styles, small images in the centre of a large black screen. This is followed by the words of the recorded message, “this is a free call” again being shown literally as typeset words on the screen, but then returns to an abstract world with a repeated siren sound against a circular, abstract white swirled pattern against the black background, swelling and contracting in time to the increase and decrease in the sound’s amplitude. The voice recording continues with a looped sample of a single syllable, against a monochrome, slightly obscure mottled image of black specks against a shaded grey plane, again fading in and out in accordance with the sound’s repetitions. This is followed by more repetitive sampled material, initially a repeated mechanical gearing noise against the cricket background from the first section, matched against a rotated repetition of an ink stream impacting and dispersing into water, reversed into a negative image with some additional colour manipulation. The next sequence shows a repeated circular fan sample, matched initially against a dropping speckled white tape falling down the left side of the standard black background, then rapidly switching to a repeated slice of a hazy video of a wave approaching a rocky shore – this is abruptly cut to the black background while the sound continues to repeat and fade in the background, opening the way to the third section.

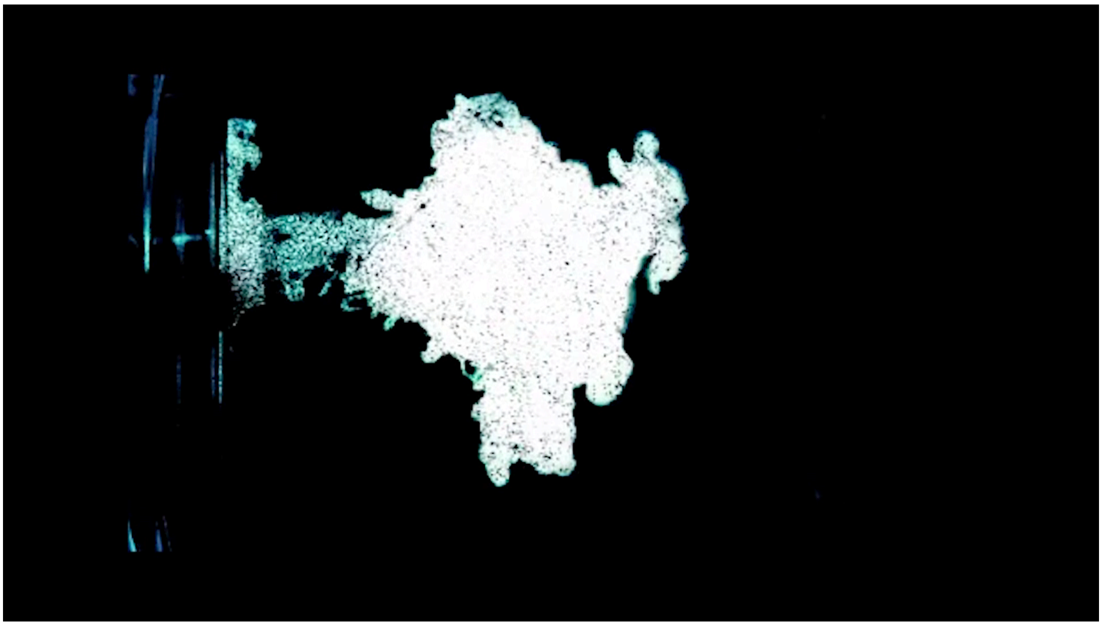


Figure 8: Manipulated negative ink droplet in second section of *See Hear (Outside)*

The third section sonically revisits the main cricket-like background texture established at the start of the piece, but highlights foreground sound objects of a rhythmical 'popping' sound. The background texture is represented visually by a white grainy noise against the black background, similar in style to the initial grainy texture of the first section, but less defined in quality and out of focus. This section marks the use of a more traditional 'drawn' animation element in the visual representation of the popping noises, centred on simple geometric forms such as squares, circles and triangles, but with added coloured surrounding 'echoes' of the base shapes as the popping noises themselves have brief repeated echoes. A brief scraping sound, repeated a couple of times, corresponds with an understated flicker of a ribbon-like texture vertically along the right hand of the screen, while the section closes with a slightly louder, more insidious repeated sound, almost like a muted siren, which is represented visually by a larger centrally placed repeated triangular motif, again using a variety of colours to provide continued interest and variety. This section does not fade out into a black screen as per the previous sections, rather it carries straight on into the final thirty second section.

Figure 9: Geometric drawn pattern and fluid vertical ‘ribbon’ in third section of *See Hear (Outside)*

The final section acts in many ways as a reprise of the themes already explored in the first minute and a half of the work, although with a few audio and visual twists. The section starts with a repeated church bell tone, matched unexpectedly yet very effectively with a series of manipulated images of white fluid dropping, expanding and contracting on top of the black background, while the background ‘cricket’ audio texture continues. As the bell fades away in repeated echoes, the background texture mutates into a muffled, aquatic texture, with slightly processed droplet sounds featured as foreground objects. The visuals reflect this water-like environment, maintaining the manipulated droplet motif but adding in small opening ‘windows’ of more ink drop in water forms, again recalling the earlier foreground objects in the second section. This moves through a couple of full-screen ink drop in water images against the aquatic background textures, into a reprise of the spoken telephone sequence from the start of the second section, in this case matched against scattered sand animations, which, rather than recreating the pasted text work from before, recall the geometric shapes from the third section, this time emerging from then disappearing back into the shifting sand textures. The piece ends with a repeated water drop recording, building in volume to an abrupt climactic halt. The sand texture is continued, but the water sounds are emphasized by using the sand to create the images of successive growing water drops, also building towards a climax, but cutting just short at the end to a black screen, which gives the sound more prominence and heightening the tension of the sudden stop after the increased building tension.

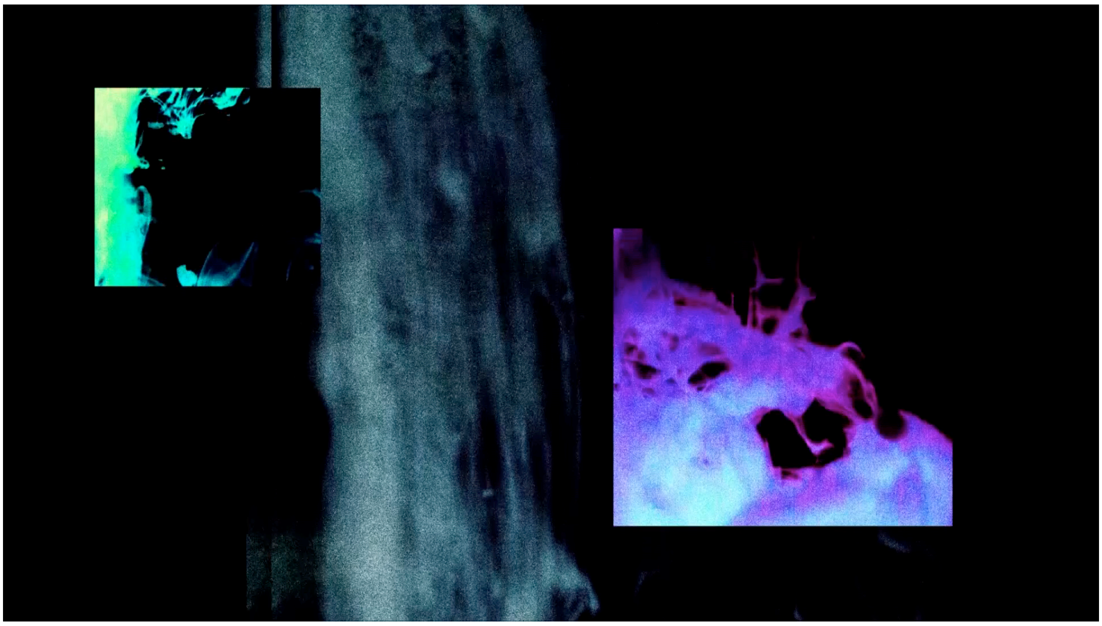


Figure 10: Ink droplet inserts and main screen in final section of *See Hear (Outside)*

The piece as a whole integrates sound and vision very well, making use of some unusual combinations which are initially surprising, yet which hold together well, such as the fluid droplets against the sound of church bells. The ongoing use of repetition in both the sound and visuals bring forward the Gestalt principles of Proximity and Similarity, where the ongoing audiovisual objects are convincing because of their synchronisation and repeated associations. There are elements of Common Fate, especially in the first section, where the synchresis between the repeated sounds and the moving repeated visual patterns also imply a shared direction and purpose behind the audiovisual combinations. Figures such as the geometric shapes in the third section exhibit the principle of Surroundedness, the shapes against their textural background on screen, and their audio equivalents, popping as a foreground object against the continued sussurations in the background.

The final section highlights elements of Emergence, both as the foreground sounds emerge from the building background textures and repetitions, and literally in visual terms as the shapes form from and dissipate back into the underlying sand texture.

This work in particular satisfied one requirement I had for my research into audiovisual composition – that the compositional techniques I had developed from the underlying electroacoustic and animation techniques could be communicated to and employed by others, particularly creative practitioners who had some familiarity with at least one aspect of audiovisual composition, but who had not previously paid great attention to the mechanics of achieving a close match between sound and vision. A positive outcome from the process was the greater awareness of the students of how they can achieve a closer relationship between sound and vision in their work - "it's been really beneficial having that bit more insight into it and [...] at a bare minimum, taking more notice of what goes on in sound and image at the same time"[224].

The audiovisual work produced by the students was of particular interest for me, as I have built up my approach to audiovisual composition based primarily on theories of electroacoustic composition, and always work initially in sound, with the visual accompaniment I envisage as I create the soundtrack being realised and altered in the secondary stage of composition. In this case, the students came from a primarily visually oriented approach, with shared backgrounds in illustration, animation and visual editing, and had limited experience in sound design. Given the structure of the module, where the sound composition came first in order to concentrate their attention on the complexities of animating to existing sound, the resulting compositions followed similar methodologies to my own compositional processes. I would be interested in experimenting further in examining how people with different backgrounds approach the composition of audiovisual works where the construction methods are not constrained by the module brief, and to see how the works would vary if an animator started with visuals rather than sound, or an electroacoustic

[224] Recorded conversation with Animation students, 10/06/2015.

composer based their sound elements on a silent visual component they created first.

Conclusions

In Chapter 2, I mentioned other approaches to perception, such as Gibson's theories of affordance in perception. This approach implies that we assign a value to a perceived object based on our experience – we see the sea, and assign it a value of being wet, as we've encountered it before. In the case of abstract audiovisual work, such as the pieces examined as case studies in this chapter, there is no specific earlier experienced value assignable to the fused audiovisual objects, although individual perceptual objects may have values based on other experience e.g. the church bells in *See Hear (Outside)*. For these works, analyses based on Gibson's theories are restrictive, as this would rely on the establishing the values assigned to perception by the individual audience members, which can vary widely. However, Gestalt Principles offer an immediately accessible form of analysis, allowing us to establish perceptual links between sight and sound, and so are more appropriate in critiquing existing work and as guidance in constructing new compositions.

The case studies presented here are only three examples of many works of visual music, abstract experimental works and narrative audiovisual pieces that I regularly present to students and peers in explaining the principles behind audiovisual composition. It can be seen from these works that not all the Gestalt principles noted in Chapter 2 have to be present in a work for it to be a successful combination of audiovisual objects, and that only a few points of convergence between sound and vision are necessary in some cases to convey an immersive audiovisual experience, particularly where there is close synchronisation achieved between the two perceptual streams in works such as *Autarkeia Aggregatum*, where the compositional techniques themselves force a very close bond between the visual and audio elements.

From the work carried out with the animation students, it is encouraging to note that the key principles identified in my research can be communicated to other creative artists, particularly to those whose primary focus is in the visual field rather than the audio I approach the subject from. As noted above, further exploration in this field would be of interest, varying the initial boundaries for creating audiovisual work, with electroacoustic composers taking an initially visual approach, and comparing that to work created by animators and other visual artists approaching their compositions from both audio and visual perceptual streams as their starting points.

From my own compositional work, and the compositions produced by the animation students undertaking the See Hear module, the importance of using both synchronisation and synchresis in crafted fixed-media audiovisual work is highlighted – too much simple synchronisation and the work becomes unconvincingly mechanistic, a sound visualisation exercise that has no depth or creative value to it, while synchresis alone can lead to extended periods where the sound and vision may appear to be at odds with each other. This may work well if part of the overall compositional vision, but where it is not planned as an integral feature, it can again lead to a loss of the immersive qualities we are looking for in a cohesive audiovisual composition.

The use of Gestalt Principles, and planning for the use of synchronisation and synchresis, within the compositional process can be seen to allow the audiovisual composer a toolkit for building interest and immersive qualities into their work from the early planning stages onward. By examining stages of composition such as the compositional plan, the storyboard, early drafts of the soundtrack, the animatic, or any other constructive method for creating the work, and relating the material to Gestalt Principles as above, the audiovisual composer is able to assess and adjust their work in progress. This enriches their final composition, and ensures that the composer is able to eliminate unnecessary or unsatisfactory material before it becomes embedded in their work.

Chapter 6

Conclusions

Introduction

The concept of the audiovisual object has its roots in Schaeffer's sound object, using a phenomenological approach to how the visual and sound elements of two perceptual streams can under certain conditions fuse into a perceived whole audiovisual gestalt. Eisenstein refers in cinematic terms to this gestalt approach as a sum greater than the parts, where a montage of sound and image combine to create a whole that "emerges perfectly as a 'third something'"[225] In terms of audiovisual composition, that 'third something' is the audiovisual object.

The audiovisual object links theories of sound design and electroacoustic composition with film and animation theory and allows the composer a basis on which to build both their own practice and to analyse the works of others, pinpointing the elements that combine well or badly to inform and inspire their own work. The use of audiovisual objects within the audiovisual environment as the basic building blocks for the composition or narrative of the composition, featuring as foreground highlight, background texture or any combination between, allows the audiovisual composer to structure their own work and have an insight into why specific elements inspire a personal or audience reaction, drawing the viewer into the immersive environment of the work, and achieving the intended emotional response and appreciation of the composer's creative purpose.

The spark for this research was initially set by Jordan Belson's aspiration for his own creative practice - "you don't know if you're seeing it or hearing it"[226]. Over the course of my research and practice, I have found my own approach and appreciation for the field to have modified from this initial position. I would now contend that the audience member can both see it and hear it, and that's exactly the way I have designed my work to be perceived.

---

[225] Eisenstein, S. (1943). *The Film Sense.* Faber and Faber Ltd. (1970 edition). p19.
[226] Brougher, K. et al. (2006). *Visual Music: Synaesthesia in Art and Music Since 1900.* London: Thames & Hudson

Through the investigation, comparison and integration of theories such as Schaeffer's definition of the sound object, Gestalt principles with regard to visual perception and their extension to audio perception, perception of colour, and cross-modal perception, this thesis establishes a theoretical basis for the concept of the audiovisual object. There is a further contribution to the field outlining how the audiovisual object can be used in creative practice, how the idea of the audiovisual object can be used in assessing the impact of composed works, and how the development of electroacoustic composition and abstract animation over time has created an artistic environment which is eminently suitable for audiovisual composition. Given the nature of contemporary work in the field, the performance space and underlying technology is also a key element in how the audiovisual composition will be received and appreciated by the audience, and therefore is also an essential component of the compositional process.

Limitations

This research outlined in this thesis is subject to certain limitations. The boundaries under which the audiovisual object has been defined and examined have been solely within the realms of electroacoustic composition and abstract animation. While my research has included theories and practice from other genres of composition and film-making, the application of this research to develop the concept of the audiovisual object and its use in creative practice has been only been explored in terms of my own creative fields of interest. It is possible that the concept of the audiovisual object can be extended to wider audiovisual practices, but that is outwith the scope of this investigation. Additionally, the principal framing of the audiovisual object has been within the confines of phenomenology and Gestalt principles. While I have included some consideration of Gibson's theory of affordances in perception in chapter 2, I have not carried out any in depth exploration of how the audiovisual object concept could be defined using this alternative approach. Other approaches to the audiovisual object could also be taken with regard to neuropsychology and perception, which would require a greater understanding of neurosciences than I can bring to bear to the field. However, the research centring on the audiovisual object within the confines as established in my practice still offer a useful approach to the practice of composing audiovisual work and creating an aesthetically enjoyable immersive experience for the audience.

The Audiovisual Object in Theory and Creative Practice

By extending Schaeffer's concept of the sound object to encompass the visual object and on to define their fusion into the audiovisual object, the audiovisual composer is able to treat the audiovisual object as a key building block in the structure of their own compositions, in a similar way to the use of the sound object as an equivalent to the note or phrase in classical music composition. Gestalt principles and associated phenomenological approaches to audiovisual perception, such as the exclusive allocation of boundaries, can be used by the audiovisual composer to create an effective environment, with elements crossing over between the foreground and background to add interest and narrative development.

The combination of sonic and visual elements to reinforce or modulate each other is an important part of the formation of the audiovisual object. The continuum explored in Chapter 2, from separation of the two perceptual streams through to full synchronisation, also offers a variety of approaches the audiovisual composer can take to create variety and stimulation for their audiences. By exploiting the combination of sound objects, visual objects and audiovisual objects in the creative process, the composer is able to maintain interest and develop their ideas throughout their work, providing an immersive and interesting audiovisual experience.

While the combination of sound and vision reflects the multi-modal approach to perception that the human mind appears to take, this also has developed from our ongoing use of technology to combine the two senses in creative performances. The focus of my research has been in the combination of electroacoustic composition with abstract animation, and the development of both art forms have encouraged experimentation and the pushing forward of boundaries.

The early animators were willing to experiment with form and perspective, mixing techniques and exploiting the plasticity of their new medium, while the development of electroacoustic and acousmatic composition likewise adopted new technologies and techniques to expand the art form. Audiovisual composition has a long history, from the visual music of Fischinger through to the planetarium spectaculars of Belson, but the latest developments in digital technology placing very powerful audio and video editing packages in the hands of the individual have allowed a new expansion of talent into this field. The techniques and exploration of theories detailed in my research allow the composers entering this field a method to examine and assess the works that they find inspiring, and to understand how they in turn can also follow the guidelines of gestalt principles and multi-modal perception to create their own immersive work.

In tandem with exploring the compositional theories and practices contributing to successful audiovisual work, the composer also needs to consider the effect the performance space has upon their work. Both the space itself and the technology employed to deliver the performance dictate different considerations that the composer has to allow for in their approach to the creative process, while both also have limitations which affect some of the impact the experience has for the audience. The other potential performance spaces explored in Chapter 4 point towards a development of the artform, and a potential new direction for further research in this area.

The case studies presented in the thesis illustrate how gestalt principles and the use of synchronisation and synchresis can be assessed in existing works to investigate the effectiveness of this approach. Not every principle has to be evident in each work, and the convergence of audio and vision to create the audiovisual objects within each work can be extremely close or quite loose, but that combination is still the driving force to convey the narrative and immersive qualities of the piece. My own practice, as assessed in my piece *Canto* in Chapter 5, and as presented in the portfolio and additional notes accompanying this thesis, explores several different aspects of synchronisation, synchresis, and the gestalt principles outlined in Chapter 2. As an audiovisual composer, each starting point of a blank page and silent score requires inspiration in some form. The use of the theories and practices outlined in this thesis establish boundaries within which the creative act can occur, giving a structure to the overall work and a spine from which the work can develop. My experience in communicating these principles and methods of working to students starting their own practice in animation and the use of sound has shown that these concepts are also of use to artists entering this field.

Developing a Research Pathway

The potential research pathway established in this PhD places creative practice and experimentation at the heart of the research, providing a basis and direction for the associated theoretical research and development of my own theories of audiovisual composition. From my research to date, two attractive directions for future research have emerged, based upon the space within which the work is performed. In each case, the underlying aim of the research would be to achieve a convincingly immersive experience, given that as “a first principle, art is an activity *intended to influence the minds of an audience*.”[227] Both potential directions of research would aim to expand upon my existing research to establish that influence, that immersion within the piece and the ability of the work to stir the emotions of the audience, with the division being whether or not that is achieved in a shared audience space or on an individual basis.

[227] Donald, M. (2006). *Art and Cognitive Evolution.* In Turner, M. ed. “The Artful Mind: Cognitive Science and the Riddle of Human Creativity”. Oxford University Press. p4.

Shared Space

As outlined in chapter 4, the performance space itself has a great impact on the effect an audiovisual composition has on an audience. While a cinema has the advantage of familiarity for an audience, it offers much less control of the auditory aspects of performing the work – conversely, an electroacoustic concert format offers great control of the sonic space, while limiting the visual space to a small flat front projection screen. However, the *Vortex* concerts carried out by Belson and Jacobs in San Francisco in the late 1950's would indicate that a planetarium offers a potentially suitable space for the performance of audiovisual compositions.

For the audiovisual composer, this format presents some challenges. The use of hemispheric projection for the visual portion of their composition provides a challenge in how to structure their animation, recordings, or processed video work to map correctly onto the curved surface. 3D CGI may be of use in this particular case, creating a virtual equivalent to the real world dome, and allowing the composer to gain some insight in to the overall composition of his work. Similarly, the multi-speaker system used in planetariums will also have an effect on the audio composition process, as the composer will need to allocate the audio material over a complete panorama of speakers, given that a standard planetarium allows up to 32 channels across the expanse of the dome[228]. This allows for a close correlation between the perceived visual and audio origins of the audiovisual objects in the composition, but also increases the complexity of the composition process.

There are limitations in the planetarium as a performance space for audiovisual compositions. The amount of data being processed for a standard set-up of 6 overlapping projectors and 32 audio channels requires a great deal of processing power and storage, and most audiovisual presentations are limited to less than 30 minutes as a result. The venue itself also has a limitation in where the 'sweet spot'

[228] Review of the Sky-Scan planetarium system presented at the Networld & Interop Show in Atlanta, Georgia, USA in 1999. Available at http://www.ips-planetarium.org/?page=a_scott1999 [accessed 28/10/2015].

covers, giving the audience the best possible combination of video and audio. The costs of performance for a relatively small audience means that the format is unlikely to ever be a profitable enterprise. However, this format does allow a shared experience for an audience, a sense of occasion, where the performance can be enjoyed with group of peers, discussed afterwards, establishing a lasting effect of the performance as a memorable experience for the group.

For the creative practitioner and researcher, this avenue for presenting audiovisual work offers several interesting research pathways. The methods of constructing audiovisual material for use in this format would require some thought, making use of the principles outlined in this thesis along with current practice in constructing planetarium display material, adapting existing procedures to allow for the different aesthetic needs of the audiovisual composer to be met. The psychological effect of experiencing audiovisual works within the planetarium environment would also be an interesting pathway to follow, understanding the audience response and correlating this to the compositional material and techniques employed by the composer.

Individual Space

The use of a planetarium as the performance space caters to a relatively small audience as compared to a large cinema. However, the second research pathway that arises from this initial research into the audiovisual object would address a performance space designed for the individual rather than a shared experience. As noted in Chapter 4, another emerging platform for audiovisual experiences is within a created virtual reality, with lightweight headsets such as *Oculus Rift* and *Project Morpheus* slated to become commercially available in 2016.

While the immediate market focus of these headsets is likely to be in the games industry, the virtualisation of individual experience is also becoming established in the cultural and heritage sectors. Guidazolli et al, writing about virtual museums accessible through the Internet, notes “in an immersive context the involvement can be deeper […] immersive digital museums enable a step further toward the achievement of sensorial and emotional experiences”[229]. Immersive VR headsets offer an immediate access to an environment where, for example, the interior of a museum can be recreated, with the user able to ‘pick up’ and examine exhibits, impractical in the real world. Lev Manovich has also touched on the development of audiovisual technology into simulated realities such as flight simulators, another step towards an immersive VR experience[230].

In terms of developing a research pathway in this area, the imminent arrival to the market of these new VR platforms presents a challenge to existing conventions of screen language. While most new media mirror their older counterparts at first, true innovation in creating content for virtual reality platforms will only occur when we drive past this to establish a new set of definitions unique to this new media. The conventions used in existing media productions allow for a non-linear progression of the narrative, using cuts to compress space and move the viewer to a new location, fades to compress time and move the story onwards. In a fully realised virtual environment, what will be the new conventions? How is the audience guided through the experience without feeling dictated to?

These questions about the aesthetics and lexicon of this new media platform can be explored through further research in combination with creative practice and experimentation in exploring how audiovisual composition can be established within a virtual world. In the same way as the early animators had freedom to experiment and trial their own creative ideas in the then new plastic art of animation, the fusion of digital technology with the developments in video and sound hardware has opened up a new opportunity for the same early development in another medium. By building upon the existing audiovisual practices and conventions, the audiovisual composer working within VR can test, challenge and develop new ideas of

[229] Guidazolli, A, et al. (2000). *From Museum to Mouseum. The Virtualisation of Daily Life Museums.* Conference: VAST2000, Volume: BAR International Series 1075, Oxford, Archaeopress.

[230] Manovich, L. (2001). *The Language of New Media.* MIT Press. Chapter 6. “What is Cinema?”.

expression and new paradigms for the successful construction of the audiovisual environment within a completely created virtual world.

This particular pathway offers a particularly rich avenue for exploration, both within the challenges of creative practice, creating successful audiovisual work to test and establish the practical production principles for virtual audiovisual composition, but also within creative media theory, establishing the new language of this medium, developing from the existing conceptual frameworks established in film, television and ludic theories. For the audiovisual composer, and for my own personal fields of interest, the future of audiovisual composition within the VR medium provides the greatest potential for research, creative practice, and development of the art form itself.

Bibliography

Bibliography

ALBERS, J. 1963. *Interaction of Colour*. Yale University Press.

ALTMAN, R. 1992. *Sound theory, sound practice,* New York, N.Y. ; London, Eng., Routledge.

ARNHEIM, R. 1974. *Art and Visual Perception: A Psychology of the Creative Eye.* University of California Press.

BARON-COHEN, S. & HARRISON, J. E. 1997. *Synaesthesia : classic and contemporary readings,* Oxford, Blackwell.

BARRIER, J. M. 1999. *Hollywood cartoons : American animation in its golden age,* New York, Oxford University Press.

BENDAZZI, G. 1994. *Cartoons : one hundred years of cinema animation,* London, John Libby.

BERNSTEIN, D. W. 2008. *The San Francisco Tape Music Center : 1960s counterculture and the avant-garde*. Berkeley; London: University of California Press.

BREGMAN, A. S. 1990. *Auditory scene analysis : the perceptual organization of sound,* Cambridge, Mass. ; London, MIT Press.

BROUGHER, K., MATTIS, O., MUSEUM OF CONTEMPORARY ART (LOS ANGELES CALIF.) & HIRSHHORN MUSEUM AND SCULPTURE GARDEN. 2005. *Visual music: synaesthesia in art and music since 1900,* London, Thames & Hudson.

BRUCE, V., GREEN, P. R. & GEORGESON, M. A. 2003. *Visual Perception: Physiology, Psychology and Ecology,* Hove, UK, Psychology Press.

CALVERT, G., SPENCE, C. & STEIN, B. E. 2004. *The handbook of multisensory processes,* Cambridge, Mass. ; London, MIT Press.

CANEMAKER, J. (ed.) 1988. *Storytelling in Animation: The Art of the Animated Image,* Los Angeles: The American Film Institute.

CANEMAKER, J. 1996. *Felix: The Twisted Tale of the World's Most Famous Cat,* New York, Da Capo Press.

CAVALIER, S. 2011. *The world history of animation,* London, Aurum.

CHAGAS, P. C. *Unsayable music: six reflections on musical semiotics, electroacoustic and digital music*.

CHION, M. 1983. Guide des Objets Sonores. Paris: Buchet Castel.

CHION, M. 1994. *Audio-vision: sound on screen,* New York, Columbia University Press.

CLARKE, D. & CLARKE, E. F. 2011. *Music and consciousness: philosophical, psychological, and cultural perspectives,* Oxford, Oxford University Press.

COHEN, A. A. (ed.) 1978. *The New Art of Colour: The Writings of Robert and Sonia Delaunay*. The Viking Press. New York.

COHEN, A. J. 2005. How Music Influences the Interpretation of Film and Video: Approaches from Experimental Psychology. *In:* KENDALL, R. A. & SAVAGE, R. W. (eds.) *Selected Reports in Ethnomusicology: Perspectives in Systematic Musicology.* Los Angeles, CA: Dept. of Ethnomusicology, University of California.

COOK, N. *Analysing musical multimedia*.

COUSINS, M. 2004. *The story of film,* London, Pavilion.

CRAFTON, D. 1990. *Emile Cohl, Caricature, and Film*. Princeton University Press.

CYTOWIC, R. E. 1997. Synaesthesia: Phenomenology and Neuropsychology: a Review of Common Knowledge. *In:* BARON-COHEN, S. & HARRISON, J. E. (eds.) *Synaesthesia: Classic and contemporary reading.* Oxford: Blackwell.

DAVID, E. E. & DENES, P. B. 1972a. *Human communication: a unified view,* New York, McGraw-Hill.

DAVID, E. E. & DENES, P. B. 1972b. Visual and Auditory Perception - An Essay of Comparison. *In:* DAVID, E. E. & DENES, P. B. (eds.) *Human Communication: A Unified View.* New York, USA: McGraw Hill.

DI SCIPIO, A. 1995. Inseparable models of materials and of musical design in electroacoustic and computer music. *Journal of New Music Research,* 24**,** 34-50.

DWOSKIN, S. 1975. *Film Is...* Peter Owen, London.

EISENSTEIN, S. & LEYDA, J. 1943. *The film sense,* London, Faber and Faber.

EMMERSON, S. 1986a. *The language of electroacoustic music,* Basingstoke, Macmillan.

EMMERSON, S. 1986b. The Relation of Languages to Materials. *In:* EMMERSON, S. (ed.) *The Language of Electroacoustic Music.* Macmillan Press.

FUXJÄGER, A. 2012. Translation, Emphasis, Synthesis, Disturbance: On the function of music in visual music. *Organised Sound,* 17**,** 120-127.

GARRO, D. 2005. A glow on Pythagoras' curtain. A composer's perspective on Electroacoustic music with video. *EMS International Conference.*

GARRO, D. 2012. From Sonic Art to Visual Music: Divergences, convergences, intersections. *Organised Sound,* 17**,** 102-113.

GEDULD, H. M. 1975. *The birth of the talkies: from Edison to Jolson,* Bloomington, Indiana University Press.

GIBSON, J. J. 1979. *Ecological Approach to Visual Perception, The,* Boston, USA, Houghton Mifflin Company.

GOERGEN, J. 1989. *Walter Ruttmann - Eine Dokumentation*. Freunde der Deutsche Kinematik, Berlin.

GOLDMARK, D. & TAYLOR, Y. 2002. *The cartoon music book,* Chicago, Ill., A Cappella.

GOMERY, D. 2005. *The coming of sound: a history,* New York; London, Routledge.

GREY, C. (ed.) 1971. *Hans Richter by Hans Richter*. Thames and Hudson Limited, London

GRIFFITHS, A. 2008. *Shivers down your spine: cinema, museums, and the immersive view,* New York, Columbia University Press.

GUIDAZOLLI, A., BONFIGLI, M. E. & LIGUORI, M. C. From Museum to Mouseum: The Virtualisation of Daily Life Museums. *In:* NICCOLUCCI, F., ed. Virtual Archaeology VAST Euroconference, 2000 Arezzo, Italy. BAR International Series 1075, 213-215.

HAGENER, M. 2007. *Moving Forward, Looking Back: The European Avant-garde and the Invention of Film Culture, 1919-1939*. Amsterdam University Press.

HARRISON, J. 1998. Sound, space, sculpture: some thoughts on the 'what', 'how' and 'why' of sound diffusion. *Organised Sound,* 3**,** 11.

HARRISON, J. E. & BARON-COHEN, S. 1997. Synaesthesia: An Introduction. *In:* BARON-COHEN, S. & HARRISON, J. E. (eds.) *Synaesthesia: Classic and contemporary readings.* Oxford: Blackwell.

HOLMES, T. B. 1985. *Electronic and experimental music,* New York, Scribner's.

HOLTZMAN, S. R. 1994. *Digital mantras: the languages of abstract and virtual worlds,* Cambridge, Mass. ; London, MIT Press.

HUSSERL, E. & HILL, C. O. 2008. *Introduction to logic and theory of knowledge: lectures 1906/07,* Dordrecht; London, Springer.

JONES, C. 1946. Music and the Animated Cartoon. *Hollywood Quarterly,* 1**,** 364-370.

JULESZ, B. & HIRSCH, I. J. 1972. Visual and Auditory Perception - An Essay of Comparison. *In:* DAVID, E. E. & DENES, P. B. (eds.) *Human Communication: A Unified View.* New York, USA: McGraw Hill.

KAES, A, JAY, M & DIMENDBERG, E. (eds.) 1994. *The Weimar Republic Sourcebook*. University of California Press, Berkeley.

KAHN, D. 1999. *Noise, water, meat: a history of sound in the arts,* Cambridge, Mass., MIT Press.

KANDINSKY, W., GOLFFING, F. & OSTERTAG, F. 1947. *Concerning the spiritual in art, and painting in particular, 1912,* New York, George Wittenborn.

KANE, C. 2011. The synthetic color sense of Pipilotti Rist, or, Deleuzian color theory for electronic modern art. In *Visual Communication. Vol 10, Issue 4*

KEEFER, C. & GULDEMOND, J. (eds.) 2012. *Oskar Fischinger 1900-1967: Experiments in Cinematic Abstraction.* Eye Filmmuseum / Center for Visual Music.

KERINS, M. 2011. *Beyond Dolby (stereo): cinema in the digital sound age,* Bloomington, Indiana University Press.

KERNELL, D. 2016. *Colours and Colour Vision: An Introductory Survey*.

Cambridge University Press.

LEHAR, S. 2003. Gestalt isomorphism and the primacy of subjective conscious experience: A Gestalt Bubble model. *Behavioral and Brain Sciences,* 26**,** 375-408.

LUCRETIUS CARUS, T. & SISSON, C. H. 1976. *De rerum natura = The poem on nature,* Manchester, Carcanet New Press.

MANOVICH, L. 2001. *The language of new media,* Cambridge, Mass. ; London, MIT Press.

MARKS, L. E. 1997. On Coloured-hearing Synaesthesia: Cross-modal Translations of Sensory Dimensions. *In:* BARON-COHEN, S. & HARRISON, J. E. (eds.) *Synaesthesia: Classic and contemporary readings.* Oxford: Blackwell.

MCGURK, H. & MACDONALD, J. 1976. Hearing lips and seeing voices. *Nature,* 264.

MCLOUGHLIN, N. P. 2003. Bursting the bubble: Do we need true Gestalt isomorphism? *Behavioral and Brain Sciences,* 26**,** 421-421.

MILLINGTON, B. 1992. *Wagner,* London, Dent.

MORI, M., MACDORMAN, K. F. & KAGEKI, N. 2012. The Uncanny Valley. *Robotics Automation Magazine, IEEE,* 19**,** 98-100.

MORITZ, W. H. 1988. Some Observations on Non-Objective and Non-linear Animation. *In:* CANEMAKER, J. (ed.) *Storytelling in Animation: The Art of the Animated Image.* Los Angeles: AFI.

MORITZ, W. H. 2004. *Optical Poetry: The Life and Work of Oskar Fischinger,* Eastleigh, John Libbey Publishing.

NEEDHAM, J., WANG, L., ROBINSON, K. & LU, G.-D. 1962. *Science and civilisation in China. Vol. 4, Physics and physical technology,* Cambridge, Cambridge University Press.

O'BRIEN, C. 2005. *Cinema's conversion to sound: technology and film style in France and the U.S,* Bloomington, Indiana University Press.

PARTAN, S. & MARLER, P. 1999. Communication Goes Multimodal. *Science,* 283**,** 1272.

PASSUTH, K. & MOHOLY-NAGY, L. 1985. *Moholy-Nagy,* New York, Thames and Hudson.

PEAKER, C 2000. *C Magazine 64*. C Arts Publishing & Production.

READ, H. 1974. *A Concise History of Modern Painting*. Thames and Hudson, London.

RILEY, C.A. 1995. *Color codes: modern theories of color in philosophy, painting and architecture, literature, music, and psychology.* Hanover: University Press of New England.

ROBB, B. J. 2007. *Silent cinema,* Harpenden, Herts, Kamera Books.

RUSSOLO, L. & BROWN, B. 1986. *The art of noises,* New York, Pendragon Press.

SCHAEFFER, P. 1966. *Traité des objets musicaux (Treatise on Musical Objects),* Paris, France, Le Seuil.

SCHEIDIG, W. 1967. *Crafts of the Weimar Bauhaus, 1919-1924: an early experiment in industrial design,* London, Studio Vista.

SERGI, G. 2004. *The Dolby era: film sound in contemporary Hollywood,* Manchester, Manchester University Press.

SMALLEY, D. 1974. Pentes. *Sources/scènes.* Montreal: empreintes DIGITales.

SMALLEY, D. 1986. Spectro-morphology and Structuring Processes. *In:* EMMERSON, S. (ed.) *The Language of Electroacoustic Music.* Macmillan Press.

SMALLEY, D. 2007. Space-form and the acousmatic image. *Organised Sound,* 12**,** 35-58.

SMITH, M. W. 2007. *The total work of art: from Bayreuth to cyberspace,* New York, N.Y. ; London, Routledge.

SOLOMON, C. 1989. *Enchanted drawings: the history of animation,* New York, A.A. Knopf: Distributed by Random House.

SONNENSCHEIN, D. 2001. *Sound design: the expressive power of music, voice, and sound effects in cinema,* Studio City, CA, Michael Wiese Productions.

SPENCE, C. & SOTO-FARACO, S. 2010. Auditory perception: Interactions with vision. *In:* MOORE, D. R., FUCHS, P. A., PLACK, C., REES, A. & PALMER, A. R. (eds.) *Oxford Handbook of Auditory Science: Hearing.* Oxford, UK: Oxford University Press.

STODDART, H. 2000. Review for *Contemporary Visual Arts no 26*

TRACY, J. F. 1999. Whistle While You Work: The Disney Company and the Global Division of Labor. *Journal of Communication Inquiry,* 23**,** 374-389.

TURNER, M. 2006. *The artful mind: cognitive science and the riddle of human creativity,* Oxford; New York, Oxford University Press.

VERGO, P. 2010. *The Music of Painting*. Phaidon Press, London

WARD, J., HUCKSTEP, B. & TSAKANIKOS, E. 2006. Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex,* 42**,** 264-280.

WELLS, P. 1998. *Understanding Animation,* London, UK; Routledge.

WERTHEIMER, M. 1938. Laws of Organization in Perceptual Forms. *In:* ELLIS, W. D. (ed.) *A Sourcebook of Gestalt Psychology.* London, UK: K Paul, Trench, Trubue & Co.

WISHART, T. 1986. Sound Symbols and Landscape. *In:* EMMERSON, S. (ed.) *The Language of Electroacoustic Music.* Macmillan Press.

WISHART, T. 2012. *Sound composition*.

WOOLLEY, B. 1993. *Virtual worlds: a journey in hype and hyperreality,* London, Penguin.

Practice Portfolio

***Study No. 3: Foghorn Abstraction* (2011)**

***Skitterling* (2012)**

***Synthonomy* (2014)**

***Canto* (2015)**

Production Notes

***Study No. 3: Foghorn Abstraction* (2011)**

Study No. 3: Foghorn Abstraction (2011)

Introduction

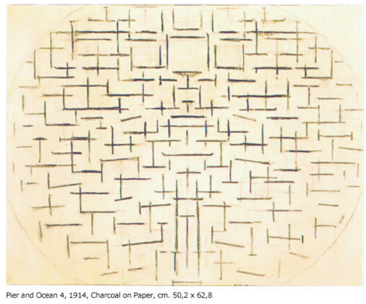
Study No. 3: Foghorn Abstraction was the second audiovisual work created in the course of the PhD, following on from *Study No. 2 – Study No. 1* had been created during my MSc in Sound Design. The second *Study* had investigated some of Kandinsky's theories on colours, their emotional associations, and the associations he had with them in conveying a physical sense of expansion, contraction or stability. In this earlier *Study*, I had used 3D animation to achieve a match between the space implied by the computer modelling seen on screen with the associated colours as prescribed by Kandinsky, but the overall composition had lacked coherence and was unsatisfying as a performance.

In *Study No. 3: Foghorn Abstraction*, I decided to continue exploring the visual ideas and methods of one of the artists I was studying at the time, Piet Mondrian. His distillation of form to the simple shapes and colours used in geometrical abstraction recalled Schaeffer's reduction in listening, leading to his theories concerning sound objects. I was particularly interested in two series of works created by Mondrian – primarily his *Pier and Ocean* series of drawings, where the initial reproductions showing the vertical lines of the pier extending out into the waves of the ocean gradually become simple lines that still convey the simple shapes of the scene he was observing. I was also intrigued by his set of works depicting trees, where the organic growth of branches were also reduced to their essential structures.

The Gray Tree (1912), Piet Mondrian

Apple Tree in Flower (1912), Piet Mondrian

The Sea (1914), Piet Mondrian



Pier and Ocean 4 (1914), Piet Mondrian

The inspiration for audio on this new piece came from one of Andrew Lewis's electroacoustic compositions celebrating four Anglesey beaches, *Penmon Point* (2002-03)[231]. I was first made aware of this work through a fellow attendee at the 2010 Sound, Space, Sight and Play Conference at De Montfort University. Ambrose Seddon presented his paper on the use of recurrence as a structural technique in acousmatic composition, with particular reference to Lewis's *Penmon Point*[232]. The use of the sea bell as a recurrent sound object every thirty seconds, and the fact it changed in sonic structure each time, sometimes close and harsh, sometimes softer and in the distance, also recalled the gradual progression of Mondrian's images, with the forms being altered and reduced to their essential elements at each stage.

Having been struck by the similarities between Mondrian's gradual exploration of visual form by gradually reducing it to its essentials, and Lewis's sonic explorations of recurring sound object, I decided to create a piece essentially carrying out the same actions on an audiovisual object, establishing it as a coherent entity at the start, then gradually breaking it apart into visual and sonic elements, using recurrence as a narrative structure throughout.

---

[231] Lewis, A. (2007). CD *Miroirs obscurs,* empreintes DIGITALes, IMED 0789

[232] Seddon, A. (2010). *Investigating recurrences in Andrew Lewis's Penmon Point.* In Wolf, M & Hill, A (eds.) *Proceedings of Sound, Sight, Space and Play 2010*. Postgraduate Symposium for the Creative Sonic Arts, De Montfort University, UK 2-4/06/2010. Accessible at http://www.dmu.ac.uk/documents/technology-documents/research/mtirc/sssp201003seddon.pdf [accessed 13/08/2015].

Concept

Having established that the work would make use of a recurring sound object, and that both the sound and visual elements would demonstrate a reduction of the foreground and background objects into their essential elements, I then started work on defining which audiovisual objects I would use. Part of the reason that Andrew Lewis's work appealed to me was the presence of the sea bell – having grown up near the sea, the sounds of the beach, ports and ships were a major part of my childhood. I had also just visited Ardnamurchan Head on the west coast of Scotland, which features both a lighthouse and a foghorn. The photographs I had taken of this foghorn would serve as a source for the visual portion of this new work.

The foghorn would be the central focus of the new piece, both as an audio feature, and as the main image source to provide a complex object that would be broken down to its elements. While foghorns used to be a regular sound along the British Coast up until the 1960's and 1970's, they have now been replaced, initially by compressed airhorns, and more recently the requirement for any audible warning signals has been superceded by the introduction of navigational aids using the GPS system. Only a few operational foghorns remain, and are only sounded on specific occasions. From my research, I identified Souter Point Lighthouse near South Shields on the North East English coast as having one of the few remaining operational foghorns. I arranged to visit the lighthouse, and the local volunteers who had restored and maintained the foghorn mechanisms very kindly gave me a full tour of the facility, and were happy to run the foghorn for as many recordings as I wanted to take.

The Souter Point foghorn is a classical diaphone, with the sound created by forcing compressed air through slots in a piston moving within a cylinder. The movement of air causes reciprocation between the piston and the cylinder, resulting in a continuous reverberant tone that is then amplified through a large horn. As the air is cut off and loses pressure, the tone drops in frequency, resulting in a short ending 'grunt' to the signal[233]. British foghorns generate a single frequency tone ending in this grunt, while those in the USA and Canada use two tones, usually a higher then a lower. The normal fundamental frequency of the British foghorn was in the region of 250Hz, with harmonics generated by the process of forcing the air through the piston. Interestingly, the frequency for the fundamental foghorn tone was sometimes determined in advance of construction by playing a piano on the beach below the lighthouse, with listeners placed on a boat out to sea to ascertain which notes carried furthest. The positioning of the foghorn itself was carefully calculated to ensure it would broadcast as far as possible across the sea. The signals made by each foghorn varied in duration and frequency, allowing coastal skippers to know which parts of the coast they were close to. The Souter foghorn, as recorded, has a fundamental frequency of 120Hz, and when in regular use, would sound a 5 second burst every 30 seconds during bad weather.

The foghorn at Souter, while offering a good source for the key sound used in the composition, is less visually interesting than single horn installations elsewhere along the coast. The Ardnamurchan Point foghorn is a single, long red horn installation, placed over a broad expanse of broken rock, which provides a much more engaging visual source. I therefore decided to use the Ardnamurchan Point foghorn as the main image in the composition.

Both the audio and visual components were established early in the work in their full normal aspects, but were then gradually broken down over the course of the work to their essential elements, echoing the approach Mondrian took with his series of drawings of the sea and apple trees.

[233] The Diaphone Fog Signal. Description by Terry Pepper. http://www.terrypepper.com/lights/closeups/fogsignal/diaphone/diaphone.htm [accessed 15/08/2015]

Audio Composition

As the main sound object featuring as the foreground image in the composition, the foghorn is the key to the entire work. From the recordings that I had made, I selected the best versions and ensured all the extraneous background sounds were cleaned up. I also wanted to create a fully realised, sonically full background over which the foghorn would feature. I had built up a library of recordings of the sea and related sounds, such as seabirds and wind, which I would draw on to create a fully -textured, complex background in the composition.

The initial decision I had to make was the delay between recurrences of the foghorn blasts. From my research, I knew that every foghorn had its own pattern of blasts and silences – ranging from blasts that lasted for only 1 second every 10 seconds, through to ones that lasted for 6 seconds every 45 seconds. Souter's pattern was a 5 second blast every 30 seconds, but this pattern did not really give enough time between recurrences for the sonic elements of the composition to be developed satisfactorily, so I decided to lengthen the delay between blasts to 45 seconds. I had set myself a time limit of around 5 minutes, so with a short 15 second introduction, a set of 7 repetitions of the foghorn theme would result in a piece lasting at most 5 minutes and 30 seconds.

As I wanted to establish the sea elements quickly in the piece, I took a recording of waves breaking against the walls of Fingal's Cave on the island of Staffa, which provides a good strong resonant setting establishing the presence of the sea from the start. I faded this in rapidly from silence to place the listener firmly within the scene. I then introduced the first version of the foghorn sound. At this point, I wanted the foghorn to be present, but in the form it would have most often have been heard – at a distance, as if from a boat out at sea in bad weather, or listening from a seaside cot a mile or so from the lighthouse.

In the first interval after this establishing sound of the foghorn, I wanted to start reducing the full texture of the sea waves to a harsher, grittier examination of the textures in the sound. I applied some distortion filters to the recording, reducing the amount of higher frequencies in the mix, and making the sound granulated. This lent a slightly ominous feeling to the soundscape, which I developed further by also distorting the next iteration of the foghorn sound, now presented close up to the listener, and filling the whole audio environment with its presence.

For the second interval, I maintained the distorted, granular seascape as the main background, but introduced a couple of foreground sounds to add interest as the piece developed. I had been able to get very close to a puffin on one of the Treshnish Islands, and had recorded its call. This is a surprisingly low growl, which rises and falls both in pitch and amplitude over its course. It is an unexpectedly sinister sound, and therefore fitted in well with the darker textures in the distorted sound of the waves. I also pitch shifted the call to drop it by just over an octave, resampling it to increase the duration and removing it slightly from its origin. The second foreground texture was provided from a recording of my wife walking across the Singing Sands on the north coast of Eigg. These quartz sand granules squeak against each other as you put pressure on them, resulting in a shirt burst of glassy pitched sounds as you walk across the beach. The original recording held its own interest, but I wanted to start to build a sense of anticipation leading up to the next foghorn iteration, and therefore resampled the recording to speed it up considerably, while also not raising the pitch of the sandy squeaks too much. I also started to fade out the background wave sounds, and stopped the squeaks just short of the next foghorn blast, to heighten the tension.

The third blast began to break the sound of the foghorn down further, in this case by making a clear division between the right and left channels of the stereo recording I had made. As the blast progressed, I panned the left channel from its normal position to end heavily to the left, and vice versa with the right channel. This gives an effect of opening up the foghorn blast around the listener, as if they are moving through it. I continued this emphasis on the separation of the right and left channels as the interval after this blast continued, with the two channels being separately echoed, resulting in slight mismatches between the left and right signals as the echoes recur and start to fade away. Against this, I reintroduced the granular sea sounds, but also with a reverberation effect added, to place the sounds of the sea in a large containing environment, as if the listener was listening to the waves move through a vast underground sea cavern.

The fourth iteration of the foghorn blast also examined the essential elements contained within, this time concentrating on the fundamental frequency and its core harmonics. The blast was passed through a set of comb filters corresponding to the key frequencies in the main harmonic series generated by the foghorn mechanism, and these frequencies were lengthened slightly and added back into the original blast, resulting in a sustained note emerging out of the blast and being held for a short sustain afterwards. Following this sustained note, I wanted to return to a quieter, more contemplative setting for the background texture, allowing some relief from the granular sea textures used earlier. I took another sea recording, this time of a more open seashore at Gullane, and applied a reverberation filter to make the waves sound even more distant, adding to a feeling of calm. I decided to add in another sound element to provide a highlight of interest in the interval, so copied the original Gullane recording and raised it two octaves in pitch. I then applied a phasing effect to the sound, and also small amount of flanging. This gave a high, vibrato effect, with increases and decreases in volume mirroring the original sounds. I then added this in on top of the distant sea sounds, fading it in and out to ensure it added aesthetically to the progression of the work.

The fifth iteration of the foghorn sound built on the open expanse of the seascape with a deep reverberation added to the blast, echoing the sound several times as if caught in a wide open space between cliff faces. This marks a halfway point in the exploration of the quieter background sea noises, with the next interval developing the vibrato sound much more, and introducing a new midrange sound object. This is a reworking of the speeded up singing sands recording featured earlier, but with a band pass filter applied to highlight only the middle frequencies of the recording. The lower frequencies of the main background sea texture are also increased gradually as the interval develops, pushing the low bass elements of the waves into more prominence. Finally, as the end of the interval is approached, the granular texture effect used earlier in the piece is also applied to a copy of the main sea texture, adding another grittier reinforcement to the increased bass sound.

The sixth iteration of the foghorn brings the blast back into dominance at the front of the composition, with a distortion effect applied to increase its apparent loudness within the mix, and to mirror the granular breakdown of all the background elements leading into this blast. The following interval continues the build-up of the distortion effects, with the underlying bass heavy sea sounds complemented by the increased use of the higher frequency phased version. The mid frequencies are also mirrored and a heavy distortion applied, creating the sound almost of heavy winds driving through the seas. A final addition to add more colour to this section is the recording of a noisy stack of seagulls recorded on the Treshnish Islands, which softens the distorted sounds of the manipulated sea recordings with organic notes, but which are still slightly disturbing due to the competition between the vast numbers of birds in the colony.

As with the earlier development of a building anticipation and movement towards a climax, the sound builds in intensity through this interval section up to a sudden drop to complete silence, again heightening the tension and anticipation. The final iteration of the foghorn then occurs, a full blast of the original recording played at maximum amplitude, placing the listener squarely right in front of the foghorn itself, and establishing its final dominance. The composition then fades gently out with a repeat visit to the recording of waves moving through Fingal's Cave, this time using a slightly quieter section of the original recording without any effects added, and fading out gently to silence.

The composition was entirely carried out using original recordings, edited, processed and mastered using Adobe Audition 1.0. This software was quite dated even at the time I started the composition, but I was very familiar with it and the various effects I could achieve using it. I had also considered using ProTools for the compositional work, but the speed with which I could apply effects in Audition and achieve the result I wanted, rather than having to experiment much more in ProTools, steered me towards the older package. I later used ProTools in remastering the original composition to better balance the right and left channels, but added no further original material using it.

Video Composition

The audio composition had created quite strong visuals in my head as I was creating it, especially as I had the theme of deconstructing the main visual element of the foghorn at the forefront of my thoughts as I worked. The main visual inspiration would be the image of the foghorn at Ardnamurchan Point, which I had photographed from several angles. The key photograph which I decided to use as the main image gave a good side view of the foghorn, showing the building housing the main mechanism, the curved large red iron horn protruding from the top, then bending over and facing out to sea, and the flat apron in front of the apron bordered by grey metal railings.

Ardnamurchan Point Lighthouse Foghorn with view out to sea

To create the majority of the visual part of the composition, I decided to use a 3D CGI animation package. I had previously used a 2D animation package (Flash) on an earlier work, which I considered using again, as the deconstruction of a starting image along the same lines as Mondrian had employed in his works could be realised quite simply using a 2D environment. However, I wanted to both show the deconstructed images and the mechanism of the deconstruction itself, with the main image being broken up and movement into and around its elements featuring as a key narrative element in the work. For this, a 3D package would be more suitable. I had made use of Blender, an open source 3D animation package for the work previous to this one, so decided to use it again as I was familiar with it as a visual construction environment.

As I had created a short 15 second 'introduction' leading up to the first iteration of the foghorn blast, I decided to use this part of the composition to create the introductory title, fading in from black to reveal the main title "Study No. 3", with the subtitle “Foghorn Abstraction” fading in a few seconds later. The two titles fade to black just before the first foghorn blast.

With the first blast, I wanted to confirm the fact that the sound was from a foghorn, and so decided to show the original photograph used as a template for the main 3D model created in Blender. I therefore made a careful 3D equivalent model of the original foghorn in Blender, taking a camera view of it that would correspond with the image shown in the photograph. I then took a frame from Blender showing that view, and combined it with the original photograph using an open source image morphing package, morphing slowly over the full 45 second interval to the next foghorn blast, to establish that we were moving from the real world into a virtual environment. The correspondence between the real foghorn in the photograph and the 3D model was not exact, and the morphing software was not capable of producing a high quality result, so the morphing sequence contained quite a few visual artefacts. However, this worked well with the gradual introduction of granular sounds in the background sea textural sound, and so the imperfections did not detract from the combination of sound and vision.

The second iteration of the sound blast marked the point where I had fully morphed from the original image of the foghorn to the 3D model. Having moved into the virtual world created in Blender, I could manipulate the foghorn much more easily. I pulled back from the close-up view of the foghorn to reveal a constructed environment around it, then made use of the implied motion through the rising then falling pitch of the elongated puffin sound to carry out a camera move up and over the foghorn. The camera finished at a position looking directly into the horn of the foghorn itself, where the interior is hidden in black shadow. I wanted to give the impression to the viewer that we were entering the interior world of the foghorn, so at this point the camera moves forward and into the horn, moving into a totally black screen. As the second foreground sound feature develops, the visuals fade back in to reveal that the viewer is now within the horn, moving forward, up and around the angles of the pipe as it moves further back into the body of the main machine. The flat red colour of the foghorn as seen from the outside has been developed further here with a texture combining the main red with highlights of orange and yellow, allowing further visual interest to engage the viewer and reflecting the increased texture in the background sounds. The speed of the camera movement increases along with the growing speed of the speeded up footsteps on the singing sands, along with tighter curves being negotiated along the pipe, up to the climax and sudden drop of all sound, where the pipe ends in another black field, and the camera passes again into a fully black screen.

The third foghorn blast is matched visually with two fields of arrays of haloed light sources, each moving to the left or right side of the screen in line with the increased panning of the two stereo channels. As they reach the side of the screen, they fade away. This visual action is repeated as the echoes of the foghorn blast continue, with the two fields keyed to the appropriate channel, and therefore taking different amounts of time to move back and fade. At the same time, the underlying sea sound textures are echoed by a visual sea of point light sources, moving in layers up and down in sympathy with the ebb and flow of the wave movements in the sound. A final visual element in this section is a recurring set of long, straightened red horn elements of the foghorn, long cones that hang in the higher portion of the visual field and repeat off to the sides and into the infinite distance, again reinforcing the idea that there are echoes of the main elements throughout this section. The background to the entire section is a deep black sky, also adding to the use of echoes to imply the viewer has entered a vast reverberant cavern. As this section finishes, in anticipation of the next foghorn blast, a small central light source appears, moving rapidly towards the centre of the frame.

As the fourth foghorn blast takes place, the central light flares and expands across the entire frame, filling it with white light. This light dissipates to a central source again, now outlined with a halo, which then fades slowly as the harmonic pitched note emerging from the blast fades away. Behind this light, a long rolling ocean is revealed, with a very pale blue sky above it. The ocean is actually constructed as a series of long cylinders placed beside each other, each with a gradient of blue tones across the surface. The cylinders are constantly rotating towards the camera, each starting from a slightly different initial position, so that a small variation of the blue colour tone occurs across the expanse reaching out in view to the horizon. The impression of an endless ocean is reinforced by the camera movement, slowly but inexorably forwards across the long expanse. As the ocean is established, the next two main visual objects of interest also start to make an appearance – two cones, the straightened horns featured in the last section but here turned vertical rather than horizontal, with their larger open ends facing upwards, start to rise out of two fixed points in the ocean, midway between the camera and the horizon. As they rise, they begin to rotate in opposite directions, showing that they have a similar textural marking to the interior of the foghorn shown in an earlier sequence.

As the camera moves forward towards the two rotating cones, the sky darkens and lightens in sympathy with the higher phased frequency foreground sounds. The darkening skies also reveal a scattering of stars across the field, again building the impression of a vast open ocean.

At the point of reaching the two rotating cones, the fifth foghorn blast is heard. At this point, circular clouds erupt from the open mouths of the cones, followed by an array of coloured materials taken from the original 3D model of the foghorn – the cylindrical grey pipes that formed the railing around the foghorn building, and several crumpled sheets of materials, in the primary colours of red, blue and yellow, reflecting the red of the foghorn itself, the blue of the sea it faced, and the yellow highlights of the concrete base in front of it. These constituent elements glide through the air until they reach a set arrangement in front of the camera, then revolve around each other in this position as the camera continues to move forward. At the same time, the skies have continued to darken and reveal the star formations in time with the higher frequency foreground sounds.

As the sixth foghorn blast is approached, the foghorn elements suddenly break away from their fixed rotating positions and move rapidly upwards and out of the frame. The blast corresponds with the explosion and rapid upwards movement of a wide cliff in front of the camera, breaking up out of the ocean and rising up, again up and off the top edge of the frame. As the camera continues forward, it also angles up to reveal the folded contours of a grey cliff, with the foghorn elements now broken out of their revolving pattern and moving at different speeds upwards and parallel to the cliff face. The camera follows these elements, shaking from side to side on occasion as the distorted wind noise gains in emphasis. Some of the incidental sounds from the seabird stack and the general sea sound environment are echoed with visual interactions, such as the yellow material bouncing against one of the grey pipes.

As the foghorn elements reach the top of the cliff, they re-orient themselves from their vertical orientation to the horizontal, then rapidly move forward over the cliff and out of sight of the viewer. The camera continues to rise upwards, pointing vertically upwards past the top edge of the cliff, and still shaking heavily in line with the increased wind noises. As the music rises to a climax then comes to a brief stop, the camera position has moved past the edge of the cliff and only shows the star-lit night sky.

Just as the final foghorn blast occurs, the camera rapidly tilts forward to reveal, not the top of the cliff and the foghorn elements that we might have expected, but an immense foghorn cone suspended in space. The piece ends with the camera moving forward into this cone, the surrounding red and orange texture of the cone interior gradually fading into shadow, and eventually into black. As the final sea sounds fade away, the credit sequence is shown, with the piece finally ending at 5 minutes and 30 seconds.

The first version of the visual composition ended with a reveal of a stylised version of the original foghorn photograph, converted into a geometrical abstraction equivalent reminiscent of Mondrian's famous work showing black lines on a white canvas with red, blue or yellow panes. However, this did not work as a satisfactory image, and the final huge foghorn cone and the implication of a further voyage through it was a more satisfying sequence on which to finish.

Audiovisual Work

The work was originally shown at the Soundings International Festival of Sonic Art at the University of Edinburgh in February 2011. This version had the original ending which featured the mock-Mondrian portrait of the foghorn photograph. The initial reaction to the piece was favourable in the main, but the ending was not well-received. The balance between the left and right stereo channels was also not properly realised in this version of the work, an imbalance which was highlighted by diffusing the work through the multiple speaker array used for the concert.

The ending was immediately corrected, with a couple of images substituted until I settled on the giant foghorn cone as an acceptable ending. This version was shown at the Edit Point concert in Glasgow a month later, although the stereo balance problem was still present. This was eventually solved using ProTools to remaster the original soundfile produced from Adobe Audition, which resulted in a much more balanced, fluid stereo mix.

Some of the problems that had occurred in the earlier version in terms of the audio balance had arisen from the choice of codec and compression for the video container in exporting the files from the video editor I had used for the original mix. This was another open-source editor, written for the Linux operating system. While the exported files worked reasonably well on the Linux platform, there were noticeable glitches in both the image and sound streams on playback on PC or Mac platforms. This problem was resolved by re-compositing the soundtrack with an exported frame sequence from the original Blender files using the Adobe Premiere Pro package and the H264 codec. This produced a much higher quality image, and removed the final imbalances in the stereo mix.

Reflections on Completed Work

The final version of this piece has been quite successful in terms of audience reaction, and in performance, having been accepted as one of the inaugural exhibition pieces in the Byte Gallery for audiovisual work at Transylvania University in Kentucky, USA, during 2011, and for performance at the Toronto Electroacoustic Symposium in August 2012. The work has been performed as both a fixed media piece with a straightforward stereo playback, and as a diffusion piece, where I have diffused the soundtrack against the projected image. Of the two, the diffusion piece allows more freedom to tailor the sound elements of the work to fuse with the visual elements, based on the differing acoustic properties of the performance venue, but a standard playback of the piece also works quite well.

In terms of my original intent, to create a continuous audiovisual equivalent to Mondrian's sequences of gradually more abstract, deconstructed images, the work fulfils the intention to a large extent, but not quite as fully as I had originally intended. My original idea had been to deconstruct the sound and image of the foghorn much further, in sonic terms to the sole fundamental frequency of the blast, and visually to the lines defining the foghorn edges, and its primary colours of red and white. However, in practice, the requirement for the overall composition to retain an ongoing engagement with the audience meant that I needed to retain a bit more structure and narrative progress than a simple full deconstruction. The piece that resulted achieves this engagement, but still addresses the original aim of exploring the essential elements of the audiovisual scene first presented to the audience at the start of the piece.

On repeated viewing, there are some elements I would like to revisit and correct, which are almost all in the visual portion of the work. This reflects my greater expertise with sound and audio composition at this stage of my research, and the need for me to develop equivalent skills in 3D animation. The starting morph between the foghorn photograph and the 3D mode contains visual glitches – while luckily the granular texture of the accompanying sound allows a certain amount of synchresis to take place, allowing the audience to connect the moving lines and blotches on screen with sound objects emerging within the granular background sound texture, I would have preferred a smoother transition into the virtual world, slowly immersing the audience into the new reality focussed on the foghorn and its surroundings. The other visual elements that I find less satisfying include the initial 'smoke' emerging from the upturned cones as the fifth foghorn blast occurs, and the implied turbulent sea 'foam' breaking across the base of the sea cliff as it rises upwards with the sixth foghorn blast. In both cases, the end result is much thinner and less convincing than my intention. This again reflects a lower skill level in the 3D animation field, and highlights my need over the remainder of the research to develop this skill further, to fully realise the creative practice that reflects the ideas and research behind the subject.

Despite my reservations about some of the animation, I do feel that the work is coherent, has a sound narrative structure, and achieves most of my original aims in applying a Mondrian inspired breakdown of a strong audiovisual object into its constituent parts, again as individual audiovisual objects rather than separate perceptual streams. I regard this piece as my first proper examination of the principles of audiovisual composition, consonant with the research being carried out at the same time, and so it marks a growing confidence in my practice as a creative audiovisual composer.

Performances

Soundings International Festival of Sonic Art, Edinburgh, UK	*Oct 2012*
Toronto Electroacoustic Symposium, Toronto, Canada	*Aug 2012*
KOMPAS Exhibition, Edinburgh College of Art, Edinburgh, UK	*Jun 2012*
Edit Point Concert, Glasgow, UK	*Mar 2012*
Soundings International Festival of Sonic Art, Edinburgh, UK	*Feb 2011*
Inaugural Exhibition, Byte Gallery, Transylvania University, USA	*2011*

Conclusions

Study No. 3: Foghorn Abstraction is the first audiovisual piece I have composed where I have achieved most of the objectives I had set out in its original concept. The inspiration from Andrew Lewis's *Penmon Point* in using a recurring sound object as a construction technique to develop a narrative allowed a coherent structure to underlie the piece, while the inspiration from Mondrian to examine the elements of a scene by deconstructing it to the bare essentials provided the impetus to drive the piece forward.

The piece also highlighted my growing familiarity with animation and 3D CGI software, but a few mistakes and glitches are very obvious to me in revisiting the work. This highlighted my need to gain more experience in using this type of technology. At this point in my research, I was very fortunate that the Edinburgh College of Art merger with the University of Edinburgh took place, leading to a change in my supervisory team. I gained Jared Taylor, Head of Animation, as my secondary supervisor, and his advice on visual technique and experimentation would assist greatly in addressing the animation practice in which I needed to gain more expertise.

Study No. 3: Foghorn Abstraction also marked the first piece that gained true international recognition – while my previous work, *Study No. 2*, had been selected for performance in the Listening Room at the International Computer Music Conference in 2011, this conference took place in the UK, and the Listening Room lacked the full exposure for assessment and critique that a concert performance would have afforded. The performance of *Study No. 3: Foghorn Abstraction* at the Electroacoustic Symposium in Toronto allowed me to present the work to an international audience of experts in the field, and the feedback received was of great use in informing my ongoing practice.

Study No.3: Foghorn Abstraction

Diffusion Score

Time	00:00:00	00:10:00	00:20:00	00:30:00	00:40:00	00:50:00
Vision	Blank Screen – Titles.	Titles Foghorn photo at 0:15	Photo morph to 3D model			Foghorn blast.
Notes	Treble - mid Dist - high Mid Dist - mid/high Main - mid Wide - low mid Bass - high (rest off)	Move from dist. & mid dist. to close front & close rear. Bring wide & Main down				
Time	01:00:00	01:10:00	01:20:00	01:30:00	01:40:00	01:50:00
Vision	Full 3D model Zoom out Puffin over + into horn	Interior horn	moving through horn	up to black exit	Foghorn + echoes	
Notes	Close in Bring rear up in line with puffin	Keep very close Bring up sides		WATCH! for end	QUICK Main & wide up	Bring close down & out
Time	02:00:00	02:10:00	02:20:00	02:30:00	02:40:00	02:50:00
Vision	Echoes continue		Light appears @ centre	Foghorn - note resonance Movement across ocean		
Notes	Keep flat front Bring mid & Mid dist. up to med.			Develop dist & mid. dist gradually bring sides in & up.		

Time	03:00:00	03:10:00	03:20:00	03:30:00	03:40:00	03:50:00
Vision	Ocean ⟶	| Foghorn blast elements erupt from cones		Central elements rotate		WATCH for cliff rising!
Notes	Keep flat distance	Flat front Some side Ride tweeters up & down ⟶				WATCH for next blast
Time	04:00:00	04:10:00	04:20:00	04:30:00	04:40:00	04:50:00
Vision	Foghorn cliff rises from sea	Follow elements up side of cliff – camera shake ⟶	//Clouds of yellow piece & tubes		Camera pause | final blast Tilt	| Move into cone to reveal cone cone turns to face camera
Notes	QUICK bring up front close	Build up rear close & sides	Wiggle up right close for clouds	WATCH for pause	Full up on all mains, wide QUICK!!	Move from mains to distance, all close out
Time	05:00:00	05:10:00	05:20:00	05:30:00		
Vision	Move into cone fade to black Titles ⟶			Out to black || end.		
Notes	Continue developing mid dist & dist	⟶	Only dist.	| END		

Production Notes

Skitterling (2012)

Skitterling (2012)

Introduction

Skitterling was originally intended to be a piece investigating the use of foreground and background in both sound and visual composition. From my research into the audiovisual object, I had determined that the use of foreground and background layers forms an important part of defining audiovisual objects in a composition. My original design brief for the work laid out my plan to create a piece where sound and visual objects would emerge from background to foreground, both in the separate perceptual streams and as conjoined audiovisual emergent objects. This would be an investigation where the planned conjunction of foreground emergence in both sound and vision could be compared against any synchretic emergence of an audiovisual object where only one perceptual stream had been deliberately designed to feature foreground emergence.

In my electroacoustic composition, I had developed my proficiency to a level where I would deliberately limit the source materials, maintaining a consistency in source sounds across the piece, while exercising my creativity in developing interesting material from my examination of the sound spectromorphologies. As a result of my conversation with my secondary supervisor, Jared Taylor, I decided to limit my use of colour in the visual part of the work in a similar manner, to achieve a greater level of consistency across the piece, and to test my skills in developing a piece with visual interest and contrast without relying on an extensive use of different colours.

In my previous work, I had avoided creating long phrases or sequences of sound. This was primarily due to the nature of my audiovisual work - as the animation takes so much time and effort, the longest work I had created so far had been 5 minutes 30 seconds in length, and I had been concerned that the audience would lose interest without a great deal of ‘action’ both sonically and visually, so I had planned my compositions to develop and move on from one sound scene to another quite quickly.

However, as part of my research, I had curated audiovisual concerts where I had issued calls for contemporary works, which allowed me to assess and critique current compositional techniques in this field. From these contemporary works, I had developed an appreciation for the work that took time to explore a sound more thoroughly, as well as allowing the accompanying images to develop and draw the viewer in to the unfolding scene, particularly where the complexity of sound and image developed over time. I decided that with this work, I would aim to give the sound and vision time to develop satisfactorily, and resist the pressure to move on quickly to maintain an audience's attention – I would trust the audience to become absorbed in the audiovisual world I was creating.

In order to limit both sound and visual palette, I decided to concentrate on a single simple substance – sand. This would allow me a variety of associated sound objects to work from, as I could use the shifting sound of sand sliding, as well as the sound of 'processed' sand, such as glass or sandpaper. In addition, visually I could make use of sand moving in patterns, both random and organised. Visually, I would limit myself to yellows, browns and ochres, which would also match in with my determination to keep to a limited colour palette.

Concept

As an initial exercise, I rapidly jotted down some thoughts on how a possible piece could take shape, using yellow, brown and red colours, the colours of sand, as basic visual building blocks, with shifting sand noises as the main soundtrack. A resolution of the sound and vision would occur at the end of the piece. These notes are reproduced below:

- Black, fade in grey.
- Resolve to horizontal wavy lines
- Zoom in gradually

- Reveal lines are sand waves and dunes
- Camera lands, site disturbed and sand spurts out from drop zone
- Start moving through sand
- Local small sand swirls overtaken by larger sandstorm into firestorm
- Light haloes appear in firestorm
- Haloes become bright metallic spheres in space
- Harmonics
- Minor third creeps in, colour changes and dissonant sphere enlarges to consume others

I developed these initial thoughts by creating a design brief and a storyboard, where a basic timeline showing the visual and audio progression of the piece was mapped out. The design brief and storyboard are included in Appendix 1.

The storyboard narrative developed both the audio and visual composition of the overall work, fleshing out the original basic ideas into a more coherent, planned full work. This was not something I had done with earlier pieces. My normal method of working was to create the sound composition first, then create a visual animation in response to the images the sounds inspired as I listened to them. In this case, as I deliberately wanted to plan for synchronisation between sound and vision in advance, I decided to plan the piece more thoroughly from the start. However, after discussing the storyboard with my supervisors, we all agreed that this approach seemed very restrictive, and I reverted to my normal practice.

Audio Composition

As per my initial thoughts, I decided to restrict my sound objects to two main sources. I wanted to achieve a contrast between the white noise of sand movements and a more pitched, melodic sound source. I recorded the sound of several different types of wine and whiskey glasses being hit with my fingernail, or having the rim ‘played’ with a wet finger. I then started to record the sound of sand being poured from one container to another. However, the sound recorded was quite unsatisfactory, as it consisted mainly of a simple bell curve envelope of white noise, with little content that could be manipulated or processed in an interesting fashion.

I therefore started to experiment with the sounds of larger grains being poured from one container to another. The sound of rice being poured between two large containers made out of bamboo held some promise, although it too provided too homogeneous a sound. I then recorded the sound of rice grains being poured within a rain stick, which gave a much more interesting sound source, which combined both the white noise of the shifting grains with some sharp, more pitched sound values, which I could blend in to the much more pitched sounds recorded from the glasses.

These two sources provided virtually the full set of sound objects then used in the composition. The pulsating sounds achieved by rubbing the top of the glasses were timestretched and resampled, to create low pulsating subharmonic noise to layer in as a base to the whole work, immediately adding an unsettling mood to the composition. The sound of the glasses being hit was also timestretched and manipulated, resulting in some tones similar to churchbells, and others of clusters of glassy, tinkling fragments.

I decided to work to a three stage structure for the composition. The first would concentrate on the bell-like tones from the glasses, building up from single separate sounds to a mass of interacting chimes. My initial approach to this was to create a collection of bell tones, each slightly higher than its predecessor, adding in higher and shorter pitches with decreasing intervals. I made several of these sequences, then layered them against each other. I then attempted to create a feeling of increasing movement and urgency by resampling the combined sequences using an increasing curve to compress the end of the combined sequence more than the start. While this did result in an increase in pitch and accumulation of bell tones over time, it was very noticeable that the initial, longer notes were discernibly gliding up in pitch over time, which was not the effect I wanted.

I then turned to the Composer's Desktop Project (CDP), which had been demonstrated in a workshop by the leading electroacoustic composer Trevor Wishart a few months prior to my work on this composition. After a few trials and errors, I found a way of using CDP to satisfactorily increase the pitch of consecutive bell sounds, and randomising the intervals between them so that the sounds would occur much closer together as the sound continued. By creating several of these sequences in CDP, I was able to achieve the complex increase of tones and pitches I had originally envisaged.

The second stage continued to explore the sound of the struck glasses. For this stage, I took the recorded sound of a red wine glass being struck and reversed it, which gave me a gradually building pulsed tone up to a climax, the reversed sound of the strike. I created several versions of this sound by resampling it. Each version was deeper in pitch and stretched from the original, and were pitched at successive octaves to each other, keeping the whole set in the same harmonic series as the unsettling subharmonic pulse underlying the whole piece.

These resampled sounds were then layered against each other to create a gradual series of interlaced sounds, moving up in pitch as the piece continued, to echo the earlier increase in pitch of the individual bell-like sounds in the first stage. The reverberations in the glass created a varying amount of panning between the left and right channels, which I planned to use as a source of tension in the visual composition.

The final stage made use of elements from the first two stages, with the primary driver of the sound being a long reversed sample of a glass being hit, again to give that feeling of increasing tension up to the climax of the reversed strike. This was combined with several unreversed strikes resampled across a related set of harmonic pitches, with the strike climaxes matched, allowing the sound to have a natural decay after the main climax.

At the same time, the underlying subharmonic pulse was overlaid with a slightly higher sample, creating more phasing and panning to accentuate the building tension. Moments of relief to offset the growing tension were provided with higher pitched reversed strikes, as a reminder of the second stage, and some of the higher pitched rising bell sounds from the first stage.

After the main climax, the glass/bell sounds were allowed to decay normally, leading to a natural fading end to the work.

Having established the main sonic sequence from the glass recordings, I turned my attention to the recordings of the rain stick. The main sequence had established the overall structure of the sonic work, but lacked contrast and interaction between competing sounds. The rainstick sounds would provide this contrast, introducing a noisier, less pitched element into this sonic world.

I edited and combined the rainstick recordings into one long source sound, which was slightly shorter than the main sequence. I resampled the rainstick source to match the length of the whole composition and matched the rainstick sound to the main sequence. I arranged the rainstick sounds to create a succession of interesting interactions with and contrasts to the main glass sequence. I also expanded the panning of the original recording, which, due to the quiet nature of the rainstick source, had been quite close up and narrow in focus.

Although the positioning of the rainstick sounds was satisfactory, the sounds themselves needed to show some development as the composition continued. I applied a pitch shift to the rainstick sample, dropping it in pitch by one and two octaves while keeping the duration exactly the same. I also convoluted the rainstick sample with the sound of the wine glass being hit, again at different octaves, giving another series of samples at the same duration, but this time with a more pitched base sound, which would tie in with the main sequence.

A final version of the sound was created, where I applied a granular filter and then a distortion effect to brighten and 'roughen' up the texture of the individual rice grains passing through the stick.

The processed rainstick samples were then layered against each other. The layers were mixed to develop the sound over time, moving from an initial 'pure' version of the original sample into a wider and deeper sound through the introduction of the pitch shifted samples. As we moved into the second stage of the overall composition, the convoluted sound became more dominant, coinciding with a quieter part of the rainstick sequence. This complemented the quieter and more contemplative part of the main sequence, with the more pitched grains of the rainstick texture blending into the background. The final stage made more extensive use of the distorted, granular sound, adding to the tension and build towards the climax. A very final short sequence reverted to the original pure sound, giving a short reprise of the contrasting granular sound against the smooth fading decay of the bell-like glass sounds.

The resulting sound composition was slightly longer than my original aim of 6 minutes at 7 minutes 20 seconds. Although I knew that this would add at least a fifth more time to the animation process, I decided the additional work would be required to allow the audiovisual world to develop satisfactorily – any major reduction in the sound composition would weaken it, while the longer duration would also test my abilities to engage and maintain interest in the visual composition as well as the audio component.

The resulting composition was discussed with my principal supervisor, Dr Robert Dow, and criticised for quality and consistency. Robert suggested several techniques to tidy up and rebalance the overall composition, with an emphasis on maintaining quality by applying EQ to the individual tracks within the work. I exported the material from my original assembly DAW, Adobe Audition, into ProTools, where I EQ'd and adjusted the stereo balance on each track. The resulting mix was much cleaner and detailed than the original work, and allowed progression to the next stage, the creation of the animation.

Video Composition

Having completed the audio composition, I spent a few days listening to it, making notes on the mental images it inspired, and noting the exact timings of the key audio events within the piece. I also discussed the composition with Jared Taylor, and his initial impressions of the piece. He agreed that the overall tone was unsettling, and likened the experience to walking along a shingle beach and finding a rotting corpse of a sea bird or similar, with insects crawling across it.

This impression linked in closely with the images I had visualised while listening to the piece, of a deserted landscape, with growing spores / mushrooms, being attacked and consumed by skittering insects.

Having determined that the visual composition would develop this theme of a bleak landscape, with smaller scuttling animals crawling around and consuming some central body or bodies, I started to consider animation techniques and packages. I considered using real-life footage, working quite literally to the sort of image suggested by Jared and using timelapse photography on a decaying animal corpse on a shingle or stony background. However, this had some immediate drawbacks to consider – the decay and dissolution of the corpse would be best served in an outside location, but there would be the protection of the camera from theft and the weather elements to consider. Also, the light would be inconsistent, which would not necessarily detract from the captured footage, but would likely require extensive grading. Finally, there was no guarantee that the corpse would attract suitable scavengers, or that they would move in a manner consonant with the composed sound, which might then require additional animation on top of the recorded image.

Moving on from this approach, I then considered a ‘traditional’ cel frame animation approach, where I would draw successive frames to match the audio composition. The use of cel animation would allow the use of several layers. While this would allow complete control of the image and the correlation between sound and vision, the number of frames involved would be excessive to allow completion of the work within a reasonable timescale.

I therefore decided to make use of computer animation, and as I had developed a good working knowledge of the open-source package Blender, I decided to use it as the main animation vehicle. My initial thought was to create the bulk of the animation using Blender, but to add some effects using a more specialist package such as After Effects. In the event, I managed to create the effects I needed within Blender, saving some assembly time, which was essential in order to complete the final render within my set time limit.

As per my decision to limit the visual palette, and in keeping with my original ideas about basing the sound and vision within the idea of sand, I created a background image for the base landscape that would be seen in the animation. As I knew I would be working in a standard 16:9, 1920 x 1080 pixel format, I created a source image for the ground using the graphic package GIMP – this image was 4000 x 2250 pixels, allowing for camera movement around the ‘set’ that would be built within Blender. I also decided to start with a very yellow/brown sandy colour to begin with, but that as the piece developed, I wanted the ground itself to grow more menacing, gradually moving to a deeper, duller red colour. I therefore adjusted the colour balance of the original image to create a series of images ranging through to a predominantly red field, and used a morphing package to create an 11,000 frame animation gradually shifting from the original image to the deep red image, with the intervening key frames placed to correspond with the main stage sequences in the audio composition.

This morphed background animation was then mapped onto a large simple plane in the Blender environment, which was then itself subdivided randomly. These random subdivisions were then linked to a wire-frame armature that was deformed vertically, creating a landscape of rills and small mounds, ideal for the semi-desert feel I was trying to achieve.

Having created the background, I then turned my attention to the foreground elements of the work. As my initial design brief had envisaged growing spores and tendrils as a foreground interest, I decided to continue with the idea of growing spores, which would match with the initial soundscape of ringing bell-tones. A random scattering of small spheres was created to rest on the landscape, again making use of a limited palette in colour, in this case a lightly textured yellow/orange. I decided that the spheres would start to emit light in sympathy with the bell-like tones, with the specific spheres emitting light chosen to match the spatialisation of the sound.

Several test versions of this idea were run, experimenting with the light emission settings of each sphere. However, the effect was unsatisfactory – at levels where the emission could be clearly seen, it dominated the sound, while lower emission levels made it difficult to identify specific spheres within the greater mass on screen. I decided to add to the emission effect by making use of particle effects. I created a small swirl of bright sparks which appeared out of the spheres in conjunction with the light emission / bell tone, which added enough emphasis to the light emission that lower levels could be used, allowing a more satisfactory blend between sound and vision.

As the soundscape of the bell tones increased in pitch and quantity, the light emissions of the spheres were also increased, both in frequency and in intensity. As the frequency of bell tones reached a level where individual tones were hard to place spatially, the spheres started to light up randomly, allowing a certain amount of synchresis between sound and vision to occur. In addition, where I had included a rising ‘tinkly’ phrase of chiming tones, clouds orange and yellow sparkles rose from the central spheres towards and past the camera.

My initial thought for the skittering sound was to create small insect-like objects that would scurry across the landscape, occasionally swarming over one of the spheres, gradually consuming them and reducing the larger collection to a smaller, more central population. I designed a small, six-legged insect analogue, with a continual 'walking' animation moving the legs in a suitable sequence. However, the walking movement was too regular for the 'skittering' motion I wanted to achieve, and the processing demand for the number of insects I would need to create a convincing swarm was prohibitive for the systems I was using. I therefore abandoned this approach quite quickly.

As the particle effects had proved useful for the spheres, I decided to make use of particle effects again, in this case to create a gradually expanding cloud of darker particles moving across screen, to give the impression that there was a creature at or just below the ground surface, raising a cloud of dust and exudates, which would accumulate and mix in the air as the spheres were attacked and overwhelmed. The colour palette was created using very dark red, orange and brown, with the occasional lighter yellow particle, and the particle emission rates were set to allow the projected creature to move at a suitable pace across the surface, matching the skittering audio sounds.

At the peaks of the skittering sounds, the particle trail would coincide with a suitably isolated sphere, enveloping it and dispersing to show that the sphere had been fully consumed.

The combination of flickering spheres and skittering insect trails created a visual tension to mirror the audio tension between the pitched bell-tones and the more granular rainstick sounds. While the gradual darkening of the ground also picked up on the underlying tension of the subharmonic pulse, the piece also required some further visual movement to build tension and the feeling of drawing closer in to the perceived action. I created a pathway for the camera viewpoint to follow throughout the piece, starting from a position reasonably wide of the central field, and following a spiral path on the horizontal axis, with the camera focal point following a similar, tighter spiral movement. At the same time the camera followed a vertical path, broadly rising higher in the air as the piece progressed. The planning for the camera positions can be seen on the scanned documents included in Appendix 2.

Moving into the second stage of the audio composition, as it concentrates on the bell-tones with the skittering sounds having much less importance, I decided to concentrate visually on the central grouping of spheres. I also revisited the pulsed phasing on both the left and right channels for the main reversed glass strike sounds. I mapped out the peaks and troughs for each channel, and matched each to the specific frame it corresponded to (see Appendix 3).

I singled out the most central sphere on screen for specific attention – this had been labelled Bell 14 at the start, out of the original 80. I matched this to the left channel, while the other remaining spheres were matched to the right channel. As the signal peaked, the sphere(s) would emit more light, and return to their darker original state as the signal hit a trough. This gave an interesting contrast between the central sphere and the others, which was accentuated visually by gradually increasing the size of the central sphere, signalling that it would become the important visual element in the final stage.

As the second stage ended and the third began, the audio included a burst of very low 'interference', high in white noise. Visually, I decided to represent this by emulating an 'earthquake', incorporating both camera shake to represent the shockwaves that would hit a real camera in such circumstances, and also by opening up a chasm in the ground which would fill with a lava analogue – adding an extra item of interest in the background, and creating an 'island' to frame the central expanding sphere. This also isolated the central sphere from the resurgent skitterlings, giving a logical explanation why the central sphere would not become another victim to the skitterlings, especially as the larger size should make it more attractive to attack.

The creation of a chasm and a lava infill proved quite problematical. I started to create the chasm by further manipulation of the wire-frame armature linked to the base ground, but the resulting changes were very jagged and jerky in animation. Despite adjusting the keyframe movement of the armature in several different ways to try and get a smoother movement, I still could not achieve a satisfactory result. I therefore created a duplicate of the original ground, and created a suitable final chasm in this by manipulating the wire-frame. I then achieved a simple transition between the original ground by keeping the duplicate in exactly the same position as the original ground up to the earthquake stage, then 'lowering' the original ground down through the duplicate, creating the impression of the ground falling down and into the chasm.

At the same time, I had created a simple smooth horizontal plane, intersecting the duplicate ground at an appropriate level to create a pool of liquid within the chasm. As I had deliberately created a gradient within the chasm, by lowering the original ground, it appeared as if the lava was seeping into the gap left by the collapsing ground.

My initial approach to the lava was to apply a very convoluted, large marble pattern to an underlying red base, and to create a rapidly juddering motion on the x and y (horizontal) axes, with the aim of suggesting a rapidly roiling surface. Although this was the effect used in a work-in-progress exhibition of the film at one of the University's "Soundings" concerts, it was obviously very artificial, and, in my opinion, detracted from the central interest of the increasingly expanding sphere. For the final version, the marbled texture was replaced. I experimented with importing a specifically designed image, again created using GIMP, but the most convincing surface was achieved by using another highly convoluted texture modifier, based on the Voronoi algorithm rather than the original marble one. I also removed the juddering motion, replacing it with a gradual movement along the horizontal X axis. The resulting pock-marked surface texture and slow gradual flow was much more convincing than the earlier attempt, adding to the background interest while allowing the foreground to remain dominant. It also provided a visual 'echo' of the central foreground sphere as it developed, which gave a more composed, completed visual look to this phase of the work.

In the foreground, the main point of interest was the central growing sphere. As the other spheres were now superfluous, I ‘dropped’ one of them into the lava to dispose of it, while the rest were gradually subsumed into the central sphere as it grew.

I exploited the basic construction of the central ‘sphere’ as an icosphere, with the individual planes making up the surface of the sphere consisting of joined equilateral triangles. Rather than smoothing the growing sphere to eliminate discernible planes, I maintained the polyhedral shape to increase interest. At the same time, I developed the underlying Voronoi texture used as a skin on the sphere, increasing the complexity and distribution of the pattern as the sphere expanded. This provided a satisfactory, engrossing evolving pattern as the piece developed, with the increasing complexity of the texture adding to the sense of an imminent climax within the soundtrack. Although there were some additional sonic features, including reprises of the reversed glass strikes, I decided against adding any additional visual components as this would detract from the central sphere, and divide the audience’s attention when I wanted it to be directed at the main foreground object.

At the same time, the contrast between the glass tones and the rainstick was still present, and the skitterlings made a reappearance. As the central spheres were isolated on a central island, the skitterling trails approached and veered off from the centre, allowing additional foreground interest to exist, but maintaining the central focus firmly on the growing sphere. As the climax approached, the skitterlings ceased to appear, although the final trail is allowed to dissipate across the screen to a much greater extent than in the past, to maintain the threat in the background even as the central sphere dominates.

At the climax, the swelling sphere eventually ruptured, in one frame exploding into a ring of minuscule fragments, which then expanded and gradually fell as the camera dipped back towards ground level. The glowing fragments will eventually fall to ground, with the audience invited to speculate that they will create the next growth of spores, and lead to the cycle repeating again and again.

The explosion is simply the substitution of the original sphere with a wire-frame duplicate, where a halo effect was applied to create clouds of particles based on the edges and vertices of the underlying structure. This duplicate is then expanded gradually as the work progresses, giving the impression of a dissipating cloud.

As the cloud starts to expand, a final reprise skitterling appears to pause briefly in place, then move on, to remind the audience that there is a counterbalance to the expansion of the spheres.

Audiovisual Work

The piece was originally shown as a work-in-progress at the Soundings Festival at the University of Edinburgh on 20th October 2012. The animation process took quite a lot of time, particularly as experimentation was often needed to settle on a satisfactory rendered result. Due to the use of particle effects, the rendering was further complicated as the entire animation had to be rendered in one long sequence, as otherwise the particle generation would result in new clumps appearing at the wrong positions on screen. The final render was used to create an image sequence of jpg files, and took just less than 20 hours from start to finish.

The rendered jpg files were then assembled in a video assembly package, Kdenlive, and combined with the mastered audio. This created a matched audiovisual sequence of 7 minutes 20 seconds duration.

Finally, the titles and credits were created using GIMP and added in as single frames, duplicated to have a reasonable duration on screen, with cross fading to blank screens between the title screen and main animation, then into and between the final credits.

From the earlier discussions with Jared Taylor, we had noted that the majority of student animations tended to use white on black titles and credits, and that he associated this with amateur work prevalent on the Internet. I therefore decided to use black on white titles and credits, to maintain a more professional approach.

There were several aspects of the work-in-progress version which I felt should be corrected, particularly in the animation, and I reworked the piece subsequently, with the final version being finished in February 2013.

The final version corrected some random movements of the spheres in the first stage of the piece, tidied up the expansion and contraction of the clouds of sparkling lights in the first two stages, and substantially reworked the lava flow in the final stage. I also adjusted the duration of the particle generation for the skitterlings throughout, and adjusted some of the audio balances between the bell tones and the rainstick sounds to achieve a better complement to the action seen on screen.

I also reworked and replaced the titles and credits. I removed a reference to this being ‘Study No. 4’, as that presupposed people would know or understand that this was part of a sequence of experimental works that I’ve undertaken as part of my studies. I retained the subtitle of Skitterling, making this the only title. I also separated the original single credit into two, finishing the work with a copyright notice rather than incorporating it into the main credit screen. Finally, I used a slightly scrawled font based on handwriting, to tie in with the idea of skittering, the random insect-like movement used later on screen.

Reflections on Completed Work

The audience reaction to the piece, even as a work-in-progress, was favourable, with several complimentary comments made after the concert. My own reaction was less favourable, as I concentrated on the errors I could hear and see during the performance. I find it difficult to be objective about a piece immediately after its creation, and I usually have to wait a couple of months before I can view the work without immediate self-criticism. The delay in reworking the piece, from October 2012 to February 2013, came from the need to distance myself from it.

With the corrections made to the piece, I can now view it in a reasonably objective light. The adherence to a restricted sound and visual palette has proved to be a strength – the combination of the two creates a more convincing and immersive audiovisual world than adding in extra sounds, colours and textures to achieve an 'easy' highlight. The necessity to use other aspects of sound and visual objects to create interest in the work has led to more inventive techniques, and more emergence of interesting audiovisual objects than I could have achieved using simpler techniques.

I was also satisfied with the longer times taken to develop individual themes within the piece. Although these are still of necessity shorter than the development time I might have taken in a purely audio composition, the slower and more complex development of themes and structures in the piece draws the audience into it much more deeply, and provides a more rewarding overall experience, particularly in the build-up to the climax/explosion. I also think that the final skittering flourish finishes the piece off well, and contains that essential reprise of the most enduring theme throughout the piece.

In terms of my original aim of exploring foreground and background, the piece does achieve this, but not according to my original plan of developing a foreground object into an encompassing background, to then be overlaid and replaced with a new foreground object, etc. I am not particularly worried by this – I think it is an essential part of my creative process to listen to the sounds that have been recorded and develop a coherent composition from them, and to follow a similar process of looking at trial images and developing them to create a sympathetic visual accompaniment to the audio composition. However, I would like to retain the idea of successive foreground/background objects for another work, should the recordings lend themselves to such an interpretation.

I also found it interesting how my concepts developed during the creation process. Many of the initial ideas are still present in the piece – the limited sound and colour palette, rooted in the idea of sand and glass, are still there. The original storyboard had a cyclical plan, with spores and tendrils growing out of a barren desert, with the camera moving along and into a tendril, following spores as they erupt out of the tendril and settle back down to the ground – a cycle that is implied, if not so overtly shown, in the completed work. The storyboard also refers to dust storms and swirling clouds of spores, which implies the use of particle effects. Although these were not used in exactly the same way in the final piece, particle effects did form a key and substantial part of the audiovisual work, particularly for the ongoing skittering presence.

The realisation of the work within Blender also had a specific effect on the visual 'flavour' of the piece. 3D CGI tends to have a specific look to it, which, although it can be adjusted by importing images and textures, gives a more synthetic, plastic texture to the final rendering than a hand drawn cel animation would normally achieve, or certainly more than a manipulated sequence captured using photography of real-life items, as had been considered at the start of the creation process. Choosing a specific package creates restraints, and although my familiarity with Blender allows me to create acceptable images and animations reasonably quickly, it does restrict the visual inventiveness I sometimes would like to apply to my work.

For this specific work, the visual aspect has worked well, particularly with the use of particle effects. I am satisfied that I would not have achieved much better effects by using a package such as After Effects on a more basic initial animation sequence, and the advantage of maintaining a holistic look within the constraints of the Blender package outweighs the identifiable look of the work as being produced in a 3D CGI package.

The synthesis between the audio and visual parts of the work is also in the main successful. I avoided my earlier tendency to mickey-mouse, and have accepted that I do not have to create a visual equivalence for absolutely every sound event, nor do I have to have a sound event present to create a visual object. I was much more willing in this piece to utilise synchresis, particularly in developing the complexity of a visual texture or particle stream in conjunction with a developing audio texture or theme. Having said that, I was also more confident in re-trying textural patterns and movements to achieve a closer match between sound and image, thereby manipulating synchresis to my own advantage.

I would rate the final work as a substantial step forward from my previous works, both in confidence of technique and audiovisual interest for the audience.

The final work was shown at the Sight, Sound, Space and Play conference at De Montfort University in June 2013, with a very favourable audience reaction. Several of the audience members had been present for the earlier concert where the work in progress had been shown, and they were impressed with the changes that had been made, noting that it was much more coherent and immersive as a result.

The work was also selected for performance at the International Computer Music Conference in Perth, Australia in August 2013. This is a leading conference in the field, and acceptance for performance is a good indication that the work is of high quality.

Performances

Black Cube Collective Exhibition, Athens, Greece	*Oct 2014*
International Computer Music Conference 2013, Perth, Australia	*Aug 2013*
Sight, Sound, Space and Play Conference, De Montfort University, UK	*Jun 2013*
Soundings International Festival of Sonic Art, Edinburgh, UK	*Oct 2012*

Conclusions

Skitterling marks two key shifts in my attitude to audiovisual composition – firstly, I started to apply specific restrictions in both the audio and visual fields, which forced a more carefully considered creative method on to my work, given that I could no longer just add any sound, colour or image that I wanted to into my work, I had to keep within my boundaries for source sounds or colour palette. This use of restrictions in sound and vision improved the quality of the work I produced, as much more thought and care was given to every element that was used. Secondly, I relinquished my earlier fears of taking too long to develop themes across my compositions, and trusted in the audience to appreciate the time taken to develop the themes of the piece. This also allowed me to make use of other elements of animation and filmcraft to highlight aspects of my work, such as using camera movements and angles to accentuate specific elements in the visual portion of the composition, which in turn made the associated sound elements take on a prominence to accompany the visual emphases.

There are some aspects of the work that do not quite work, especially on repeated viewing – in particular, the rising 'sparkles' of yellow and orange colour that accompany the 'tinkly' glassy sounds during the first part of the work dominate the field of view too much and distract from the main interest of the glowing spheres, but this is a minor problem that does not warrant another reworking of the piece.

Moving forward from this work, there are some of the initial ideas, such as time-lapse photography, which I would consider using for a future work, but the most successful aspect was the time taken to fully develop the narrative of the work at the earliest stages, which strengthened the whole process, and gave a strong narrative structure to the work. While the audiovisual objects and environment depicted in the work are all essentially abstract in terms of not being a definable real-life analogue, the nature of a composed piece means that a coherent narrative structure provides a strong ongoing impetus to the overall work, and allows the longer sequences to develop while maintaining the audience's interest. This work reinforced that awareness of the importance of strong structure for me, and marked a renewed confidence in my own abilities in audiovisual composition.

***Skitterling* (2012)**

Original Design Brief

Original Storyboard

Skitterling – Figure and Ground

Andrew Connor – PhD Creative Music Practice

Design Brief

The next audiovisual work I aim to complete for my PhD will make use of the concepts I have started to explore in my paper to SSSP 2011 on perspective and emergence.

The key element identified in this paper, which will form the cornerstone of my thesis, is the idea of the 'audiovisual object'. This is a distinct phenomenon, separate from simple synchronisation of sound and vision, or the more fortuitous synchresis identified by Chion.

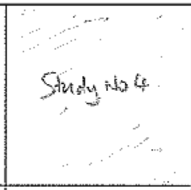
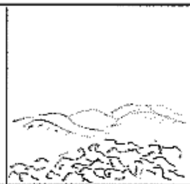
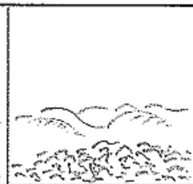
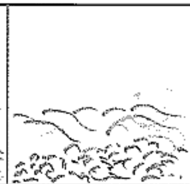
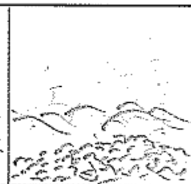
In the sound, vision and audiovisual examples referred to in the paper, I have noted various perceptual conditions contributing to the emergence of the audiovisual object. One key condition is that of figure and ground – the emergence of both sound and vision objects as identifiable figures against a background allows for a potential audiovisual object to be formed as the figures combine. At the point where the sound object or the vision object becomes just discernible from the surrounding ground, there is an exclusive allocation of a boundary to the object.

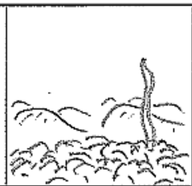
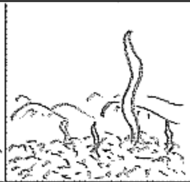
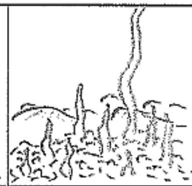
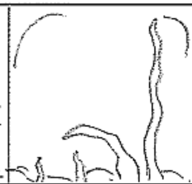
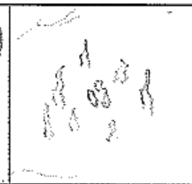
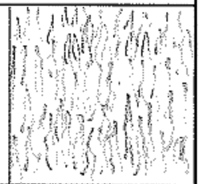
Skitterling will examine figure and ground primarily as a technical exercise, but will maintain a narrative and aesthetic approach to ensure the full audiovisual piece has an appeal to the viewer. The piece will contain several instances of figures emerging against the ground simultaneously in both audio and vision, and will also include individual emerging figures in both perceptual streams without a designed counterpoint, to examine the effectiveness of planning for a dual emergence against the fortuitous random combinations of sound and vision (synchresis).

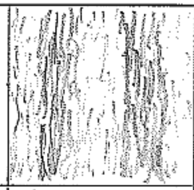
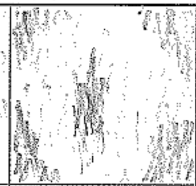
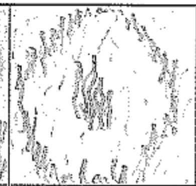
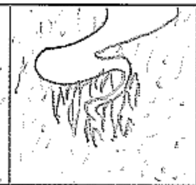
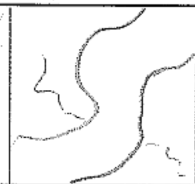
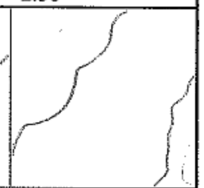
The piece will be a maximum of 6 minutes in length, to allow the animation to be carried out in a reasonable period of time. The current intention is to use Blender 3D as the main animation software, although parallel test work using other animation techniques may lead to other packages being used in conjunction with Blender.

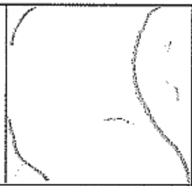
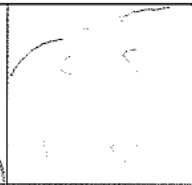
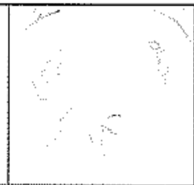
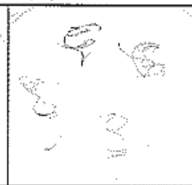
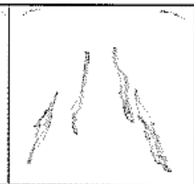
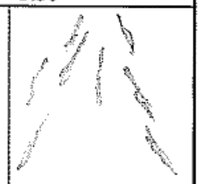
From initial discussions with ECA's Head of Animation, I will also adopt a more coherent specifically designed colour palette for the visual part of the work, consonant with the specific sound palette I would normally adopt for the compositional portion of the work.

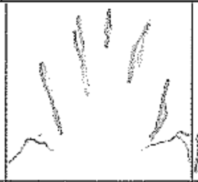
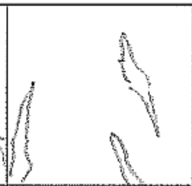
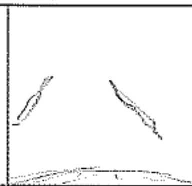
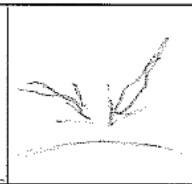
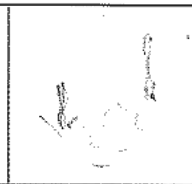
A rough storyboard for sound and vision is attached to this design brief.

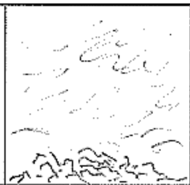
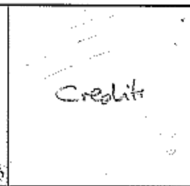
Time	0:00	0:10	0:20	0:30	0:40	0:50
						
Vision	Opening title – fade in from black	Title cross fades into starting scene, a barren landscape / desert dunes	Continuing barren landscape	Rising winds causing sand flurries to start to rise and cross in sympathy to panned sound	Sand flurries gain in speed and fury – a windspout centres on position screen right	A glowing tendril erupts from the ground and starts to grow in time with the growing tonal sound
Sound	Silent, gradual fade in of white noise / very granular, sandy sound	The quiet hissing white noise continues and strengthens	Within the noise, individual directions of more discernable, textured wind segments begin to form	The wind noise becomes more prevalent above the background hiss	The wind noises become more dominant and centred on the right.	A comb filtered pure tone emerges from the swirling noise, gaining prominence against ground, the wind flurries begin to die

Time	1:00	1:10	1:20	1:30	1:40	1:50
						
Vision	The tendril continues to grow – the sand windspout and flurries die down as the background sound subsides	The tendril grows and expands as the tone grows. Other tendrils start to appear and grow with the additional tones	Tendrils continue to appear and grow	Camera viewpoint explodes upwards, spinning as ‘pushed’ by tendril growing from directly below	Camera spin starts to settle down, ends up focussed on landscape below	Ground now carpeted with tendrils, all glowing the same colour, expanding and obscuring the original ground
Sound	The pure tone strengthens and moves slightly from side to side	The main tone continues to pan slightly, joined by growing harmonic tones	Harmonic tones dominant against fading background noise, although it is still discernable	Bass fundamental of tones appears suddenly, along with hard crash and noise	The bass fundamental fades to match others, some oscillation of tones	Tones continue to dominate, oscillations begin to increase

Time	2:00	2:10	2:20	2:30	2:40	2:50
						
Vision	The tendrils start to glow in related colours, striations appear and move across the screen with the panning tones	The moving colours and tones continue and grow, centring on a spot mid-lower screen	Colours continue moving, spot begins to coruscate and brighten, tendrils there swell and pulsate	As the colours reach a climax, the spot explodes and a long tendril explodes upwards past the camera	Camera is buffeted in wake of tendril's passage, spinning and bouncing to be carried along with tendril	The camera continues to settle as the tendril 'grows' to occupy the bottom half of the screen
Sound	Oscillations grow and pan across from left to right etc. as tones start to vary	Variation and oscillations in tones grow	A central pulsating tone gradually deepens while other tones start to rise in pitch	Main pulse deepens, strength-ens to a room-shaking sub-harmony, other tones rise above audible reach	The main tone re-emerges from the climax, accompanied by shimmering harmonics, some granular noise	The granular noise grows into a base rich with sub to give a ground to the dominant tone

Time	3:00	3:10	3:20	3:30	3:40	3:50
						
Vision	The tendril starts to switch directions and snake wildly, the camera jounces along with it	At a particularly rapid jounce, camera spins and falls 'into' tendril, surrounded by its basic colour and swirling gasses moving with the tendril growth	The camera settles into the stream, bobbing gently. The gasses swirl and create small congregations of colour	The gasses swirl more furiously, the congregations of colour grow brighter and more coherent	The colour congregations spark and fuse to form small spears of colour, spores moving rapidly, speeding past the camera and forward	The camera speeds up in the wake of the spores
Sound	Main tone is panned to left, and to right, with Doppler if that adds to the effect	Brief point of silence, followed by ripping sounds and 'underwater' transition	Main tone is softened and added to with high freq harmonics, panned at sides	Harmonics grow, become more strident, narrow-ing to central discordant tones	Small explosions and rapid forward Doppler of discordant tones	Sub component of main tone develops, grad-ually increasing pitch overall

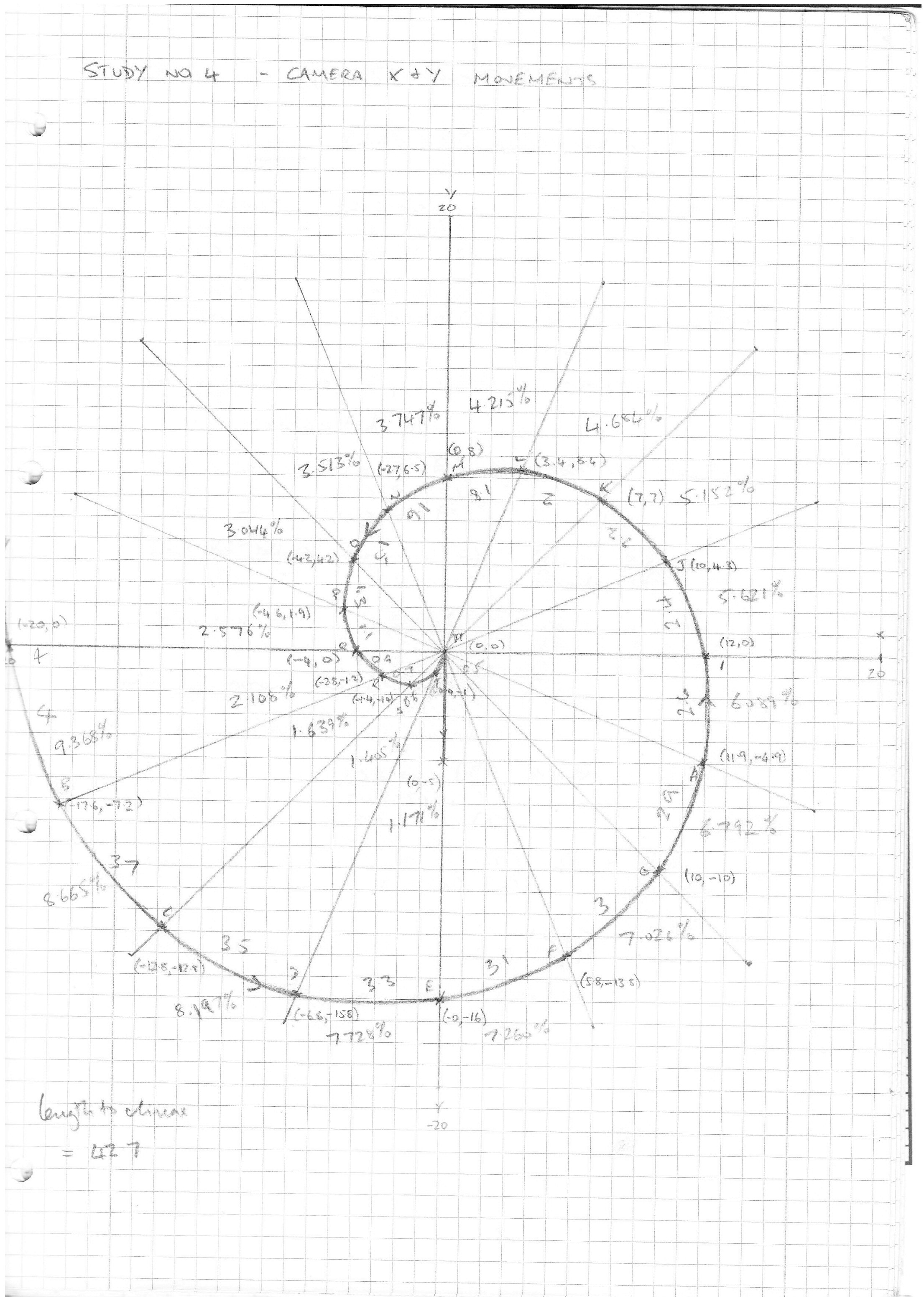
Time	4:00	4:10	4:20	4:30	4:40	4:50
						
Vision	The spores and camera burst forth from the forward end of the tendril, heading off into a free, open expanse of nothingness	The camera weaves in and out between the spores	A landscape/surface appears at the lower edge of the screen	The spores start to head downwards to the surface, lengthening and 'burning' at the fore as the velocity increases	The camera follows the spores down to the surface, landing in a crash as the sound reaches a climax	The screen is a haze of dust and particles thrown up by the collision
Sound	Climax of tone, ripping/ tearing sounds, followed by silence with glassy high freq bursts	The high frequency glassy/icy tones fly past the audience on left and right	Silence begins to be overlaid with a growing low granular / noisy texture	Glassy bursts lower slightly in pitch, overlaid with turbulence/ fiery noises	Low granular noise rises rapidly in pitch and speeds up to a loud climax	Aftermath of climax – settling static from high to low frequencies

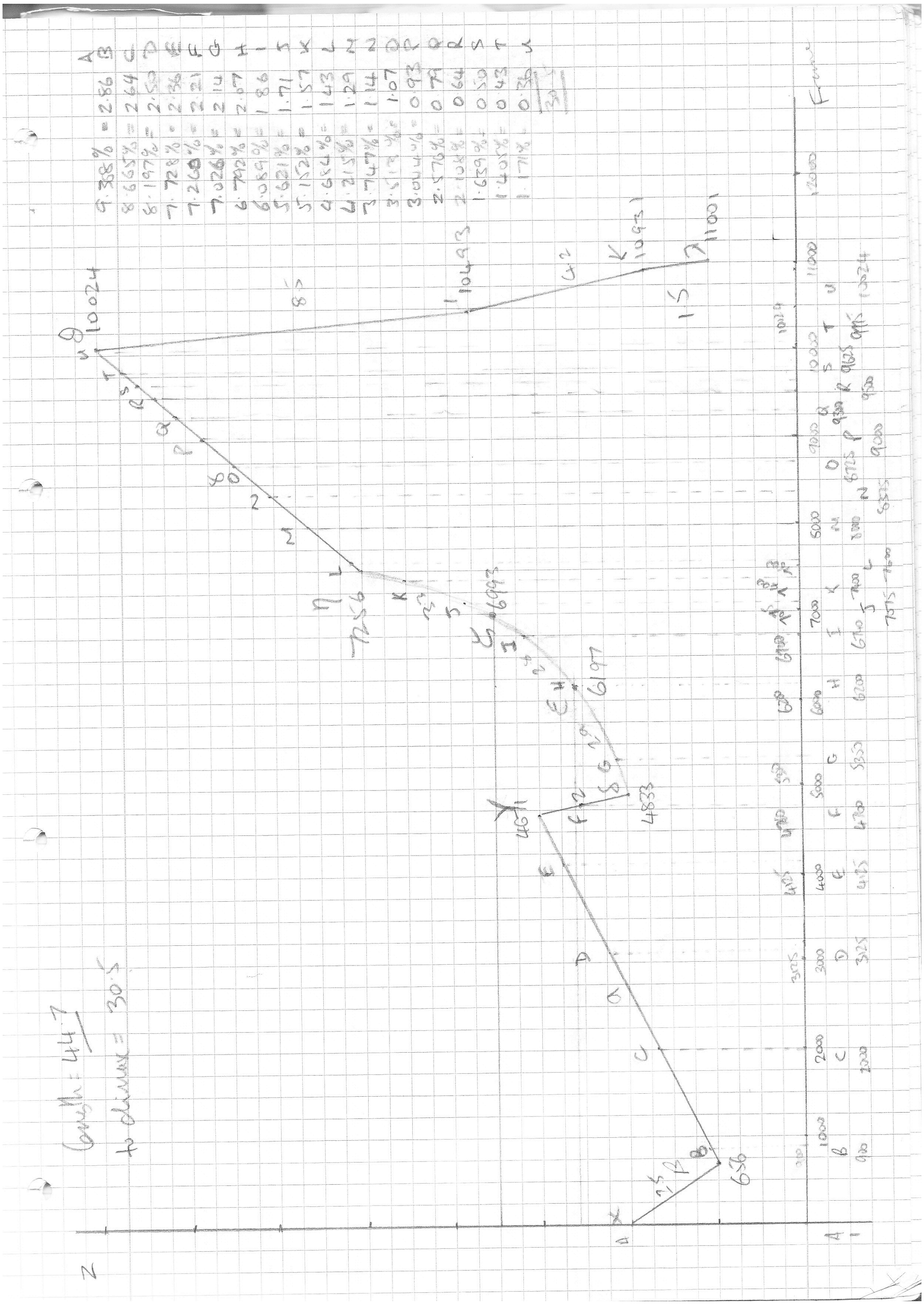
Time	5:00	5:10	5:20	5:30	5:40	5:50
						
Vision	The dust begins to settle, the landscape starts to be revealed as a bare, barren desert scene.	The dust finally settles, the impacts of the spores settle from bright glowing spots to the standard surface	The landscape reverts fully to the same bleak desert as at the beginning	Fade to credits	Credits	Copyright
Sound	Static falls to sub harmonics and out of range, sandy hissy grains emerge from the dying static	Short almost subliminal reprise of the glassier tones, dying into a gradually settling sandy hiss	Sandy hiss dominates and also begins to die	Sandy hiss dying away	Sandy hiss dies away to silence	Silence

Skitterling (2012)

Camera Movements in X & Y planes

Camera Movement in Z plane


STUDY NO 4 - CAMERA X & Y MOVEMENTS
3.747%
4.215%
4.684%
3.513%
5.152%
3.044%
5.621%
2.576%
6.089%
2.108%
1.639%
1.405%
9.368%
6.742%
1.171%
8.665%
7.026%
8.197%
7.728%
7.260%
(0,8)
(3.4, 8.4)
(7,7)
J (10, 4.3)
(12,0)
(11.9, -4.9)
(10, -10)
(5.8, -13.8)
(-0, -16)
(-6.6, -15.8)
(-12.8, -12.8)
(-17.6, -7.2)
(-20,0)
(-4.6, 1.9)
(-4,0)
(0,0)
(0,-5)
length to climax
= 42.7



Skitterling (2012)

Frame listings for amplitude peaks and troughs (sample)

TIME	FRAME	LEFT	RIGHT
88.115	2204	-27.43	-27.99
88.158	2205	-29.96	-30.82
88.182	2206	-26.84	-27.75
88.374	2210	-28.72	-29.12
88.411	2211	-27.17	-28.43
88.452	2212	-26.26	-25.9
88.554	2215	-26.1	-27.76
88.622	2217	-27.65	-26.38
88.721	2219	-25.51	-24.5
88.832	2222	-25.46	-27.19
89.051	2227	-25.7	-26.74
89.163	2230	-28.38	-27.39
89.206	2231	-26.69	-26.94
89.299	2233	-28.7	-26.61
89.328	2234	-25.91	-26.52
89.399	2236	-24.61	-24.16
89.511	2239	-26.81	-25.5
89.623	2242	-28	-24.13
89.648	2242	-25.42	-25.78
89.683	2243	-25.98	-26.23
89.786	2246	-27.87	-24.58
89.943	2250	-28.1	-26.7
89.984	2251	-28.78	-25.4
90.161	2255	-32.26	-27.35
90.174	2255	-28.45	-27.52
90.221	2257	-28.44	-26.86
90.331	2259	-28.67	-27.45
90.372	2260	-28.93	-27.24
90.495	2263	-29.18	-26.54
90.592	2266	-30.87	-31.92
90.600	2266	-26.38	-27.05
90.743	2270	-30.25	-26.63
90.766	2270	-28.36	-27.35
90.826	2272	-28.58	-28.61
90.938	2274	-26.43	-25.97
91.003	2276	-25.91	-25.08
91.038	2277	-27.03	-26.52
91.137	2279	-27.16	-27.45
91.224	2282	-26.88	-28.97
91.254	2282	-26.84	-27.43
91.323	2284	-30.1	-30.3
91.337	2284	-27.95	-27.05
91.425	2287	-26.12	-28.82
91.458	2287	-26.92	-27.5
91.562	2290	-27.94	-28.15
91.610	2291	-27.91	-28.87
91.659	2292	-28.27	-30.84
91.698	2293	-24.57	-28.74
91.900	2299	-25.99	-29.47
91.910	2299	-26.51	-28.01
91,938	2299	-27.38	-27.14
91.977	2300	-27.72	-27.71
91,998	2301	-26.75	-27.47
92.206	2306	-24.31	-24.24
92.305	2309	-27.57	-27.34

TIME	FRAME	LEFT	RIGHT
92.374	2310	-25.57	-29.87
92.409	2311	-25.34	-29.07
92.487	2313	-27.57	-28.78
92.544	2315	-27.02	-25.64
92.604	2316	-26.83	-28.31
92.774	2320	-33.54	-29.21
92.785	2321	-29.88	-26.37
92.808	2321	-29.73	-24.45
92.834	2322	-26.5	-27.05
93.000	2326	-30.18	-28.41
93.078	2328	-28.4	-24.41
93.130	2329	-27.14	-27.35
93.224	2332	-31.27	-28.33
93.260	2333	-28.34	-25.51
93.363	2335	-30.2	-27.23
93.376	2335	-27.32	-25.75
93.476	2338	-30.11	-29.59
93.529	2339	-27.8	-30.48
93.553	2340	-26.36	-26.48
93.684	2343	-29.83	-32.38
93.690	2343	-28.3	-27.8
93.784	2346	-27.12	-30.1
93.807	2346	-28.78	-27.87
93.896	2348	-30.55	-29.04
93.922	2349	-27.66	-28.0
93.958	2350	-28.84	-28.68
93.998	2351	-33.32	-28.01
94.004	2351	-27.17	-30.21
94.077	2353	-31.01	-27.32
94.098	2353	-29.01	-27.64
94.168	2355	-27.98	-28.17
94.222	2357	-26.45	-26.48
94.254	2357	-26.92	-27.04
94.450	2362	-30.01	-28.8
94.478	2363	-27.21	-30.08
94.538	2364	-30.11	-29.87
94.562	2367	-27.61	-28.13
94.644	2367	-29.43	-28.15
94.665	2368	-28.83	-28.66
94.698	2368	-27.33	-30.3
94.765	2370	-28.46	-29.96
94.783	2371	-29.07	-28.48
94.810	2371	-30.58	-25.51
94.818	2371	-26.85	-27.31
94.954	2375	-24.5	-28.16
94.970	2375	-26.75	-31.33
95.006	2376	-29.86	-27.1
95.016	2376	-26.81	-28.54
95.123	2379	-26.74	-27.74
95.156	2380	-27.47	-31.52
95.167	2380	-26.42	-29.36
95.201	2381	-27.92	-26.87
95.364	2385	-28.85	-29.73
95.383	2386	-28.0	-28.32
95.397	2386	-29.27	-25.66

Production Notes

Synthonomy (2014)

Synthonomy (2014)

Introduction

Synthonomy was created as part of the 2014 celebrations marking a centenary since Norman McLaren's birth. Although McLaren is perhaps better known for his work for the National Film Board of Canada, he was born in Scotland and studied at the Glasgow School of Art. His experiments in film and animation have strong resonance with the field of Visual Music – he once remarked that Oskar Fischinger's early abstract animation "made an indelible impression on me, excited a yearning in me, and was to have a profound, long-lasting influence on many of my films"[234].

In his centenary year, 2014, the Scottish animator Iain Gardner co-ordinated a series of events to celebrate McLaren's work, in association with the National Film Board of Canada. The University of Edinburgh, through the Edinburgh College of Art, took part in these celebrations, both by hosting an exhibition of McLaren's work at the University's Talbot Rice Gallery, but also by organising a set of workshops for ECA students to develop creative works inspired by Norman McLaren, which would then be performed at a concert linked to the gallery exhibition. Two workshops were organised in early June, one focussing on the creation of experimental compositions for abstract animations, organised by Yati Durant of the Music department and led by the guest composer Dr. Joseph Hyde from Bath Spa University, and another focussing on experimental animation techniques, organised by Jared Taylor, Head of ECA Animation, and led by the animator Jim LeFevre[235].

The workshops were aimed at creating material that would be assembled for a concert to be held in the University's Playfair Library, adjacent to the Talbot Rice Gallery. There was a very tight deadline to meet, as the workshops ran over the two days immediately prior to the day of the concert. The intention was to marry

[234] Moritz, W. (2004). *Optical Poetry: The Life and Works of Oskar Fischinger*. John Libbey Publishing, Eastleigh, UK. p166.
[235] The project description for the workshops and the concert is included in Appendix 1.

experimental musical compositions created by students on the MSc Composition for Screen course, led by Yati Durant, with looped animations created by undergraduate animation students in Jared Taylor's department. As my PhD research examined the creation of both musical compositions and animation for experimental, abstract audiovisual compositions, I was invited to contribute both a composition and the accompanying visuals. In the event, I created the short piece *Synthonomy* and also edited together the animated sequence to accompany Aleksandra Kovac's composition *One*, which was accompanied in the concert with additional live music performed by the Edinburgh Film Music Orchestra (EFMO), conducted by Yati Durant.[236]

Concept

This project was very different from the other audiovisual work I have produced over the course of the PhD in that the project brief was set by someone else. However, the brief was quite wide-ranging, and allowed me considerable editorial control over the content I created, which meant that the challenge was more in the time limit allowed for the work rather than feeling too confined in what I could do. In terms of time, the entire project took place over a 54 hour period, from the initial workshop at 10am on Thursday 5th June 2014 to 4pm on Saturday 7th June, as the work had to be completed in time for a rehearsal with EFMO prior to the concert start time of 7.30pm.

The first workshop was aimed primarily at the music composition process, although it was open to both animators and composers. Professor Joseph Hyde, based at Bath Spa University, had just completed a research trip to California to study papers and works by McLaren and others, and therefore had a wealth of knowledge about McLaren's approach to his film-making. McLaren often created his own soundtracks, often by drawing directly onto the sound strip of the film stick he was

[236] Concert footage of *One* is available at https://www.youtube.com/watch?v=XsFU-kwH4IY [accessed 17/08/2015].

using, effectively synthesizing his music by mapping out the sound wave patterns he wanted to explore. This workshop provided a solid background in the correlation between sound and image that McLaren sought to achieve, which had strong resonances with my own approach to creating strong bonds between sound and vision to create audiovisual objects within my own compositions.

The second workshop, run by Jim LeFevre, concentrated on a specific form of experimental animation, using a device called a phonotrope. This is in effect a large cardboard disc placed onto a record player, which is then set to rotate at 45 rpm. A camera set to record at 25 fps is aimed at the disc. The disc itself has 33 divisions on it, on which successive images can be drawn, in the same manner as a zoetrope or praxinoscope. The combination of the rotation of the disc and the camera frame rate means that an animation is created across the disc divisions passing by the camera lens. The 33 successive images form an animation loop, which lasts for 1.3333 seconds. This basic loop duration was very important, as it would form a specific constraint on all the composition and animation that I was to do for the piece.

Following the workshops, the individual composers and animators had created some short musical pieces or looped animations. The animations were either recorded sequences from the phonotrope experiments, or individual loops constructed using other animation techniques. The challenge then was to marry the two together. I took on the role of editor for one of the compositions, by Aleksandra Kovacs, and combined it with looped animations provided by the animation students. In addition, given that my expertise was in combining sound and vision, I also took on the task of creating my own audiovisual work in tribute to Norman McLaren, while observing the restraints based on the phonotrope animation loop. In this case, given the time constraints, I decided to use sound synthesis for my audio composition rather than the acousmatic techniques of manipulating recorded sounds that I had used up until this point, and while I created several animation loops for use in this piece, I also made use of animation created by one of the animation students, Abigail Lamb, as it proved to fit extremely well with my audio composition.

Audio Composition

As the composition had to be completed in a very short time, I eschewed my normal approach of recording specific source material to then manipulate, process and edit. I decided to use a digital audio workstation package that was suitable for sound synthesis – Cubase. This software allowed me to set up a specific tempo and use various synthesis algorithms to create a suitable soundtrack for a McLaren tribute.

From the phonotrope workshop, I had established that my basic time unit was 1.3333 seconds, the duration of one typical animation loop created using this method. This meant that there would be exactly 45 loops in 60 seconds – if I counted 1 loop as equivalent to a bar in common time, this gave me a tempo of 180 beats per minute, This is a rapid tempo – equivalent to *allegro vivace* (very fast) or *presto* (extremely fast). If I included a percussion beat to match this tempo exactly at 180 beats per minute, the piece would be fast and driven throughout. However, by working to a pattern of only 1 beat per bar at the start, then layering in more sounds as the piece continued, I could create a growing dynamic in the piece, adding a sense of drive as the work progressed. I also decided to use the same technique in reverse to bring the piece to a close, gradually reducing the beats per bar until I returned to the steadier beat of 1 beat per bar that I started with.

This basic structure gave me a platform on which to develop the piece. I decided that the overall length would be 120 bars, or 480 beats long – this would produce a piece with a duration of 2 minutes and 40 seconds, with a total of 120 animation loops in the visual perceptual stream.

I started the compositional process by adding a single bass drum beat at the start of every bar, using a percussion synthesizer in the Halion plug-in available within Cubase. This gave an immediate anchor point for the start of each loop duration throughout the piece. I decided that I would group bars together in sets of eight – this is in line with the basic repetitive pattern of much of Scottish and Irish music, reflecting both McLaren's origin and my own, and also meant that there would be 15 of such 8 bar groupings throughout the piece, outlining a mid-level structure that would be useful in grouping the animation loops together as the piece progressed. For the second set of 8 bars, I added the sound of a wooden block being struck on the third beat of every bar, adding more drive to the sound. I continued this pattern, adding in other percussive elements such as a cymbal, shakers, tom-tom, etc. I continued to build the driving impulse of the beats in the work up until the 80th bar, the tenth 8-bar cycle, and then started to gradually reduce the driving rhythms, returning at the end to a simple bass beat alternated with a cymbal, for the final 8 bars, culminating in a single sustained cymbal hit.

Having created the backbone of the piece using percussion synthesis, I then decided to introduce a long held note to balance the driving percussive beat. I had downloaded various free plug-in synthesizers for use in Cubase, and started experimenting with the preset instruments and voices provided in each one. I found an effect in the Alieno plug-in which worked well in combination with the percussive beat, an eerie vibrato with whistling overtones called ‘Mars Volta’. I introduced a long held note over 3 bars from bar 5 onwards, pitched at C2, i.e. one bar into the second half of the first 8-bar phrase. The nature of the note meant that it continued to decay over into the 1st bar in each half-section of the basic 8-bar structure, allowing a continuity to lead into the next half-section. I continued this simple C2 pitched note throughout the rest of the piece, finishing at bar 112 and allowing a final 8-bar phrase with just the percussion to close out the work. I also added in another note using the same synthesis effect, starting at bar 32, this time at G3, an octave and a fifth above the original note, for a duration of two bars at the end of each 4-bar half-phrase. This continued through to bar 96, and added a major melodic harmonic to the main underlying C2 pitch. Finally, I also added a single bar-long note pitched at E ♭3 from bar 67 onwards up to bar 96, which added a minor melodic variation to the final bar of each 4-bar half-phrase.

Having established the beat and a longer held note to provide a continuing theme throughout, I wanted to add some more variation to the piece to add interest and colour. I returned to the Alieno synthesizer plug-in, this time using the ‘Serpent’s

Egg' preset, which created a pitched note accompanied by a rhythmic white noise element, reminiscent of the dry noise from a rattlesnake's tail. I continued to work in the C chordal family by pitching this effect at a consistent C5, three octaves above the core C2 used for the fundamental continuing note. I also wanted this sound to break up the normal 8-bar pattern I had established, and so placed it to start at the 5th bar in the second 8-bar phrase, bar 13, for a duration of 1 bar, and repeated it every other bar for an 8-bar duration (i.e. bars 13, 15, 17 and 19). This meant that the sound would play on the 'off-beat' bar of the 8-bar cycle, and would also cover the end of one cycle and the start of the next, adding to the slight imbalance conveyed by the sound. I continued this pattern over the rest of the piece, placing this 8-bar sequence to start at bars 29, 45, 61, 77 and 93, in each case starting in the second half of a normal 8-bar phrase and continuing into the first half of the next one.

I continued to explore the idea of setting additional sounds off the expected pattern of the underlying 8-bar phrases, adding in another instrument using the Crystal synthesizer plug-in. The 'Entropic' preset provided a mixture of an almost vocal phrase combined with a slithering synthetic glissando, which promised an opportunity for synchresis against some of the animation loops that Abigail and I had created. Again, I decided to insert the sound as a longer held note, this time at C4, lasting for a full 8 bars. However, I offset the start by 5 bars from the normal 8-bar cycle, again intending a slightly unsettling effect. As I wanted an emphasis to end the note with, I added a very short 1 beat repetition of the note at the start of the following 9th bar. I also added in an additional note using the same preset at G3, starting at the 5th bar in from the first pitched note, but finishing on the 3rd beat of the 4th bar, just short of a full 4-bar half-phrase. This added to the interest of the instrument, while also adding to the emphasis of the added note at the start of the 9th bar. A standard 8-bar space was left between iterations of this phrase, with the phrases inserted at bars 21, 37, 53, 69, 85 and 101.

The penultimate effect added was another instrument from the Crystal suite, the 'Swell 7' preset. This provided a short electronic trill for one bar, followed by a percussive snap just as the second bar started. As I had already placed the other additional instruments off the main 8-bar phrase pattern, I returned to the standard 8-bar phrases for this sound, placing it to play twice at the start of an 8-bar phrase, pitched at C5, with the trills in bars 1 and 3, and the snaps starting bars 2 and 4. I started iterations of this pattern from the 3rd 8-bar phrase, bar 17, and continued to add it in on every other 8-bar phrase, on bars 33, 49, 65, 81, 97 and 113 – making this sound pattern continue on to the quieter section at the end of the piece.

The final sound addition came from the DSK Ethereal PadZ plug-in, using the ‘Sweeping’ preset. This provided a padded synthetic chord, panning gradually from left to right over the 2 bars in which I placed it, culminating in a short lower tone. As with other instruments, this was pitched within the C chord at C4. The two bars of this note were placed at the end of the 8-bar phrase, with the first insertion at the end of the second 8-bar phrase at bars 15 and 16. This pattern continued with repetitions at bars 31, 47, 61, 79, 95, 111, with one final out of sequence insertion in bars 117 and 118, the 5th and 6th bars of the final 8-bar phrase.

The full audio composition therefore combined a lot of different synthesized instrument voices, which worked together satisfactorily to create a platform against which the animation loops could be added. The increase of the percussion beats from the start through the first two-thirds of the piece provided impetus, while the reduction across the final third allowed for a natural closing down of the work. The instrumental voices, while containing interest in themselves, nevertheless allowed sufficient audiovisual ‘room’ to allow synchresis and links to be made to the visual objects in the animation loops, allowing the fusion and creation of audiovisual objects throughout the piece.

Video Composition

This audiovisual work also departed from my normal practice in that I created the visual elements of the work before the audio composition, going against my usual approach. Although my usual practice of creating the sound first is predicated on the idea that this allows a stronger presence of the audio perceptual stream against the more naturally dominant visual stream, in this case the piece does not seem to have suffered from the reverse approach.

Given that the inspiration for the piece was Norman McLaren's work, I referred to several of his animations as inspiration for the animation loops I created. In his more abstract pieces, he made extensive use of both geometric patterns and more organic, drawn shapes. In *Lines: Vertical* (1960)[237] and *Lines: Horizontal* (1962)[238], the dominant elements are, unsurprisingly, lines in the appropriate orientation, moving to and fro against each other to match the accompanying soundtracks. In other films such as *Loops* (1940)[239] and *Dots* (1940)[240], McLaren drew squiggles, blobs and small dots directly on to the film negative, marrying the shapes up to marks he made directly on to the sound strip at the side of the negative.

I used these works by McLaren as inspiration for my own animation. Following the example of the *Lines* films, I created short 33-frame loops by drawing lines on a basic 1920 x 1080 pixel field using the graphical manipulation package GIMP. By moving the lines, saving a frame in jpeg format, and continuing on to move the lines again, I created a repetitive cycle across the requisite 33 frames. I continued to follow a similar approach for circles and blobs, and created a simple sequence showing a square frame containing four shaded circles rotating around its centre point, as well as a very basic loop with a central white circle with a black centre growing and contracting within a black field. I then decided to also create 3D CGI spheres, as a modern development of McLaren's pioneering work. I made use of 3DS Max as a 3D animation package to create a set of 3 loop sequences featuring spheres. The first loop showed a single central sphere, with a deformation modifier applied over the 33 frames to convert it to a misshapen blob. The other 2 sequences started

237 McLaren, Norman. (1960). *Lines: Vertical*. Animation, accessible at https://www.youtube.com/watch?v=LnbavAYULUU [accessed 17/08/2015].
238 McLaren, Norman. (1962). *Lines: Horizontal*. Animation, accessible at https://www.youtube.com/watch?v=qJwfeG3Mntk [accessed 17/08/2015].
239 McLaren, Norman. (1940). *Loops*. Animation, accessible at https://www.youtube.com/watch?v=6JvOqeEtFRY [accessed 17/08/2015].
240 McLaren, Norman. (1940). *Dots*. Animation, accessible at https://www.youtube.com/watch?v=E3-vsKwQ0Cg [accessed 17/08/2015].

with 4 spheres jammed up against and overlapping each other, which then either consolidated into one central sphere, or shrank into 4 small but distinct separate spheres.

Due to time constraints, I was unable to complete any further animated loops myself, and had to move swiftly on to create the synthesized composition that would be the sound track to my finished work. The animation students involved in the project had consolidated all the loops they had created into a shared folder and from those sequences, I decided to use some created by Abigail Lamb, which worked well with the composition I had just completed.

As per the construction calculations I had made in approaching the audio composition, I knew that I had to combine 120 loops in sequence to match the 120 bars of the audio composition. Given that I had repetition of themes in the music, I felt that I should also rely on repetitions of the visuals as well, meaning that I could select the strongest animated loops that suited the tone of the music, and repeat them as necessary to match up against the soundtrack. Due to the rhythmic nature of the music, I knew that synchresis would come into play as long as the animated sequences matched up exactly with the bar durations.

The assembly of the animated sequences was accomplished using a non-linear editing package, Adobe Premiere Pro. This allowed me to 'drop in' the loops I felt best matched the individual bar sequences in the soundtrack, without having to assemble them in an exact chronological order. This allowed me to experiment with different animated loops against the different bars and sound phrases, making sure that a close audiovisual match was created at every point of the work. Eventually, I matched all the soundtrack bars against individual loops, with a 'chorus' effect where the 3D spheres regularly re-occurred throughout the piece. I also made repeated use of the line sequences, but added visual interest by changing the base colours of the lines for different repetitions. Abigail's animated loops were different in style from mine, having been hand-drawn and animated by photographing the individual frames, but melded well with my loops and the underlying music.

Audiovisual Work

The editing process allowed a good match of sound and image, making use of the basic calculations for bar and loop durations to allow close matching of sound to an appropriate image. Chion's principle of *synchresis* also came into play, as the movements in some of the loops resonated with the movement, in pitch or stereo panning, heard in the music. The repetitions in the music were matched with repetitions and reiterations in the visual sequences, again reinforcing the matches between sound and image. The correlations between the two perceptual streams allowed audiovisual objects to fuse and be created from the two elements, showing that it is possible to create 'building-blocks' of audiovisual objects that can be slotted together.

The limited time available in which to create and assemble all the elements of the work meant that there were compromises made in the compositional process, particularly at the assembly stage. The matches between the sound and image were fixed on a basis of 'it's good enough', which fortuitously worked in the majority of matches, but closer matches may have been achievable using a wider range of animated loops had more time been available for experimentation. However, the finished piece as it stands works quite well, and closer matches may not have been possible without additional time to specifically craft new animation loops to fit more closely with the music.

The title of the work, *Synthonomy*, was derived from another of McLaren's animations, *Synchromy* (1971)[241]. This animation gave a literal visual illustration of the patterns that McLaren used on the sound strip of the film stock, coloured to add visual interest. As the sound I had created had been completely synthesized rather than manipulated from real-world recordings, I decided to reference both the synthesis and tip a nod to McLaren's work in the title. The piece was performed at the McLaren Tribute Concert on the 7th June 2014, and received good feedback from

[241] McLaren, Norman. (1971). *Synchromy*. Animation, accessible at https://www.youtube.com/watch?v=UmSzc8mBJCM [accessed 17/08/2015].

the audience. The standard of all the work was high, particularly given the tight time constraints for composition, animation, and video assembly. The works have not been performed as a complete suite again following that concert, but the individual pieces have been exhibited, mainly by the animators, at other shows and animation festivals. *Synthonomy* received a second showing as part of the Black Cube Collective's exhibition of various artworks in Athens, Greece in October 2014, again receiving positive feedback from the audience.

Reflections on Completed Work

Synthonomy is very different from the other works I have created over the course of the PhD, but is actually one of my favourites from my output. The incredibly tight deadline for its creation and assembly forced me to take a completely different approach to making audiovisual work from my normal approach, and opened up other possibilities to achieve my goals using quicker methods requiring lower levels of processing power.

My practice up until this point had been to create the audio component of my composition first, then the video portion, although the visual elements were mostly formed in my imagination as the audio composition took form, so the animated composition was used mainly to transfer that imagined visualisation into the closest I could achieve using the software available to the best of my ability. The approach taken here, to create the video portion and then the audio component without any specific reference to each other apart from the duration of the individual loops, left a lot of experimentation at the assembly stage, which was both frustrating and quite liberating. While I could look at individual loops and imagine which ones would fit best against the audio composition at specific points, the only way to test this was to try the two out against each other, and measure the effectiveness of the synchronisation and synchresis between the two perceptual streams. While I would not want to do this with the majority of my work, I would recommend this approach as part of a workshop on creating audiovisual compositions, as it exercises the ability of the composer to identify effective audiovisual fusion of audio and vision into convincing audiovisual objects.

The use of synthesized sound was also a new approach for me. While I had made use of synthesis in projects outwith the scope of my PhD, this had mainly been in the creation of spot effects or a musical score for short films, and I had avoided its use as part of my own compositional practice. The effective use of synthesis in this work opened up the possibilities of adding to my compositional techniques by including synthesis where appropriate, although my preference would still be to make the majority of my soundscapes using acousmatic composition, manipulating real-world recordings.

The use of repeated sound and visual motifs in the piece work quite satisfactorily in creating a cohesive structure, but on repeated viewing the 'chorus' that repeats, alternating the conjoined 4 large 3D spheres alternately coalescing to one central sphere, or into the separate smaller 4 spheres, followed by the deforming sphere, occurs too often. If I revisited the piece, I would retain the original audio composition, and most of the existing animation loops, but would add in much more variation for that repeated 3D sphere sequence leading up to the deforming sphere.

Performances

Black Cube Collective Exhibition, Athens, Greece	*Oct 2014*
Norman McLaren – Animated Composer Tribute Concert, Edinburgh, UK	*Jun 2014*

Conclusions

The constraints under which *Synthonomy* was produced put a lot of pressure on all the participants to create work to a suitable standard to perform publicly at a concert celebrating the work of one of the most prolific pioneers of experimental film-making, but in the event we all rose to the challenge. As with imposing my own restraints on the earlier composition *Skitterling*, the boundaries placed on my work on this occasion resulted in increased creativity to solve the problems I encountered.

This exercise also allowed me to work closely with other composers and animators, helping to realise their own creative visions, but also to inform my own work, and in the case of Abigail's animations, to include another person's visual work against my original audio composition, and for the piece to be greatly strengthened as a result. While I find it easier still to create and hold the two perceptual streams in my head as I start the composition process from the audio element, the experience of working with Abigail reassured me that the principles I apply to audiovisual composition are also valid in collaborative compositions.

Synthonomy has been publicly performed fewer times than any other of my works, with the exception of the very recently completed *Canto*. This reflects the fact that it lies slightly outside my usual practice, but it has influenced my later work, especially in my reconsideration of the merits of using synthesis, and of trying a variety of visual sequences against an audio sequence to discover the closest audiovisual 'fit'. It is therefore significant in terms of the development of my practice and my understanding of the techniques that can be used in creating audiovisual objects.

Synthonomy (2014)

Original Project Description

Norman McLaren – Animated Composer Workshops and Concert

A project designed by Yati Durant and Jared Taylor, presented by the Edinburgh College of Art, the Reid School of Music and the University of Edinburgh Challenge Investment Fund

Project Description:

This project focuses on the creative experiments in animation and music by Scottish-born animator Norman McLaren (1914 – 1987), and aims to present 3 new works for animated film and live orchestra based on creative phases of McLaren's life. Active participants will work in 1 of 3 small teams of animators and composers, and take part in two workshop days, followed by a concert. The workshops will present modern work and research by contemporaries of McLaren, in ways that can draw a parallel to his creative style and innovation. Works-in-progress for the concert event will be honed and analyzed by the guest speakers, using hands-on practical techniques and examples. Guest speakers will also share their personal creative work and how it has been influenced by McLaren's creative legacy.

The concert event is free and open to the public, and tickets will be made available via Eventbrite.

Programme of events:

June 5th 10am-1pm and 2pm – 4pm Talbot Rice Gallery, Georgian Room

Workshop on experimental composition for animation with guest composer Joseph Hyde (Bath Spa University)

This event is for participants only

June 6th 10am-1pm and 2pm – 4pm Talbot Rice Gallery, Georgian Room

Workshop on experimental animation techniques with animator Jim Le Fevre

This event is for participants only

June 7th 7:30pm Playfair Library, Old College, University of Edinburgh

Reception starting at 6:30pm in the Talbot Rice Gallery

Concert of new works for string orchestra and animation, inspired by creative phases in the life of Norman McLaren

The Edinburgh Film Music Orchestra
Conducted by Yati Durant
www.efmo.co.uk

Free tickets available via Eventbrite

Production Notes

Canto (2015)

***Canto* (2015)**

Introduction

Canto is the final piece composed over the course of my PhD practice. The majority of the pieces I had produced up until this point had been designed to explore specific themes or ideas – *Study No. 2* had explored Kandinsky's theories of colour and their effect on the viewer's perception of being in a constricting or expanding space, with the sonic realm contributing to this by using recordings of confined or open spaces. *Study No. 3: Foghorn Abstraction*, as detailed elsewhere in these production notes, was inspired by combining recurrence in sound composition and geometric abstraction in the visuals. *Skitterling* was originally designed to investigate foreground and background elements in both sound and visuals, *Synthonomy* was designed as a tribute to Norman McLaren without any specific underlying investigation, and *Light, Dark Balance (Study No. 5)* looked at extensive use of synchresis, and the matching of colours to frequency ranges.

For *Canto*, I decided not to have any specific goal in terms of illustrating one or other of the principles I had covered during my research. As this was to be the final film of the degree, I decided to select a specific narrative, to create the piece to follow that guidance rather than any technical exploration, and to examine how much of the principles I had researched could be experienced naturally in the work without a pre-existing plan to highlight them.

A recurring influence on my life during the degree had been depression – not only had I been affected personally by this, many of my friends and family had also undergone severe stress or depression. I decided to create a piece that would acknowledge the powerful effect depression has on those who suffer from it, but also to include an optimistic element to round off the work, to reflect my own recovery from the illness.

Concept

As I had established that this work would be narrative based, concentrating on depression but leading to redemptive finish, I decided to continue on with a theme of a triptych, as I had already used in the previous piece, *Light, Dark, Balance*. This would provide a simple three-stage structure, common to many dramatic narratives. Given the underlying theme of depression, I decided that the three stages would be 'sinking into depression', 'recovering from depression' and 'normality'. These stages reminded me of the *Divine Comedy* by the Florentine poet Dante Alighieri (1265-1321), an epic poem of 100 cantos describing Dante's journey through hell (*Inferno*), purgatory (*Purgatorio*), and finally heaven (*Paradiso*). While these may seem extreme concepts to apply to the stages of depression that I wanted to chart with my composition, the emotions experienced during depression are also extreme, and Dante's journey would provide a rich source of imagery to inspire my own creative work. The opening lines of his first canto recalls the feelings of lack of purpose and clarity that depression brings:

> In [the] middle of the journey of our days
> I found that I was in a darksome wood –
> The right road lost and vanished in the maze.[242]

The division of the poem into 100 cantos, one introductory then 33 each for *Inferno*, *Purgatorio* and *Paradiso* also gave me the name for the work – while used to denote a subsection within an epic poem, it derives from the Italian word for song[243], and therefore alludes to both the compositional process and the inspiration from Dante.

The final decision to make was the duration of the finished work. From the concerts I had curated for the *Soundings International Festival of Sonic Art* held at various points during the first years of my PhD studies, as well as my general research, I had

[242] Dante Alighieri. (1884). *The Divine Comedy of Dante Alighieri* [HTML]. Tr. Sibbald, J.R. Available at http://www.gutenberg.org/ebooks/41537 [accessed 18/08/2015]

[243] Etymology of 'canto' taken from the Oxford English Dictionary online. Available at http://www.oed.com/view/Entry/27287?redirectedFrom=canto#eid [accessed 18/08/2015].

assessed a large selection of contemporary works in audiovisual composition, and had observed the audience reaction to them. From this, my personal judgement was that the longest successful abstract audiovisual compositions had a duration of up to 10-11 minutes – longer pieces where there was no clearly strong narrative or identifiable characters to empathise with seemed to lose the audience's interest. I therefore decided to aim for a total duration of 10 minutes, with a bit of leeway for titles and credits. I was not proscriptive in the time I would allocate to each part of the work, as I wanted the various stages to progress naturally from one to another as the piece developed. In the event, the composition itself lasted for 10 minutes and 10 seconds, with another 50 seconds taken up by titles and credits, resulting in a work with an 11 minute long duration.

Audio Composition

At the start of my PhD, I generally worked in a very specific form of electroacoustic composition, creating acousmatic music where I would take location recordings and apply processes and effects to create a sound object that would fit a specific purpose in my composition. The recording could be altered very slightly, perhaps only to remove extraneous background noises, or have a succession of complex operations carried out on it, making it completely unrecognisable from its source. Over the course of developing my practice, however, I developed an appreciation for other methods of creating sounds for use in my work, particularly computer synthesis either through plug-in effects or signal processing. For this piece, therefore, I decided to exercise a catholic approach and make use of any and all techniques that produced the specific sound I needed, either through manipulating recordings or synthesizing new sounds. As I worked through the piece, I found that I was initially quite reliant on synthesis, but made a greater use of location recordings in the second portion of the work. The final section employed a very simple but powerful combination of one virtually unaltered sound recording and a synthesized slow moving chordal sequence.

To start the recording, and to set up the initial entry into the Inferno, I determined that I would use an initial explosion followed by a long, ominous low frequency continuous roll of thunder. Although I had one field recording of thunder, it lacked an explosive start, and the thunder itself was largely overshadowed by the sound of heavy rain on a nearby tin roof. I therefore made use of a specialised plug-in for Cubase to create the initial explosion and continuing rumble, adapting the preset values for thunder effects to ensure the sounds conformed to my need for a deep ongoing rumble. I wanted to develop even more unease in the lower frequencies throughout this first section depicting hell, and therefore added in two more variants of the low rumbling thunder, adding phasing and discordances throughout the section. In order to achieve a similar sense of unease in the higher pitch registers, I made use of other Cubase plug-ins to synthesize an electrical hum and a cyclic 'chopping' noise, which also interacted with underlying lower pitched rumbles to add to a feeling of discordance.

While these sounds gave a good ongoing textural base to the work, recalling the general unease and malaise as part of depression, I also wanted to add in some foreground sounds to reflect the sudden attacks of anxiety experienced during depression. In terms of the *Divine Comedy*, I found inspiration in the 4th and 5th Cantos of the *Inferno*, describing the initial descent into Hell through the First and Second Circles, aspects of limbo where the cries of unbaptised infants and (relatively undamned) heathens are heard. I created a high-pitched, unearthly scream by convolution, combining the sound of air escaping from a balloon's tightly stretched neck with a recording of my own screams. Depression also makes you antisocial and argumentative, again recalling later passages in the *Inferno* where those condemned to hell strike out and rend each other. I wanted to include sounds where a crowd of voices could be heard, though manipulated to sound guttural and animalistic. I took a recording made in the main railway station in Edinburgh, and applied several filters and effects to achieve this purpose, while still retaining a recognisable humanity in the voices. This recording also featured the approach and rapid transit past of a train, which added a climax and release that I could use to end this particular section. As Dante passes down through to the lower levels the fire and brimstone of Hell die down to bone-chilling cold, until he reaches the final depths where sinners are encased in ice, with even Lucifer trapped there and gnawing on the ultimate traitors Judas Iscariot, Brutus and Cassius. I wanted to reflect this change from heat to cold, and a final pit of despair, so reduced the level of the interlacing rumbles to a lower, more stable form, and lower in amplitude. At the same time, I introduced a high pitched vibrato note, intended to signify the loneliness of those locked in the ice, or sunk into the isolation of their own depression. This would mark the end of the *Inferno* section and the transition into *Purgatory*.

The concept of Purgatory is an interesting one for me – I was brought up in the Presbyterian faith in Northern Ireland, where there is no recognition of a purgatorial state, you are either saved and will enter heaven, or you are not and are condemned to hell. The Catholic view of Purgatory is that it is a place where the impure but not necessarily damned will go, a clearing-house in eternity, where the prayers of the living can improve the individual soul's balance sheet and allow entry to heaven, while the unlamented soul faces a long slow slide downwards to hell. As a fairly devout agnostic, I have no belief in either of these positions, but the idea of Purgatory as a slow journey towards something (hopefully) better resonates with the slow progression from depression towards a more normal interaction with the world. Dante's *Purgatorio* is an island set in an ocean, with several terraces he ascends to reach the gateway to Paradise. This also has the implication of a journey, and also raised reminiscences of my childhood in Northern Ireland, where there were everyday atrocities due to the 'Troubles'. Throughout my teenage years, travel away from Northern Ireland to go on holiday or eventually to attend University represented a feeling of progress away from a violent and unhappy state to a place of enjoyment or release – although I should also note that I lived in a relatively peaceful village, with very tolerant parents, so I was never exposed to the worst of the incidents.

Most of the travel I undertook at this point was by boat, across the Irish Sea to Scotland and then further afield. I therefore always associate the sound of the sea as a sound of progress, of moving on to another place, and therefore I determined that the sounds of water, the sea and associated noises would dominate in this central section.

Having used synthesis to establish the continuing background textural sounds in *Inferno,* I continued to use synthesis to establish a background texture in *Purgatory*. I created a set of phased, synthesized pulsing sounds to convey the longer swelling sounds of ocean waves, still incorporating some of the sense of unease to continue on from the *Inferno* sequence. I wanted to reinforce that sense of unease, and a feeling that the traveller through this section was surrounded by other unfortunates, and made use of a recording of a sea stack colonised by large groups of seagulls and other seabirds, which I had also used in *Study No. 3: Foghorn Abstraction*. I processed this recording slightly to place it at a distance from the listener, with the sounds of the birds distorted and blurring into each other, which increased their otherworldliness and added an ambiguity to the source of the sounds, allowing the listener to imagine they might be hearing the sounds of a distant crowd of people, arguing and bewailing their fate. However, as this was a journey towards a brighter goal, I introduced another recorded sound to overlay the seagull voices. This is a recording of attendants singing in the highly resonant Baptistery of St. John, in the Piazza dei Miracoli in Pisa, Italy. I reversed the original recording, taking out most of the background noises, and then applied several processes to make the sound more resonant and unrecognisably human in origin, although I retained sense of space and reverence gained from the location. To close the section, I wanted to focus the listener's attention on a more dynamic sense of movement and direction while retaining the theme of water, so made use of a recording of water trickling over stones on a beach at Gullane, running down the sand and into the sea. Apart from removing the background sounds of wind and microphone rumble, the only processing applied to this recording was a time compression as the end of the section approached, resulting in a gradual build in tempo and pitch, presaging an upwards movement from *Purgatory* into *Paradise*.

For *Paradise*, I wanted to create a sense of peace and well-being, a paradise of the individual that is achievable in the present moment. I decided to use a location recording of a small forest by a sea loch on the Knoydart peninsula on the west coast of Scotland. This area is accessible only by boat, which provided a continuity from the setting for *Purgatory*, and as there are no roads on the peninsula, the only extraneous noises in the recording are a small amount of microphone handling and an occasional gust of wind. The recording featured a recurring call from a willow warbler as well as a fortuitous fly by of a bee close to the microphone. I decided to use this as the main establishing sound for the section, but also wanted to include a nod to the ideas of heaven established in classic western thought, particularly in Hollywood films such as "A Matter of Life and Death" (1946) and "It's a Wonderful Life" (1946). I composed a sequence of chords, moving from minor to major, using a set of glassy, angelic voice instruments available in plug-ins for Cubase. I gradually brought these chords in under the recorded birdsong, then allowed them to dominate, reducing and then fading out the birdsong. I closed the section and the piece with a slow fade of the final chords.

Video Composition

As I was creating the sound composition for this piece, I had some very strong images forming in my head. The music for the *Inferno* section, with its explosive start and continuing thunderous rumbles, produced the image of a huge fireball and continuous burning flames. The rolling, watery sounds at the start of *Purgatory* inspired visions of a deep underwater setting, while the birdsong and music in *Paradise* made me think of clear blue skies and peaceful open green countryside. Following my normal practice, I created a rough timeline of the textural background movements and foreground gestures in the soundtrack, and noted where specific visual events would occur to match.

For this work, I knew that the animation package that I had used in earlier compositions, Blender, would not afford an immediately achievable path to create the positioning and image manipulation that I would need. I had made use of 3D Studio Max as an animation package for *Synthonomy* and *Light, Dark, Balance*, and considered using this for the majority of the animation for *Canto*, but, having compared the package with other software, settled on Maya as the most appropriate system for the main animation work, which I would then modify in Adobe After

Effects.

The start of *Inferno* is a strong, explosive thunderclap, and I wanted an image that would immediately engage the audience and draw them into the unsettling world implied by the soundtrack. Dante’s Inferno describes a series of circles in hell, each featuring specific punishments for the sinners who have been condemned there. The first sections describe violent, fiery settings, with spirits subjected to burning winds, or being encased in burning mud. Combining the circles, the initial explosion, and the burning mud and storms, I decided that the main central image for my *Inferno* would be a sphere, hanging in the middle of a large shell or cavern, where the size would be impossible to judge but the implication would be that it is immense. The viewpoint would always remain the same, and the sphere would not move itself, hanging static in the void, also conveying that the audience is caught immobile in hell and unable to escape. I added a cracked mud texture to both the sphere and the surrounding shell, to reflect the desolate burning mud in Dante’s hell, and used particle effects to create both an initial fireball and continuing flames across the surface of the sphere. The screaming sounds were matched with vaporous white wisps fading into view, passing around the sphere and fading out again, recalling the souls buffeted in the burning winds. The electrical hums and chopping noises inspired a visual recreation of the plasma globe effect, with glowing purple tendrils extending out from the central sphere, and bright electrical arcs connecting the sphere with the surrounding shell.

As the section continues, and reflecting Dante’s journey further into the frozen depths of hell, I wanted to convey the idea that the heat and flames were dying down, withdrawing into the centre of the sphere. As the background sound textures recalled a cyclical sound of crashing, grinding rocks, I associated that with a movement of lava and magma within the sphere, and used After Effects to cycle through a change of colour within the surface cracks, producing a sullen red/yellow glow on both the sphere and shell, pulsing in time to the textural shifts. At the same time, I started to thin out the flames burning across the surface of the sphere, making the surface initially paler, then lowering the overall light levels, as the light source of the flames disappears and the sphere becomes a dark mass lit by the glowing material within.

The distorted railway noise forms part of this descent into darkness, but with one last flicker of life in the guttering lava glows as the sound of the train rushes past. As this dissipates, the entire screen darkens to a gloomy black, with even the final glowing cracks fading out. The emptiness of the scene is emphasized by the persistent high frequency vibrato note, which marks the transition to *Purgatory* and the start of the next section.

The soundtrack to *Purgatory* was designed to convey the feeling of motion, a journey, specifically through water or ocean. The section fades in from the dull black at the end of *Inferno*, revealing a shifting interlace of green and blue wave-like forms. These were created using particle effects in Maya, recreating the effect of ink drops dropping into and swirling around in a large vessel of water. Three versions of this effect were rendered, then the sequences reversed and layered over each other in After Effects, resulting in a rich interaction between the swirling ink effects. These sequences were particularly resource intensive to create – each one required over three days to render using a render farm composed of 8 interlinked processors. In order to create a dappled effect underlying the ink swirls, one of the sequences was duplicated and a layer effect added, with the effect faded in to coincide with the introduction of the seabird colony sounds in the soundtrack.

Originally, the waving tendrils were intended to also emphasize the seabird sounds, but on viewing, seemed to work well with the processed singing voices from the Pisa Baptistery, and so were adjusted to match this sound object instead. These were created using a lightning special effect available in After effects, but with some extreme changes to the normal settings to achieve the organic tendril effect. I adjusted the parameters for these strands so that they would move slowly, with an occasional swift jerky movement, and change colour in line with background changes in the moving ink swirls.

The reversed ink drop sequences were synchronised so that they would all coalesce and reverse flow up and off screen, timed to coincide with the upwards rising pitch at the end of the recording of the running stream of water. This left a blank clean white screen to end the *Purgatory* section, clearing the way for *Paradise*.

The combination of birdsong and synthesized music in the Paradise section reminded me of blue skies and a sweeping landscape. Accordingly, I decided that rather than trying to recreate nature using a synthetic animated environment, I would use footage of the real world to fuse with the soundtrack in this section. Having explored various options of working with still photographs or video recordings made at variable speeds to allow some manipulation of the visuals, I settled on timelapse photography as providing a suitable image sequence that I could adjust to match movements in vision with movement in the soundtrack.

My first choice of location for the timelapse sequence was in one of the Glens of Antrim, an area of natural beauty close to my childhood home, and recalling summer trips around the region. Unfortunately the DSLR camera I was using at the time developed a fault, which, when combined with the sporadic but insistent rain showers, meant that the footage captured in this session was unusable. The second attempt involved a visit to the Royal Observatory on Blackford Hill in Edinburgh, where I was able to find a suitable image of a large tree framed on each side by Edinburgh Castle and Salisbury Crags, which combined both the beauty of nature, and the man-made vista of Edinburgh, my adopted home. I had calculated a time-line for this section which would start with a vertical shot of the sky, which luckily on an unusually sunny day, was a wide blue expanse with a few white clouds moving away from the camera position. This establishing shot matched the wide open skies I envisaged on hearing the birdsong, while also acknowledging the general belief in Christianity that Heaven is in the skies above us.

The timelapse sequence was planned to feature a gradual tilt down from the blue sky to reveal the green landscape of the tree and the wider city, which provided some complications for the shoot. Motorised tilt heads for camera tripods are difficult and expensive to obtain, and are also too heavy to easily transport to the location I had selected. I decided instead to measure the shots taken closely, and at appropriate intervals, to manually tilt the camera by approximate declinations, to gradually reveal the tree until the camera was fully horizontal. The timelapse interval was 1 second, which resulted in a full shoot of over an hour, with specifically timed points at which I would start and finish the incremental tilt, calculated to coincide with specific moment in the soundtrack, such as the sound of the bee buzzing past the microphone.

The timelapse shoot resulted in a good sequence of high-quality images, although the method of tilting the shot meant that some extensive work was required using After Effects in post-production to lace the tilt downwards together into a relatively smooth motion. Some jerky movements remained, but these were acceptable as a normal feature of timelapse photography, and therefore not particularly distracting from the overall sequence. The photographs also included a number of intrusive momentary images, such as birds flying past, and a few photographs where a family with small children could not be dissuaded from walking across shot. The birds were removed digitally and painstakingly from the initial blue sky sequences, as I wanted the audience to concentrate on the ‘clean’ view of the blue sky and clouds. As the tree comes into shot and begins to fill the screen, however, I retained the birds in shot, as part of the now established natural world. I also digitally removed the family walking across the shot. Some imperfections remained – there were some small marks visible on each photograph, which will have been due to small smears on the camera lens. The whole sequence was graded in After Effects to reduce the flicker from the changing light conditions over the shooting period.

In order to accent the feeling of an abundant natural world, I increased the saturation of the green and yellow tones in tree and surrounding plants, creating a lush, vibrant feeling to the scene. As the music passed its climax and started to fade out, I also applied an edge detection effect and a gradual fade to white, allowing the scene to dissolve gradually into an outline before completely fading out into silence.

Audiovisual Work

On completing the main animation to accompany the soundtrack, I revisited the audio composition to introduce a small crescendo into the opening explosion, as I had decided to establish the opening scene of the central sphere as a lead in to the fireball explosion and subsequent continuous flames. This lead in required an accompanying introduction in sound terms as well, priming the audience for the explosive start. I also created the opening title and the closing credits, which used the same font and text placement, but made use of a contrasting monochrome scheme, with the opening title white text on a black background, and the closing credits as black text on white. This is a deliberate reference back to the earlier research I had carried out on Kandinsky's thoughts on colour. His view of black was that it "is something burnt out, like the ashes of a funeral pyre, something motionless like a corpse"[244] while white by contrast "is not a dead silence, but one pregnant with possibilities. White has the appeal of the nothingness that is before birth, of the world in the ice age"[245]. As the start leads into the void and lack of hope that is hell, it seemed appropriate to use black as the dominant colour at the start, with white as the dominant colour at the end of the *Paradise* section. This colour scheme is retained internally as well – *Inferno* starts and ends in black, *Paradise* starts and ends in white, while *Purgatory* as the bridge between the two starts in black but ends in white. This does, of course, rely solely on a Western artistic and mythological tradition – in some Asian countries, white is associated with death and mourning.

The completed visuals for each section were rendered out from After Effects as png sequences, allowing the transparency and levels of each frame to be preserved in the individual files. These were then imported into Adobe Media Encoder along with the png files for the opening titles and closing credits as well as the completed audio

[244] Kandinsky, W. (1977). *Concerning the Spiritual in Art*. Dover Publications Inc. New York. p39.
[245] *Ibid.*

composition. The entire work was then composited and rendered into a complete assembly, using a high-end codec (Apple ProRes 422). This created a very high quality HD format file, but one which was extremely large at over 7GB in comparison to my earlier work. Another render was later completed using the H264 codec, resulting in a smaller file at about 750MB, which will be the version available for most exhibitions.

Reflections on Completed Work

There is a lengthy examination and critique of *Canto* as a case study in Chapter 5 of the main thesis text, and so I will not include a lengthy analysis in this section. In terms of overall success, I believe *Canto* achieves the original intention of depicting the journey from the bleakness of depression to the normality of recovery, and makes good use of some of the symbolism and inspiration gained from Dante's *Divine Comedy*. Of the three sections, I find *Purgatory* to be the most satisfying in terms of an immersive audiovisual experience – the use of synchresis in this section works particularly well in binding the soundtrack with the swirling, overlapping and branching images. *Paradise*, though my second favourite section, lacks the same sense of immersion, especially as the tree becomes the dominant image – the audience has moved from being a participant in the scene to looking in at a moving photograph, a representation of the world seen in a frame. There are successful moments of synchresis in this section, particularly at the start where the cloud movements appear to match to sonic movements. The appearance of an aeroplane's contrail just as the scene settles to the horizontal view of the central tree is particularly fortuitous, as this also synchretically interacts with the synthetic chords at this point.

The section I find least satisfactory on repeated viewing is *Inferno*, although my reservations concerning it are relatively minor. In comparison to the other sections, it is too long, which means that the central sphere is too dominant. The lack of any development or movement of that central object other than the flames moving over its surface means that it is relatively easy to lose interest in the work, although the sections where the purple plasma glow appears, or where the flames change colour to white in pulsing phases, do counteract this problem to a large extent. The flaws in this section are not sufficient to warrant a re-working of the piece, but will inform future compositions.

The duration of this work is the longest I have attempted over the course of the PhD, and although I feel the work retains enough interest throughout that an audience will not feel it outstays its welcome, I think the duration has reached the limit appropriate to this sort of work. The potential loss of interest in the *Inferno* section is luckily counteracted by the more immersive *Purgatory* and contemplative *Paradise*. As a final composition to round off the research and practice of my course, I am content that this work is consistent and engaging, and that it represents a lot of the skills and techniques I have developed over the duration of the PhD.

Performances

Edinburgh College of Art Masters' Degree Show	*Aug 2015*
Animation Showcase, Edinburgh College of Art Degree Show	*Jun 2015*

Conclusions

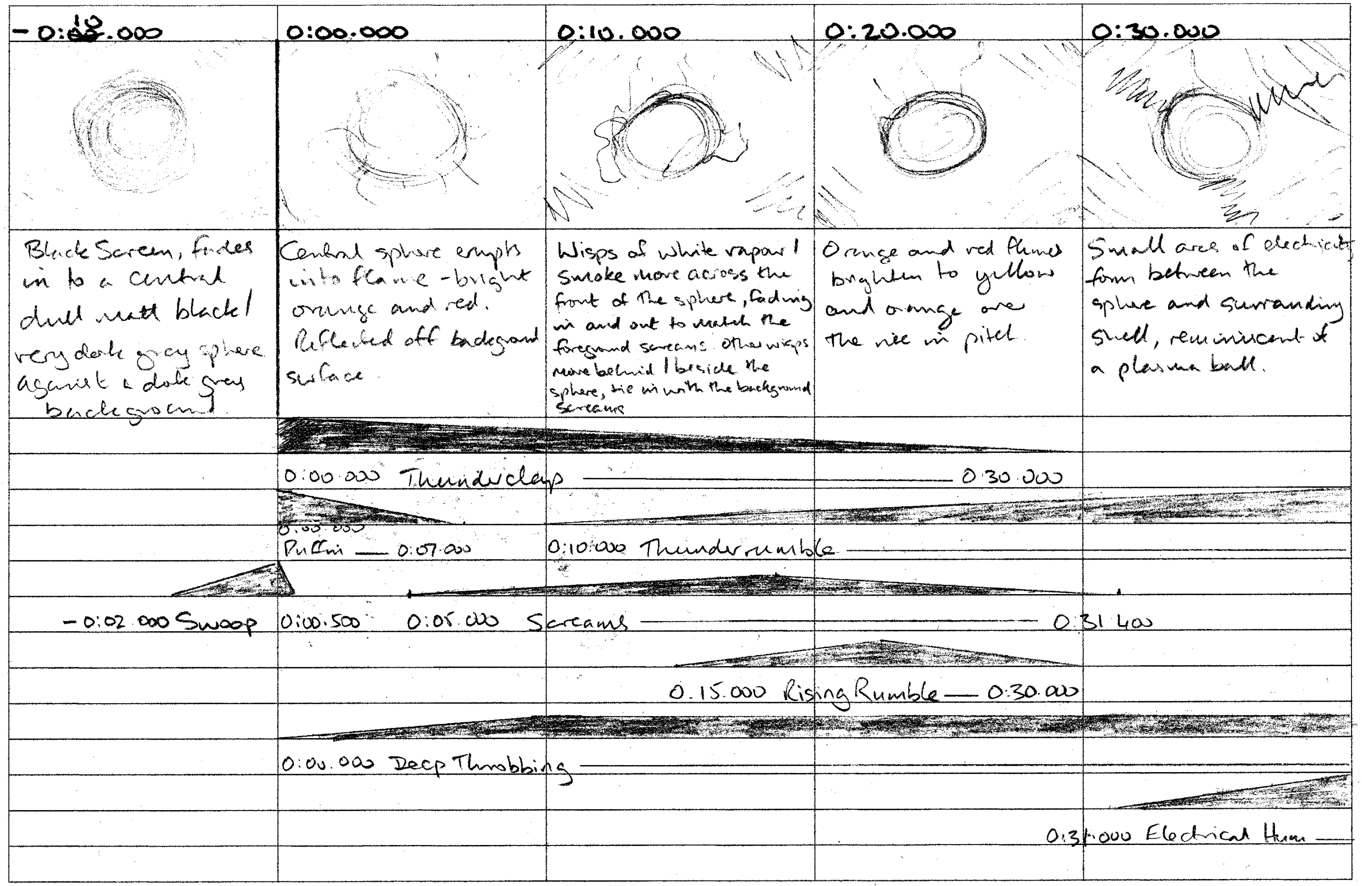
As the final audiovisual composition to be created over the course of my PhD, *Canto* ran the risk of having too many concepts and techniques forced into it. However, part of my development in compositional terms was a confidence in only including audiovisual objects in my compositions that would add to the whole immersive experience, rather than overload the audience with sounds and images as I had done with some of my earlier work. *Canto* was designed to strike a balance between content and technique, allowing some experimentation, primarily in the combination of sound synthesis and recordings on one hand, and the combination of animation created in Maya and the effects available in After Effects on the other. However, the driving principle in the composition was to develop a cohesive structure referencing Dante's *Divine Comedy*. In this, the piece has succeeded in meeting its brief.

As a marker for the end of this period of research and practice, *Canto* also meets my aim of creating an immersive audiovisual work, of sufficient duration to develop and explore a concept or narrative, but not so long that it will lose the audience's attention. While the piece has not yet been performed outwith the Edinburgh College of Art's Animation screening and Masters' Degree Show, I am confident this work will be successful in being selected for concert performances and animation festivals, and provides a solid foundation for further refinement of skill and technique in my future audiovisual compositional work.

Canto (2015)

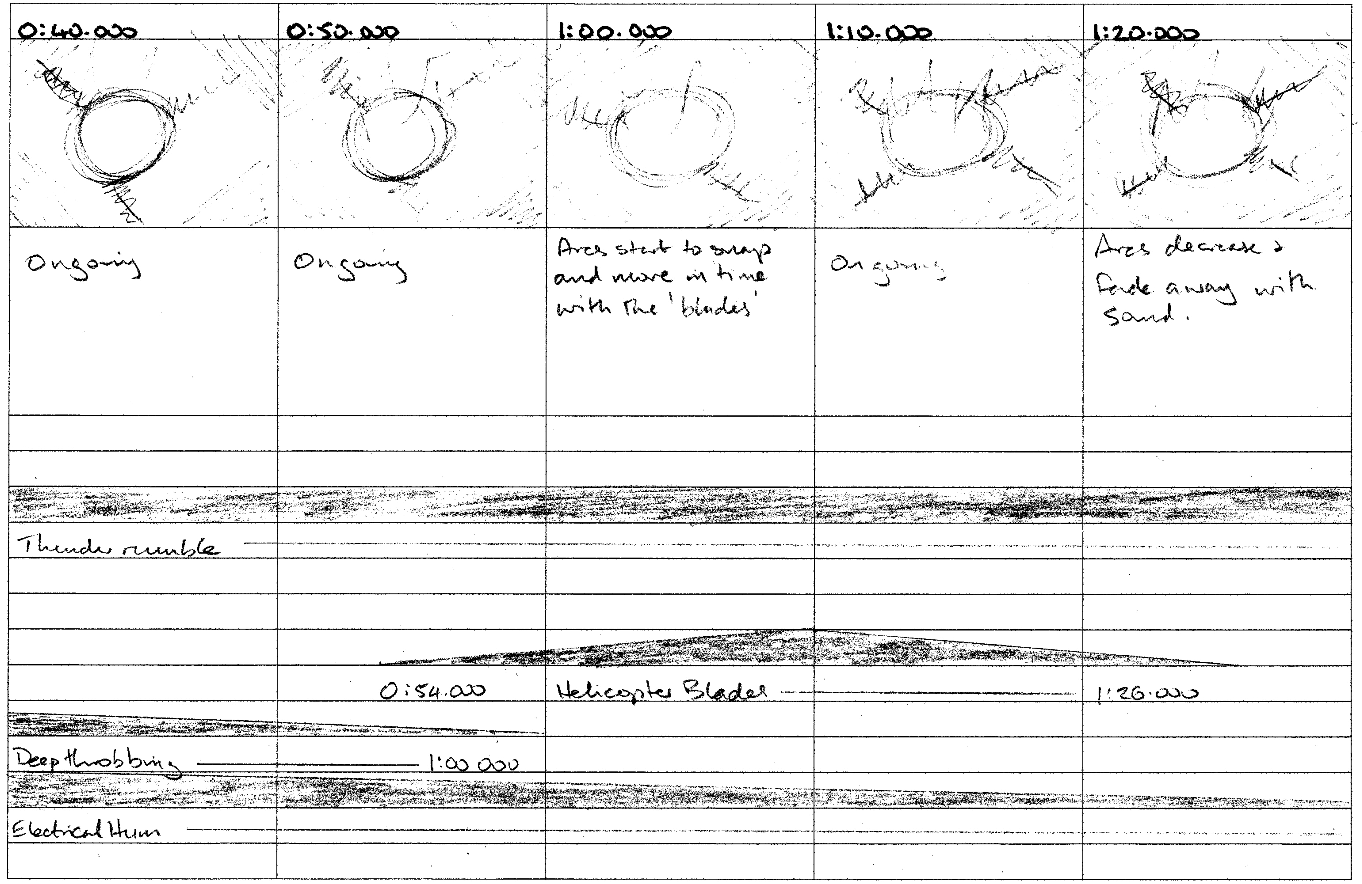
Storyboard

Title: Study #6 (Paradise Lost I:284-5)



1/13

Title: Study #6



2/13

Title: Study #6

1:30.000	1:40.000	1:50.000	2:00.000	2:10.000
Flames move across the sphere in phase to the windy surges	Wisps of white vapour/ smoke move across the front of the sphere, fading in and out to match the foreground screams. Other wisps move behind/beside the sphere, the in sync w/ the back-ground screams	Ongoing	Cracks appear in the surrounding background, with dull red magma glow behind, pulsing in intensity with the phased noise	Ongoing

Thunder rumble ——————————

1:36.000 Screams —————— 2:00.000

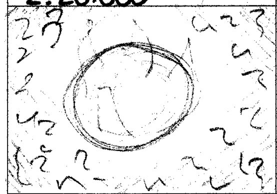
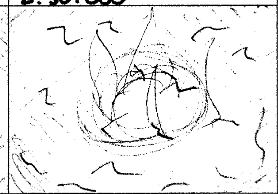
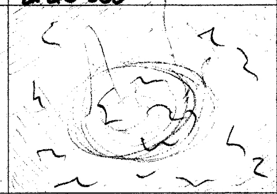
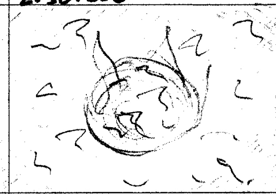
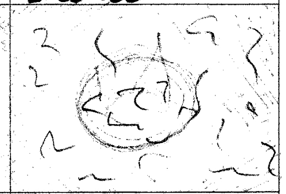
1:35.000 Windy Surge Phased —————— 2:10.000

2:01.500 Crashing ——————

Hum 1:35.000

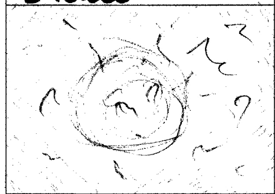
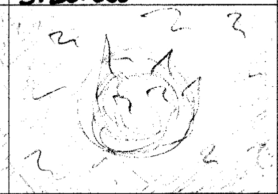
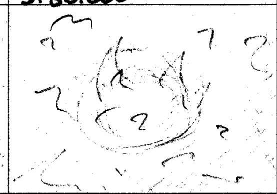
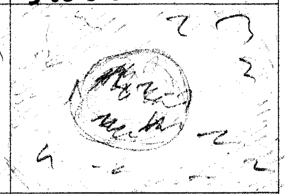
3/13

Title: Study#6

2:20.000	2:30.000	2:40.000	2:50.000	3:00.000
				
Ongoing	Cracks also appear on the central sphere, with dark red magma glowing behind, pulsing in rhythm	Ongoing	Ongoing	Flames start to die down, with embers replacing them as main source of light, but darkening to dark reds
				3:00.000 Distorted Train
Thunder rumble				
	2:33.500 Windy Phase Noise			
Crashing				

4/13

Title: Study #6

3:10.000	3:20.000	3:30.000	3:40.000	3:50.000
				
Ongoing Flames dying down.	Ongoing	Ongoing	Ongoing	As train noise peaks, there is a surge of intensity in the glow coming through the cracks, then fade all out to dull murky black

Distorted Train

Thunder rumble — 4:00.000

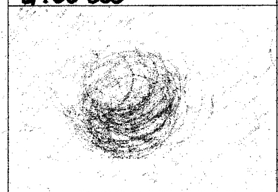
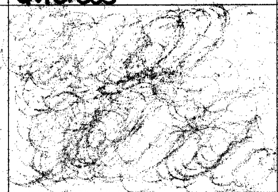
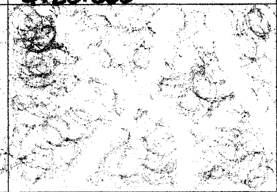
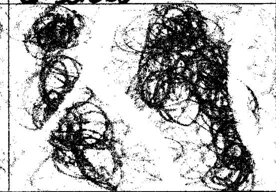
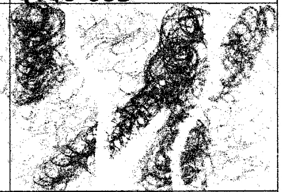
3:58.500 High Pitch

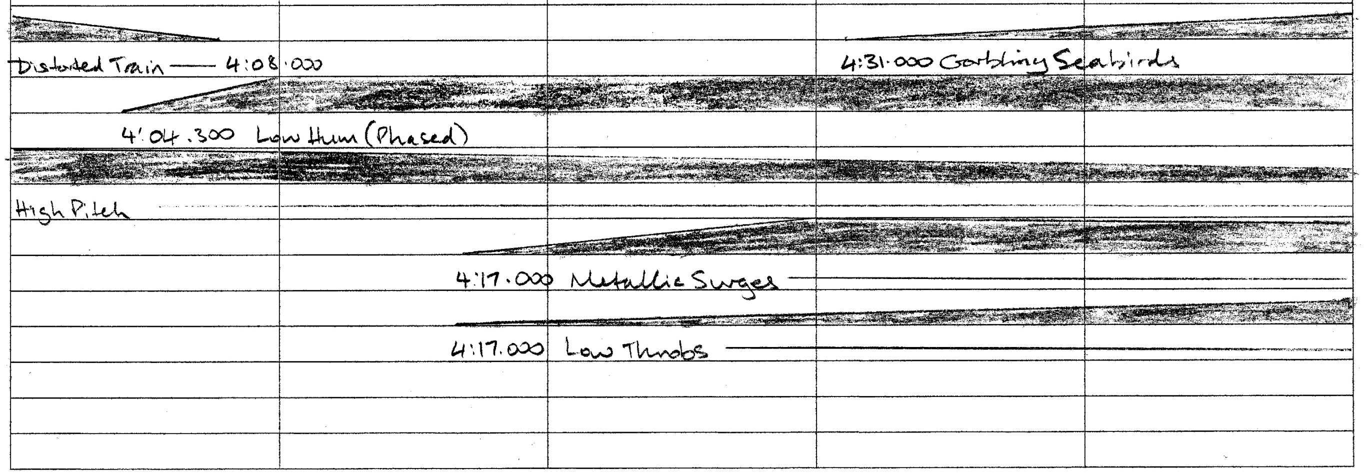
Windy Phase Noise — 3:43.000

Crashing — 3:43.000

5/13

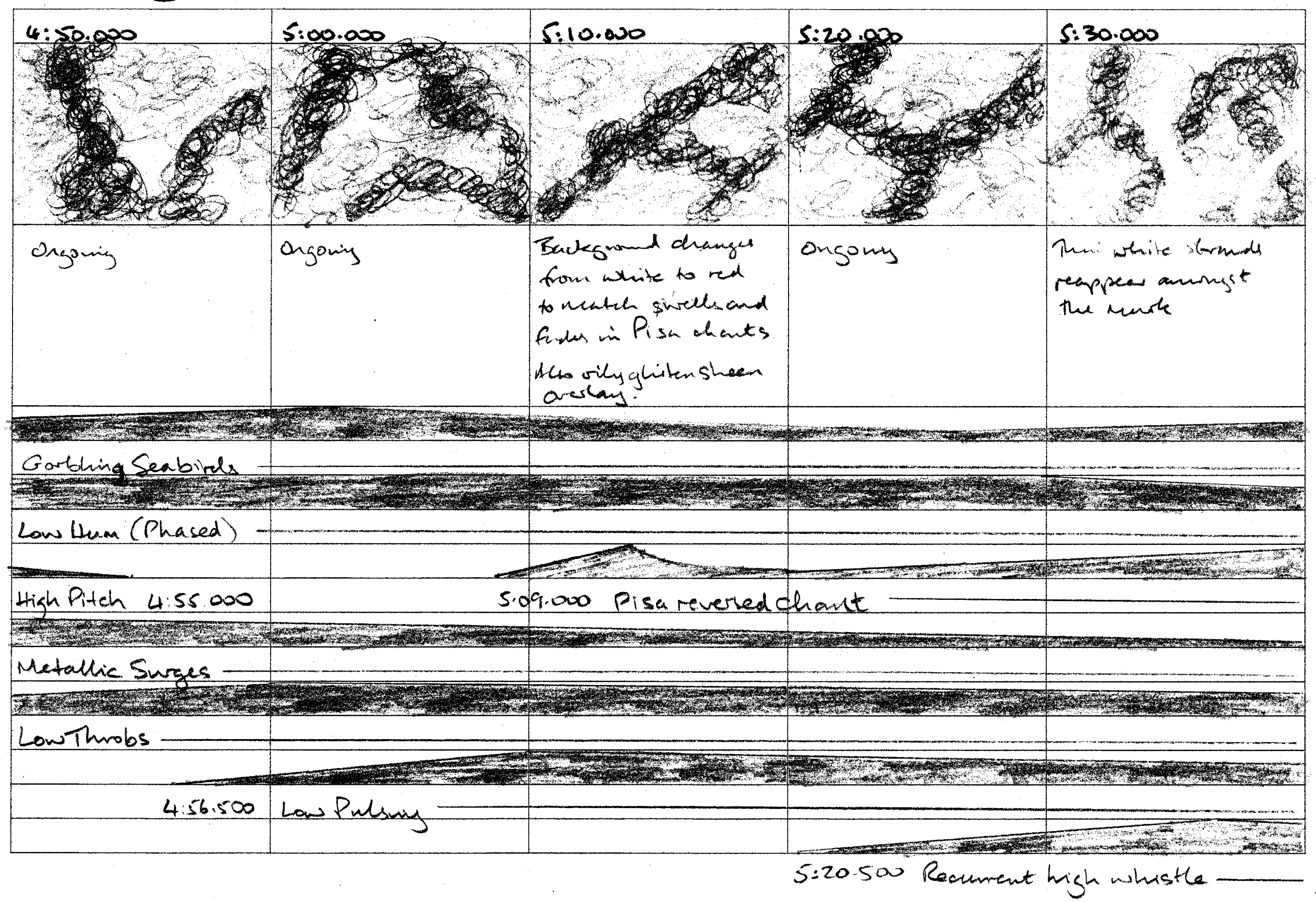
Title: Study#6

4:00.000	4:10.000	4:20.000	4:30.000	4:40.000
				
Against the background of murky black, an overlay of a glistening oily sheen fades in slowly then out slowly in time with high pitch.	Faint swirls of black move within the murk as the black ink begins to coalesce	An overlay of thin white ink strands form amongst the black ink strands, moving in time to the metallic surges	Central black ink strands become more discernable. Pan around and swell to match the seabirds	The background begins to pulse and shift in phase with the low throbs



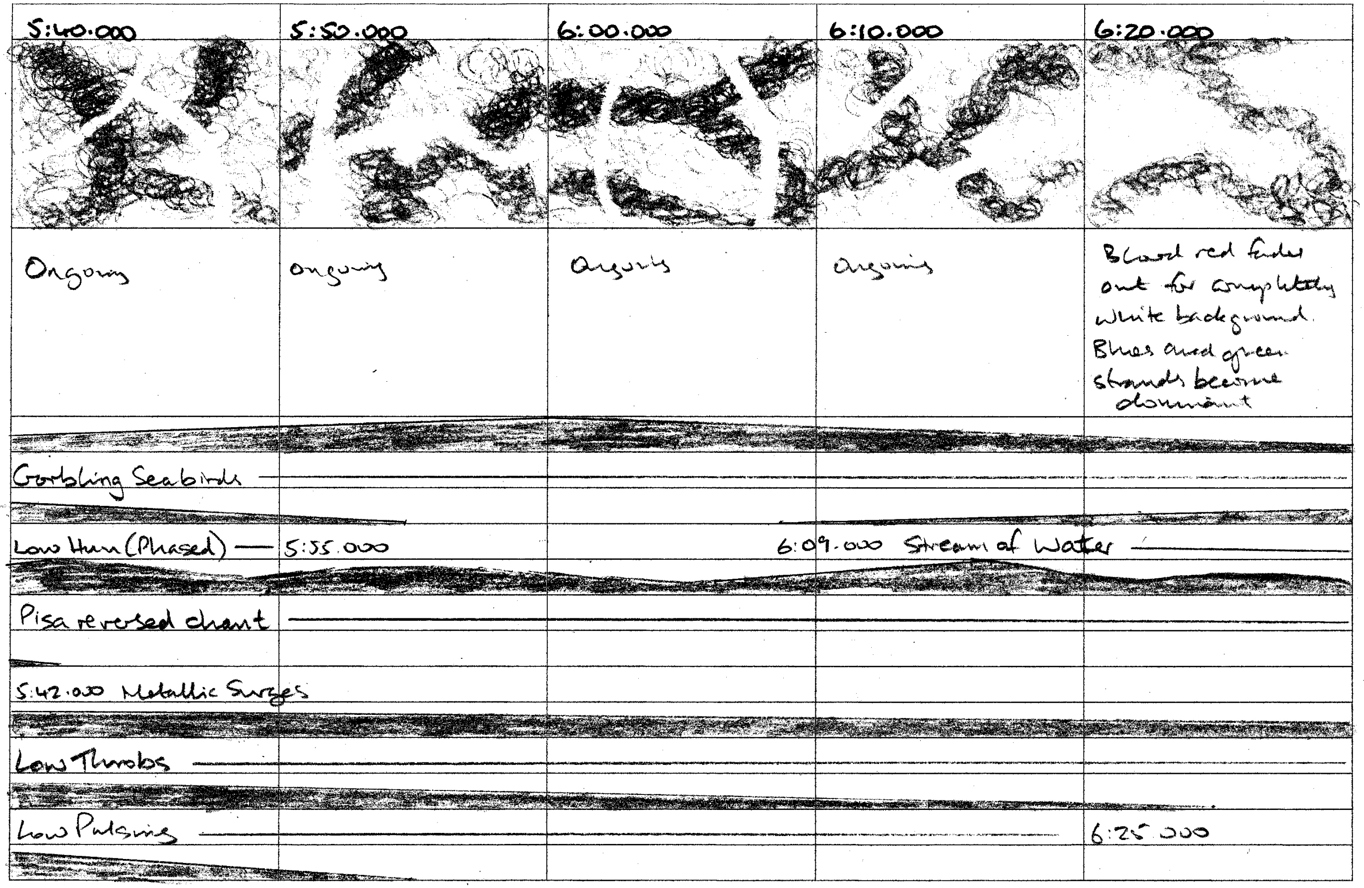
6/13

Title: Study #6



7/13

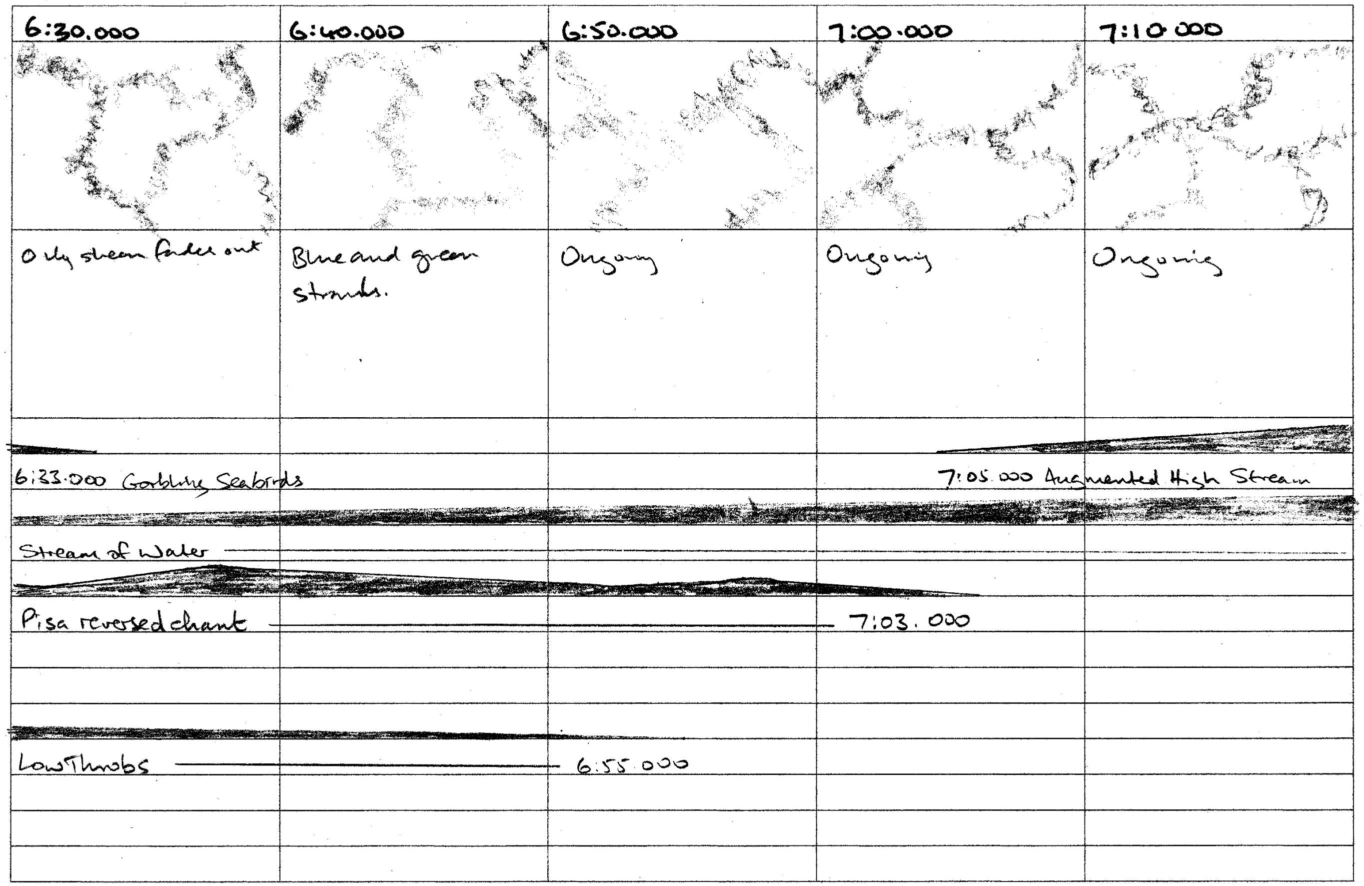
Title: Study #6



Recurrent High Whistle —— 5:55.000

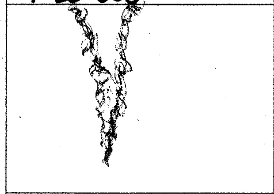
8/13

Title: Study #6



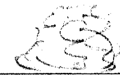
9/13

Title: Study #6

7:20.000	7:30.000	7:40.000	7:50.000	8:00.000
				
'Drop' disappears up and off top of screen	White screen, fades slowly to blue sky with clouds.	Clouds moving quickly across blue sky	Ongoing	Ongoing
7:26.550 Augmented High Stream				
7:26.550 Stream of Water				
	7:30.300 Recurrent Bird Song			
		7:38.000 Rising minor chord		

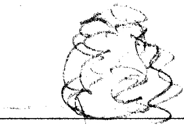
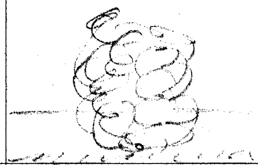
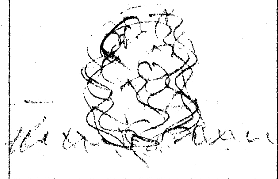
10/13

Title: Study#6

8:10.000	8:20.000	8:30.000	8:40.000	8:50.000
				
Ongoing	Tilt starts - the top of the tree comes into sight at the bottom of the frame.	Ongoing	Ongoing	Ongoing
Recurrent BirdSong				
Rising Minor Chord				

11/13

Title: Study#6

9:00.000	9.10.000	9:20.000	9:30.000	9:40.000
				
Tree continuing to be revealed by tilt - supersaturation of greens and yellows starts.	Ongoing	Ongoing	Tree fully revealed	Gradual fade to white begins, with edge detect visual effect.
Recurrent Bird Song			9:36.000	
(Modulation) Major chord				
			9:29.000 Glassy Chimes	

12/13

Title: Steady #6

9:50.000	10.00.000	10.10.000	10.20.000	10.30.000
		CREDITS →		
Gradual fade to white	Gradual fade to white & out	CREDITS →		
Major Chord	10:06.000			
Glossy Chimes	10:06.000			

13/13