



Universiteit
Leiden
The Netherlands

Imagined Voices : a poetics of Music-Text-Film

Kyriakides, Y.

Citation

Kyriakides, Y. (2017, December 21). *Imagined Voices : a poetics of Music-Text-Film*. Retrieved from <https://hdl.handle.net/1887/58691>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/58691>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/58691> holds various files of this Leiden University dissertation.

Author: Kyriakides, Y.

Title: Imagined Voices : a poetics of Music-Text-Film

Issue Date: 2017-12-21

Imagined Voices

A Poetics of Music-Text-Film

Yannis Kyriakides

Imagined Voices

A Poetics of Music-Text-Film

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden

op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,

volgens besluit van het College voor Promoties

te verdedigen op donderdag 21 december 2017

klokke 15.00 uur

door

Yannis Kyriakides

geboren te Limassol (CY)

in 1969

Promotores

Prof. Frans de Ruiter

Universiteit Leiden

Prof.dr. Marko Ciciliani

Universität für Musik und darstellende Kunst, Graz

Copromotor

Dr. Catherine Laws

University of York/ Orpheus Instituut, Gent

Promotiecommissie

Prof.dr. Marcel Cobussen

Universiteit Leiden

Prof.dr. Nicolas Collins

School of the Art Institute of Chicago

Prof.dr. Sander van Maas

Universiteit van Amsterdam

Prof.dr. Cathy van Eck

Bern University of the Arts

Dr. Vincent Meelberg

Radboud Universiteit, Nijmegen/ docARTES, Gent

Faculteit der Geesteswetenschappen

Academy of Creative and Performing Arts (ACPA)

Table of Contents

Acknowledgements	1
Introduction	4
PART I: Three Voices	16
Chapter 1: The Mimetic Voice	17
1.1 Art Imitates	18
1.2 Cognitive Immersion	24
1.3 Vocal Embodiment	27
1.4 Subvocalisation	32
1.5 Inner Speech	35
1.6 Silent Voices	38
Chapter 2: The Diegetic Voice	41
2.1 Narration	43
2.2 Paratext	46
2.3 Narrational Network	48
2.4 Temporality	51
2.5 Frames	54
Chapter 3: The Multimodal Voice	58
3.1 Opsi Melos Lexis	60
3.2 Media Correlation	65
3.3 Metaphor Hierarchy	77
3.4 Asymmetrical Balance	82
3.5 Case Study: <i>Subliminal: The Lucretian Picnic</i>	83
Chapter 4: Historical Perspective	96
4.1 Image as Language	98
4.1.1 Intertitles	98
4.1.2 <i>Anemic Cinema</i>	100
4.1.3 <i>Television Delivers People</i>	102
4.2 Language as Image	103
4.2.1 <i>Zorn's Lemma</i>	104
4.2.2 <i>So Is This</i>	108
4.3 Music as Language	111
4.3.1 Surtitling & Music Video	111
4.3.2 <i>The Cave</i>	113
4.3.3 <i>A Letter from Schoenberg</i>	114
4.3.4 Other Examples	114
4.3.5 <i>Ballade Erlkönig</i>	115
4.4 Language as Music	119
4.4.1 <i>Perfect Lives</i>	119
4.4.2 <i>Ursonography</i>	120
4.4.3 <i>Traité de have et d'éternité</i>	121
4.4.4 <i>Le Film est déjà commencé?</i>	125
4.5 Image as Music	127
4.5.1 <i>I...Dreaming</i>	128
4.5.2 <i>Newsprint</i>	129
4.6 Music as Image	131
4.6.1 Three Music Videos	132
4.6.2 <i>It Felt Like a Kiss</i>	133
4.7 Summary	136

Part II: Music-Text-Film	138
Introduction	139
Chapter 5: Internal Monologues	141
5.1 <i>Dreams of the Blind</i>	143
5.2 <i>Mnemonist S</i>	151
5.3 <i>Memoryscape</i>	158
Chapter 6: Unanswered Questions	167
6.1 <i>Machine Read</i>	169
6.2 <i>Dodona</i>	174
6.3 <i>Norms of Transposition (Citizenship)</i>	176
Chapter 7: Voiceprints	183
7.1 <i>Wordless</i>	184
7.2 <i>Varosha / Disco Debris</i>	189
7.3 <i>Der Komponist</i>	197
Chapter 8: Interactive Scores	204
8.1 <i>Karaoke Etudes</i>	207
8.2 <i>Trench Code</i>	212
8.3 <i>Oneiricon</i>	221
Conclusion	228
Appendix: Additional Works	233
<i>Simplex</i>	234
<i>The Queen is the Supreme Power in the Realm</i>	236
<i>Scam Spam</i>	238
<i>QFO (Queer Foreign Objects)</i>	240
<i>RE: Mad Masters</i>	242
<i>The Arrest</i>	244
<i>Circadian Surveillance</i>	246
<i>Nerve</i>	248
<i>True Histories</i>	250
<i>8'66" (or everything that is irrelevant)</i>	252
<i>Walls Have Ears</i>	254
<i>Music for Anemic Cinema</i>	256
<i>MacGuffin</i>	258
<i>The Lost Border Dances</i>	260
<i>The Musicians of Dourgouti</i>	262
Bibliography	264
Audiovisual Media Citations	272
Links to Online Media of Music-Text-Film	274
Summary	276
Samenvatting	279
Biography	282

Acknowledgements

I would like to dedicate this thesis to the memory of my dear friend and colleague Bob Gilmore, who passed away several years after I began this research. Without his initial encouragement, support and insight I would never have got off the mark. I still remember, during the first years of research when we would bump into each other on social occasions, he might ask with a pinch of his trademark polite irony: "Yannis, dare I ask how the research is going?" to which I would answer with an expression of discomfort, and an excuse about composition deadlines. Now having finished the thesis, I have a tinge of regret that he couldn't see more of it, or have more of an influence on it in the last years. However, in writing this thesis I would occasionally punctuated it with an imaginary footnote that would read: "...wonder what Bob would've made of that?" His voice is somehow still present in many pages.

Having decided to continue with this research after Bob's passing, I am indebted to my three supervisors for helping me finish the work. Catherine Laws, who I was very happy to be in contact with again, after being undergraduate students together at York University, helped me enormously with structuring the ideas and getting me to think in ways that are more academically acceptable. Marko Ciciliani, one of my closest friends and musical colleagues, is the one person I could think of, who has such a deep affinity and understanding of this kind of music and multimedia, and is an innovative composer himself. I have valued our conversations and collaborative work throughout the years. The comments and mark-up of the thesis were of great help, both to understand what is important in the arguments and giving significant suggestions and contributions. And naturally, Frans de Ruiter, without whom I would not have been able to get this over the line. I am also indebted to him for arranging the practical possibilities allowing me the possibility of following this research trajectory, while maintaining composing and teaching commitments.

I would also like to express my gratitude to Henk van der Meulen, Henk Borgdorff and Martijn Padding from the Koninklijk Conservatorium for supporting this research trajectory with time and money, as well as the rest of the composition faculty there, Peter Adriaansz, Cornelis de Bondt, Calliope Tsoupaki, Guus Janssen and Diderik Wagenaar, and to the many amazing students who I've had the privilege of teaching and discussing these ideas with.

I have had the chance to publish two articles in the last year based on material from this thesis: *Hearing Words Written*, in *Organised Sound*, Volume 21, Issue 3 (Sound and Narrative), Cambridge University Press, dealt with some of the ideas about narrative found in Chapter 2 and 3, as well as focusing on a number of my

compositions. Thanks to Leigh Landy and James Andean for this opportunity. A part of Chapter 6: *Unanswered Questions* was also published in Tijdschrift Kunstlicht, Volume 37, 2016 No. 3/4 (*Translation as Method*) at the end of 2016, many thanks to Marianna Maruyama (editor) for inviting me for this issue and to Steyn Bergs (co-editor) for reading and suggesting improvements to the text.

As a composer I have been privileged to be able to work with incredible musicians and artists in the last years, some of whom are responsible for the work presented here. Musical experimentation of the most enjoyable order with my MAZE fellow members: Anne Le Berge, Reinier van Houdt, Dario Calderone, Wiek Hijmans, Gareth Davies, and the previous incarnations as Ensemble MAE with Bas Wiegers, Fedor Teunisse, Michel Marang, Barbara Lueneburg, Koen Kaptijn, Karolina Bater, Jelte van Andel, Noa Frenkel, and Kristina Fuchs. Other ensembles I have had the pleasure of working with mentioned in this thesis are: Veenfabriek, ASKO | Schoenberg, musikFabrik, Champ D'Action, Okapi, The Electronic Hammer, Slagwerk Den Haag, Ragazze Quartet, Kronos Quartet, Brodsky Quartet, Ergon Ensemble, Ensemble Artefacts, Jeugd Orkest Nederland, Philharmonie Zuidnederland and the Athens State Orchestra. Organisations such as Gaudeamus, November Music, Suspended Spaces and Holland Festival. Individual musicians such as Takao Hyakutome, Lore Lixenberg, Ilan Volkov and always and everywhere with Andy Moor. Thanks also to some of the other artist, writers and software designers that I collaborated on this work include Paul Koek, HC Gilje, Joost Rekveld, Isabelle Vigier, Reinaldo Laddaga, Mehmet Yashin, John Mcvey , Darien Brito, Andrea Vogrig and Mirko Lazovic.

Finally I would like to express love and gratitude to my wife Ayelet Harpaz, not only for supporting this additional workload the last years, for giving constant artistic feedback on my projects, but also for finding time to proof read and correct the manuscript.

Introduction

Words is a two edged sword,
That block the road and open door.¹
(Lee Scratch Perry)

In 2002 I started exploring a form of music-multimedia combining text projection with music. At first this was a way of exploring what would occur when meaning or narrative is imposed on a musical structure, where the music itself has no pre-defined narrative intention. The idea was that the on-screen words were not being used to either amplify or translate what was being spoken, played, sung or visualised, but rather to interfere with the listening perspective of the music. From these initial experiments two questions arose, and eventually became points of reflection for further development: Who is narrating? And where is the voice located?

These questions became more pertinent after I noticed a strange phenomenon occurring during performances of these works: that when we read text synchronised to music, we become very aware of an inner voice silently reading along. The recognition of this phenomenon (that it was not just me imagining it) became clearer in discussions with audience members after performances and presentations of these pieces. While the difficulty of processing words and music varied between native and non-native speakers, and between musicians and non-musicians, the awareness of the act of reading was an experience shared by the majority of them.²

In these music and text pieces, words are read by the spectator at the pace set by the music. Sometimes the speed of text is too slow or too fast for any realistic inner utterance, so that the text becomes suspended or attaches itself to whatever layers of music are being heard at that moment in time. In this way, meaning comes and goes. It is constantly being reconstructed without forcing any clear-cut interpretation. The avoidance, in most cases, of a clearly perceivable outer voice, is one of the strategies that opens up the space for the music to become a surrogate voice. The voice-like properties of music, and the semantic arsenal that it carries, give a sense, or perhaps the illusion, that something is being uttered by 'the music'.

¹ Lyrics from *Two Edged Sword*, 2011, Lee Scratch Perry & Warrior Queen on: EP *Profit*, Have-A-Break-Recordings.

² Musicians often have more problems processing language and music at the same time than non-musicians, possibly because music and language have been shown to trigger activity in similar parts of the brain, and because musicians tend to have a more articulated, syntactic reading of music, which might conflict with the processing of language. This has been explored in neurological studies by Mireille Besson and Daniele Schön in their article: "Comparison Between Language and Music" (Peretz & Zatorre 2003: 269).

This effect of hearing one's own voice in the music, becomes an added dimension that I had not initially predicted. It was a discovery that had many consequences for the ways in which I subsequently approached composition and ideas about listening. For many years I held the belief that a musical work should always be approached or analysed from the perspective of the listener. I felt that only the listening experience gives music any sense at all, as it is the listener that is supplying the meaning. The revelation that some of the side-effects of these projected text works could possibly lead to an increased self-awareness by the listener of their role when listening to music was an exciting one. More questions followed: Which form does the subjectivity of the listener take? Is it only the extent to which a musical narrative is interpreted in a different way by each individual? Beyond the so-called 'subvoicing' of words, does the listener's inner voice also filter and echo the sonic information being received by the senses? Is there a case to be made, stating we use our inner voice to test out certain sounds we hear for syntactical/melodic structures, and that we also map these sounds onto our own imagined vocal apparatus? Perhaps a silent imitation of the music is also taking place in our minds while we listen? These questions were part of the reason I wanted to embark on the research trajectory, contemplating the different implications of these kind of multimedia pieces and seeing where it could lead.

Definitions

One of the first quandaries, which persisted throughout the course of the research, was what to call this genre of work in which music is played along-side projected text. As a hybrid art form that is not so widespread, there is no generally accepted terminology. In the beginning of the process, I described the pieces as 'music and videotext'. I liked the way the blended word 'videotext' looked on the page, but it was confusing as a label because of another association, it is often used as an alternative name for 'teletext', the information system that delivers (or used to) computer-style text on television.

In 2013 the German media centre ZKM (Zentrum für Kunst und Medientechnologie) curated a major exhibition called *Schriftfilme, Schrift als Bild in Bewegung*, focusing on what we might translate to English as 'artistic word-films', or 'scripted-films'. This was an overview of the major analogue and digital films using animated words, from early silent film intertitles to new media. Even though 'Schriftfilme' in German has a persuasive ring to it, for the catalogue publication it was translated as 'typemotion', which seems to have a more design-oriented use, describing the animated word and letter sequences as used in advertising and film titling. In the realm of digital poetry or 'videopoetry' there have been various names given to these types of works: 'video-visual poetry', 'poetronica', 'poetry video', 'media poetry',

'Cin(E)-Poetry' or 'video poem opera'³. In music, some of the names used by individual composers include 'moving word sequence', used by John Oswald to describe his work *Homonymy*, or 'reading piece' used by Peter Ablinger for works such as *A Letter From Schoenberg*. Other possible descriptions that have been used include: 'on-screen' words or text which are generally used for text on computer and hand-held devices, 'kinetic typography', and 'moving text' which both highlight the animation of letters and text.

I pondered for a long time whether it was still relevant to use 'video' as terminology, motivated by Yvonne Spielmann's definition of video as not just an analogue technology, but a medium in itself which also encompasses the digital realm (Spielmann 2008). But her arguments are almost ten years old now, and from what I understand of current use of the term, it is almost out of use. So why reference an even older technology in the use of the word 'film', when these works do not even make use of this analogue technology? The reasons have partly to do with the fact that the historical examples I draw from, are largely from an era when film was indeed the relevant medium – partly also to the ubiquity of the term 'film', that it is no longer used just for the medium itself, but to describe a process and an object of moving images – and partly to the more poetic reference of its original meaning, a thin layer of material, that in this context might cover the music with a trace of textuality. 'Music-text-film', I therefore felt, highlights the relationship between the trichotomy of media, with the emphasis on music as the dominant medium.⁴

Defining what kind of work music-text-film encompasses was also important as a way of demarcating the territory of pieces I would be examining, both my own and those made by others. I came to define it as: **the use of projected text together with music, where the words are not amplifying or translating what is being spoken, sung or visualised**. This latter point is vital in that it eliminates its predominant use in film, pop video, and theatre, where there usually occurs a doubling or redundancy in the overlaying of a text, carrying a meaning which is already being expressed in another medium. This often results in over-emphasis of meaning, or a 'sandwiching' of one medium between two others, that might be said to constrain or force one particular point of view. Similarly, with surtitling in opera, the text is projected both visually and through the voice of the singer onstage, as a result of which the audience does not have to look for the narrative source; it is being reinforced through many other layers of media. What I wanted to emphasize in my own music-text-film pieces is that there is a discernible dynamic between the media: the text functions as an independent voice, not as a double.

³ This is used specifically to refer to Gianni Toti's multimedia works.

⁴ The sense of hierarchy of media is an important aspect that will be developed later in the thesis, mirrored in Aristotle's 'melos opsis lexis' and Barthes' 'music image text'.

At this point in the introduction, as a way of clarifying the definition of music-text-film, I would like to take a short interlude to present a work of mine for cello and text-film that highlights these issues, and also opens the question of what is being communicated with words or music.

Words and Song Without Words

Words and Song Without Words is a short work for cello and text-film. Originally commissioned by the Amsterdam Cello Biennale and first performed there by Larissa Groeneveld on 29th October 2012.⁵

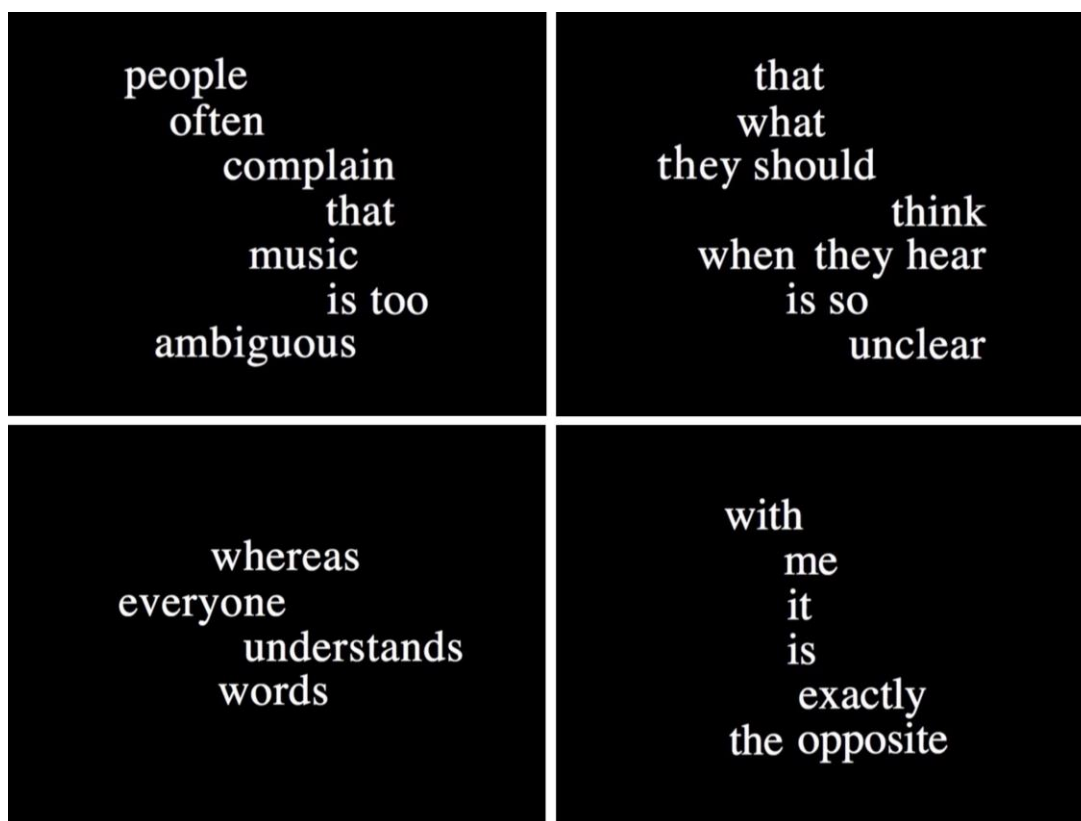


Figure 1 Stills from *Words and Song Without Words*.

In the course of this research, I came across an interesting statement by Felix Mendelssohn. This statement grew to become a question in my mind and eventually resulted in a new composition; I cannot not say I agree with Mendelssohn's point of view, but the citation and the subsequent piece provide a useful introduction to the concepts and compositional techniques I have been exploring in this thesis.

⁵ It subsequently featured in a documentary portrait on my music made for NTR Dutch television by Interakt (directed by Ditteke Mensink), and was awarded the International Rostrum of Composers (IRC) prize in 2014, by the International Music Council, NGO official partner of UNESCO. The link below is from a recording by Francesco Dillon. [weblink: <https://vimeo.com/54731855>]

The statement is from a letter to Marc-André Souchay dated 15th of October 1842, where Mendelssohn clarifies his position regarding his ongoing *Songs Without Words* cycle. Souchay suggests that Mendelssohn can add words to these pieces to make them into actual songs, prompting this reply:

People often complain that music is too ambiguous; that what they should think when they hear is so unclear, whereas everyone understands words. With me it is exactly the opposite, and not only with regard to an entire speech, but also with individual words. These, too seem to me so ambiguous, so vague, so easily misunderstood in comparison to genuine music. The thoughts which are expressed to me by music that I love are not too indefinite to be put into words, but on the contrary, too definite. The same words never mean the same things to different people. Only the song can say the same thing, can arouse the same feelings in one person as in another, a feeling which is not expressed however by the same words. Words have many meanings, but music we could both understand correctly.⁶

Discussions of the ways in which music conveys meaning have persisted for centuries, and Mendelssohn's remarks offer a refreshing viewpoint. Specifically, he points to the notion of the indefinite relation between a language and meaning itself, and to the subjectivity of that position. Where Mendelssohn is more in line with his peers, is in his idea that music communicates emotion, and that the emotional states music can conjure are, to him, more relevant, coherent and meaningful, than if they were described in words. I stress the point about the 'emotions' because, whereas he begins by talking about "the thoughts which are expressed to me by music", Mendelssohn moves on to underline the idea of "feelings that are aroused". What Mendelssohn takes a stand against, is the idea that music can indeed mean something, but it can never be as defined and context specific as language.

Mendelssohn prioritises the emotion over other semantic functions, and compares the immediacy of music in conveying this to the perceived ambiguity of words. Implicit in this is the viewpoint that the context or narrative described in the words of a song are not necessary to convey an emotion.; that emotion is more immediate when it is unmediated by language.⁷

My own position is that both music and words, equally ambivalent, and equally embedded in a multiple layers of context, are communicating *something*; and this communication is facilitated by the voice of the words, the voice of the music and, most importantly, the voice of the listener. I overuse the word 'voice' here very

⁶ Letter to Marc-André Souchay, October 15, 1842, cited from *Felix Mendelssohn: Letters* (Mendelssohn 1946: 313).

⁷ In literature on this subject, a similar attitude is expressed by many classical musicologists, namely that the main characteristic of music is that it expresses emotion, and that there is a common language of tonality in much music that can trigger similar emotional responses, but for there to be a meaning in this emotion, a context is required (Cooke 1959: 21).

deliberately, as I wish to stress the importance of subjectivity in the act of communication and that, to paraphrase Marshall McLuhan, the 'voice is (also) the message': 'voice' is a useful metaphor, because it refers to the act of speech or song, as well as to its use in the creation of identity and narrative. In this thesis and in my compositions, I try and highlight how voice, in the narrational sense, emerges between the levels of different media; how voice is transformed and embedded in a music, and how this can activate inner voices during the act of listening.

What Mendelssohn's *Songs Without Words* and my music-text-film have in common, is that they both highlight the absence of voice. One could further claim that this absence leads the listener to an even greater awareness of voice because of, and not in spite of, its absence. In Mendelssohn's piece, the case of our inner voices following the contours of the music is even more pronounced, because melody is so clearly emphasised and given a vocal quality. The absent vocal lines, as suggested by the piano, find their completion in the subvocalisation of the listener, who might trace the melody with their inner voice.

In the case of my piece *Words and Song Without Words*, the subvocalisation occurs between the projected text and the contours of the cello voice, which could be said to exist in a space somewhere between singing and speaking. Like in Mendelssohn's piece, it is a vocalise, a song without words, but because the words are re-injected back into the mix in another form, in an act of translation from one communication medium to the other, the song breaks down into speech, and the semantic meaning of the text comes under strain when forced into such a direct relationship with music. Loading each word with the rich possibility of musical meaning opens it up to many other possible interpretations; the flow of syntax is slowed down and deconstructed until ambiguity starts to creep in. We understand that communication is taking place, that a voice is addressing us - Mendelssohn, the composer, the musician or his instrument; but because of the slow pace of unfolding, we have time to hear our own voices reflected back into the space between the text and the music.

The type of signification occurring in music-text-films such as this, across the main media, and within ontological thresholds, vary enormously. When the word 'meaning' is used in this thesis, it is meant not only in the linguistic sense that Mendelssohn has misgivings about, but broadly to the many types of meaning discussed within semiotic theory, such as conceptual, connotative, affective, social, reflected, collocative and thematic (Leech 1974: 9). Furthermore, the three main distinctions of meaning set out by semiotician Charles Morris are crucial in understanding the subtle difference of signification generated between various ontological levels: *semantic* meaning, generated between the sign and entities in the world, *pragmatic* meaning, between the sign and the user, and *syntactic* meaning, between different signs of the same or similar order (Morris 1938: 6). These categories are useful in distinguishing between what kind of semiosis occurs

between the various media and the audience. For instance, taking the first phrase in *Words and Song Without Words*, 'people': the *semantic* meaning could both refer to the image we have of a collection of humans, and at the same time we hear a sound associated with the instrument we know in the world that is called 'cello'. The *pragmatic* meaning could be the emotional or expressive effect this combination of signs has on us, and the *syntactic* meaning could refer to both the relation between one word and the next, one musical phrase and the next, but also between music, word and visual representation. In this way, my use of *syntactic* meaning in the thesis, is not limited solely to the semiosis generated between words, but between different signs of words, sounds and image, both interlingual and intralingual. What I would like to further underline, and which will be explained in detail in Chapter 3, is that the type of meaning generated is largely dependent on what the dominant media hierarchy is at any given moment.

Structure of Thesis

I have chosen to call this thesis a 'poetics.' There exist two thoughts behind this: first, I wanted to place the focus primarily on the form, meaning and implications of music-text-film. Rather than deal with theoretical, aesthetic or other philosophical questions discretely, I wanted to approach them as they arise out of, or through commenting on, this particular artistic practice. Many of the questions I deal with in the thesis are born directly out of my practice of creating these pieces, though they also naturally feed back into them; sometimes as a way of testing out a theory or exploring another possible path, which has emerged from the research. Secondly, the themes that I have chosen to structure the theory around come directly from the mother of all poetics: Aristotle's⁸. The keywords for the first three chapters are based around terms strongly associated with the *Poetics*: 'mimesis', 'diegesis', and the trichotomy of media: 'melos, lexis, opsis'.

The first two chapters elaborate on Plato's binary distinctions of art: mimesis and diegesis (imitation and narration). In Chapter 1, I begin with the basic definitions, centred on the idea that art is by nature imitative, and develop the idea of mimesis, not in terms of how art mirrors the world, but how the spectator mirrors the artwork. The question of to what extent the spectator is implicated in the artwork, the relation of immersion versus critical distance involved in music-text-film, is defined as an intermediate state of 'cognitive immersion', not fully immersed but engaged on a certain cognitive level, where the spectator is projected into the artwork. I employ psychologist James J. Gibson's theory of 'affordances' (Gibson 1977) so as to be able to discuss the embodied experience of the spectator not simply as subjective experiences but, rather, framed in terms of the possibilities afforded by

⁸ *Poetics* (Περὶ ποιητικῆς) by Aristotle from around 335 BCE. In the thesis, I largely refer to George Whalley's 1997 English translation.

the musical works. Proceeding from various theories of embodiment, I discuss subvocalisation, silent reading and the inner voice from historical as well as psychological perspectives.

The following are the three forms of inner vocality that, I argue, are activated by music-text-film: 'silent reading' as in the reading of the text; 'silent singing' as in the tracing of melodic contours with the inner voice; and 'silent discourse' the hidden dialogue of thought that occasionally surfaces during overt self-reflexive moments in the works, or when the half-completed syntax of words triggers a myriad of possible answers.

Chapter 2 develops Aristotle's conception of diegesis, the art of narration, as developed through the work of literary theorist Gérard Genette and the field of narratology. This is elaborated into questions about how narrative operates in a musical context and specifically a music-multimedia form such as music-text-film, where one can say that there is an overt narration, but no single narrator. The necessary conditions for narration are discussed, specifically highlighting the relation between narration and voice: the focus of the narrative that is then given over as perspective to the spectator. How is this perspective created in music-text-film? The 'paratexts' of an art work are discussed – the texts or the frame existing around the work – and what kind of perspective or narration they offer. Can we see the words in music-text-films as paratexts in some way? The concept of the narrative voice is developed, as is the question of the extent to which the spectator is implicated in this. The idea that for narration to exist there have to be two distinct ontological levels is one of the conclusions that drawn from this. However, while it is much easier to detect these levels in literature, when temporal differences exist between the narrator and the narrated, how can we define this music? One of the concluding observations is that ontological levels are also demarcated by differences of media. The chapter ends with a reference to cognitive scientist Lawrence Barsalou's theory of 'frames', as a basis from which to formulate a model of how meaning might be generated between different perceived layers in music, as well as in music-multimedia.

Multimedia art is the principal focus of Chapter 3. Aristotle's trichotomy of media, *melos*, *lexis*, *opsis*, forms the basis of a discussion of the history, hierarchy and opacity of media, as well as notions of what in fact constitutes a medium. I go on to propose two different models of analysing multimedia: the first based on the correlation of six different aspects of media, which I define as sync, scale, space, story, style and sentiment. I analyse selected works as a way of examining what kind of understanding this can grant us about the relation of media, how some aspects converge or diverge on different levels.

A second model of analysing media is proposed by looking at how hierarchies of media are manifested in the artwork. This is a model inspired by the idea of how hierarchies of metaphor are constructed, and is specifically drawn from ideas of conceptual metaphor developed by philosophers George Lakoff & Mark Johnson (1980). The creation of media hierarchies within a work, is in my view one of the principal ways by which focus is created, which in turn affords perspective for the spectator. In order to explore both methods, I analyse one of my early music-text-films, *Subliminal: The Lucretian Picnic*, a work which has a clear and distinct use of at least three principal media.

In Chapter 4, I trace a history of text-film organised not in chronological order but in terms of metaphoric relations between the two dominant media. This, again, demonstrates the way in which perspective is dependant on the particular art practice these works emerge from, as well the cultural context. These include pieces that have had a significant influence on my own work: Marcel Duchamp's *Anemic Cinema*, Hollis Frampton's *Zorn's Lemma*, Michael Snow's *So Is This*, Dick Raaijmakers' *Ballade Erbkönig*, Robert Ashley's *Perfect Lives* and Isidore Isou's *Traité de bave et d'éternité*.

The second part of the thesis is devoted to the discussion of my own music-text-film pieces. In recent years I have written about thirty works which use projected text in some form, in a music or sound art context. I decided to include a reference to all of these in the thesis, some through only a few lines of commentary in the appendix, others with a more in-depth analysis using the tools discussed in Chapter 3. I include such a broad representation of work, to demonstrate the many possible perspectives inherent to this form of work. These pieces have some shared themes and some similarity in techniques, but broadly speaking I see a development in the main concerns, from the earlier pieces, dealing with dream or memory in first-person narratives, through interactive sound art installation work and the manipulation of the spoken voice, up to my most recent music-text-film works, dealing with the interactive video score. I have charted this progress in the four chapters that make up this part of the thesis.

Chapter 5, 'Internal Monologues', deals with three ensemble works that highlight first-person narratives derived from conscious or semi-conscious discourse taking place within the mind: *Dreams of the Blind*, *Mnemonist S* and *Memoryscape*. All these works use projected text in different ways, with sober to more playful graphics, and all have prominent electronic soundtracks relating in different ways to the music for ensemble. There is an interesting reconfiguration of the hierarchies of the media in these three pieces, and thus this is one of the issues that are discussed in depth.

Chapter 6, 'Unanswered Questions'⁹, deals with a video and two installations that explore question and answer structures across media. It begins by discussing ideas concerning the translation of language to music and vice versa, various encoding practices and question structures. In the pieces cited, *Machine Read*, *Dodona*, and *Norms of Transposition (Citizenship)*, the translation of words into music and back again reinforces the ambiguity inherent to these specific question and answer structures. In all three cases, a request for information and the ensuing answer, translated into another medium, reinforces the problematic communication structure involved, where the questions are never answered coherently.

'Voiceprints', Chapter 7, concerns work where the material is based on the manipulation of spoken voices. The projected text in these pieces provides a hint as to the spoken content, or the context in which the voices are found. *Wordless* is constructed out of interviews from which the words are removed to leave only paralinguistic. *Varosha /Disco Debris* are two versions of the same concept, assembled from a mass of granulated voices, with a narrator leading the listener through the remains of a ghost town. The orchestral work *Der Komponist* is based on a short time-stretched fragment of speech by composer Helmut Lachenmann.

Chapter 8, 'Interactive Scores', looks at my recent work dealing with text-films, that verge towards musical notation. The idea being, that instead of having one visual code for the musicians and a separate one for the audience, the two could be merged. These works are *Karaoke Etudes*, *Trench Code*, and specifically *Oneiricon*, which is an interactive app-score that acts as notation, sound-generator, book and source for visualisation, for both audience and musicians. The act of reading is shared; the veil between musicians and audience is lifted. These latter works represent an interesting direction in the use of technology which I see opening up in the future, wherein the roles of musicians, the function of notation, sound generation, processing and audience environment become blurred and interchangeable.

The initial three theoretical chapters of the thesis provide a foundation by which to discuss the music-text-films. They are intended as way of understanding these works using three distinct perspectives: the '*mimetic*', how the spectators mirror themselves in the artwork, project their own voice within the frame of the piece; secondly the '*diegetic*', how narration, focus, and meaning is constructed in these works, specifically the relationship between the narration within the music and the narration in the projected text; and thirdly, how the kind of meaning that is generated is depending on the hierarchy of media, and how this hierarchy is constructed. All these aspects are in some way contingent on the point of view of the spectator, yet they are treated as three distinct ways of approaching the discussion of

⁹ A reference to both Charles Ives famous work and to Leonard Bernstein's 1973 Harvard interdisciplinary lectures.

music-text-film. The theoretical foundation of the thesis is not intended as seamless and over-arching theory of how inner voice, and narration are intertwined within the listening experience, the scope of such an endeavour lies outside my expertise. The difficulty of explaining the listening experience outside an authoritative semiological framework is undeniable, but I hope that I have compensated this by providing enough examples and discussion of a variety artwork, including my own, where these issues are brought to the fore and problematised.

Imagined Voices

Voice, then, arises at the crossroads of words and music not because it can, or may, do so but because voice in general just is that which arises at the crossroads... Voice must always be understood as a plural - even at the cost of grammar: what *arises* at the crossroads of words and music *are* voices, and the threads binding one voice to another are always tangled. (Kramer 2014: vii)

When I set out to write this thesis, I began with a title set in the singular, 'the imagined voice'. I thought that what I was searching for in music-text-film was a singular inner voice, a bridge between the words we read and the sounds we hear. Gradually, a multiplicity of voices started making their appearance: the mimetic voices that I call 'silent reading', 'silent singing', and 'silent discourse'. Though sometimes too reticent to make a very conspicuous presence, these voices still interchange roles, and as musicologist Lawrence Kramer suggest above, are always entangled. When I started digging into the diegetic voice, the same thing happened: there was not one singular voice that I could pinpoint in a musical narrative, but many voices with many viewpoints. The narrational perspective also began to be entangled in my mind with the listening perspective, and I began to ask myself the question: how can a voice simultaneously see and listen; how is a voice an ear as well as an eye?¹⁰ From this was born the idea of a narrational network of voices, created between different ontological levels within the music alone, and across various media and senses.

Perhaps I should start this thesis by defining exactly what I mean by 'voice', as the word is already packed with so many interpretations as a noun, verb and even an adjective. Though there are definitions aplenty here, perhaps it is exactly this mutability of meanings that 'voice' represents: both the essence of expression and of listening, inextricably interconnected. As Mendelssohn stated: "the same words never mean the same things to different people".

¹⁰ It was around this time that Bob Gilmore penned the title to his short monograph on my work, *The ear of the voice of the eye*, published by teleXpress for November Music 2011.

PART I:

Three Voices

Chapter 1: *The Mimetic Voice*

One of the bedrocks of the history of aesthetics in western culture is Plato's formulation of 'mimesis' and 'diegesis'. In a dialogue between Socrates and Adeimantus in Book 3 of *The Republic*, Plato sets out to differentiate all forms of poetry in 'mimesis' and 'diegesis', which roughly translate respectively into 'showing' and 'telling':

One kind of poetry and story-telling employs only imitation - tragedy and comedy, as you say. Another kind employs only narration by the poet himself - you find this most of all in Dithyrambs. A third kind uses both - as in epic poetry... (Plato, *The Republic*, 394c)

Here Plato sets out the basic duality between narration and imitation. In diegesis the poet or narrator is speaking in their own voice, never leading the audience into thinking they are anyone other than that. In mimesis, the poet utilises imitation, and takes on the persona of another, by voice or gesture to show, to act out, as is the convention of much staged drama. In later chapters of *The Republic*, Plato expresses his prejudice against mimetic art, which he considers inferior, because in his view it simply copies the appearance of the real, reproducing shadows rather than shedding light on truths. This is somewhat ironic, because Plato himself utilises the form of the dramatic dialogue in much of his writing, using the voice of Socrates as a medium to channel his ideas. (Farness 1991: 23)

These Platonic definitions of narrative are clear-cut and to some extent polarising in their categorisation, especially when much poetry (defined by Plato as everything from comic drama to lyric poetry) can embody varying degrees of these functions, let alone when we discuss more contemporary art forms. It is useful, therefore, to consider that art embodies varying degrees of both mimesis and diegesis, and that their functions are deeply entangled.

In this chapter I would like to appropriate the word 'mimesis' as a way of describing the process by which the voice is engaged in listening and reading music-text-film. I will begin by giving an overview of the concepts associated with classic 'mimesis' and how it informs our understanding of what art is. In the twentieth century, 'mimesis' has been associated with ideas of the 'simulacrum' and the 'hyperreal', reflected in the immersive experience of interactive art or art practices, which call for a high degree of engagement from the listener. In this sense, 'mimesis' is not only exclusive to something enacted by a performer on stage, as was intended by Plato's original use of the word, but is used to describe the process whereby the artwork is transferred to the body and mind of the spectator. I explore how instead of the artwork being a mirror of reality, (one's subjective) reality becomes a mirror of the

artwork. As Plato asks: "Or do you think that someone can consort with things he admires without imitating them? I do not. It's impossible" (Plato, *The Republic*, 500c).

In the section 'Cognitive Immersion', I highlight the apparent contradictions in the experience of the music-text-film, namely of being absorbed or immersed in the musical reading of the text, while at the same time having the critical distance to generate meaning about the interaction of words and music.¹¹ The inner voice of the audience is participating in the silent reading while at the same time remaining at a distance, making sense of the narrative that is being generated between words and music.

The last three subsections of this chapter deal with what I perceive to be three distinct inner voices generated by music-text-film. Firstly, I borrow ideas from music theorist Arnie Cox's essay "The Mimetic Hypothesis" (2001), to highlight how our voices are activated when listening to music, how the voice follows the contours of melody and gesture in the music we are listening to or recalling from memory, in a form of 'silent singing'. Secondly, I deal with the widely discussed phenomenon of 'silent reading', which sometimes entails a complex interchange and modulation of voices, moving from the image we have of our own voices to that of the imagined author or characters in the text. Lastly, the third and more elusive inner voice is the voice of thought; the often dialogic interactions taking place under the hood of our brains, between different mental processes, between different aspects of the self.

1.1 Art Imitates

Plato's original conception of the term 'mimesis' underpins his view, that most art concerns itself with the imitation of nature and reality. The word appears in Chapters II and III of the *Republic* in a general discussion about poetics and education, as he tries to show how it can undermine the state's ideals of truth and justice. According to Plato, the mimetic aspects of art are considered inferior because they merely imitate reality and truth. Although Plato advocates the use of certain stories to educate the young, he reflects an idea cherished by many a totalitarian regime, suggesting the censoring of narratives depicting depraved, violent, sexual or politically sensitive material. Furthermore, his notion of certain kinds of mimetic art as useful comes close to a concept of propaganda: art at the service of political utility.

¹¹ This dichotomy is broadly reflected in the polarity of mimesis and diegesis, as the former is usually associated with immersion and the latter with critical distance.



Figure 2 Plato's Allegory of the Cave by Jan Saenredam, according to Cornelis van Haarlem, 1604, Albertina, Vienna.

According to Socrates, 'mimetic narrators' are not to be trusted, as they undertake an act of concealment which they use to create the basis for further deception, so that the persona they impersonate is fragmented into a manifold. Beyond the conception of mimesis through political ideology, in book X Plato sets out his infamous metaphors on illusion versus the real in the allegory of the cave, where the dangers of taking merely shadows as the representation of reality are indicated. In this allegory, Socrates portrays consciousness through the image of a group of people chained to the wall of a cave. They watch shadows of things happening outside projected on the wall of the cave, mistaking this for reality. Only the philosopher, free from the prison of the cave, has the insight to see the world beyond the shadows on the wall, and perceive true reality.

In another metaphor, mimesis is a mirror, inadequately reflecting what already exists in the world, and so failing to offer anything in terms of essence on its own:

of turning a mirror round and round –you would soon enough make the sun and the heavens, and the earth and yourself, and other animals and plants, and all the, other things of which we were just now speaking, in the mirror. (Plato, *The Republic*, 596e)

In the allegory of the couch, Plato goes on further to suggest, that since art imitates appearances it is twice removed from the real; removed from the world of pure idea and also that of form.

Plato argues that the 'mimetic' artist should be banished from the state. He seems to consider art only through his political lens, since controlling images and words is at the heart of political power. Philosopher Giorgio Agamben, in *Man without Content*, turns this idea into an interesting suggestion, arguing that the fact that the artist has no place in the ideal state stems from Plato's fundamental understanding of the power of art, rather than his misinterpretation of it (Agamben 1994: 4).

René Girard, who in his anthropological philosophy redefines the 'mimetic' as a basic mechanism by means of which desires are borrowed from others, also writes of Plato's recognition of the importance of mimesis, together with his hostility towards it:

If Plato is unique in the history of philosophy because of his fear of mimesis... he is also deceived by mimesis because he cannot succeed in understanding his fear, he never uncovers its empirical reason for being. (Girard 1978: 15)

Plato's identification of poetry with the concept of imitation, secondary knowledge, femininity, emotional depravity and suchlike terms, has its roots in the patriarchal tendencies of his society, which was just beginning to move away from traditional oral culture, as manifested by Homer, to the burgeoning technology of the written word. This itself reflects an interesting cultural shift from the power of the voice in oral tradition to the power of written text.

In the *Poetics*, Aristotle addresses the issues brought up by Plato and the problem of the too-politicised interpretation of the arts. He veers towards the side of poets, stating:

These representations or imitations are communicated in language which may be through terms in current usage or include foreign words and metaphors: these and many modifications of language we allow to the poets. In addition, the same standard of correctness is not required of the poet as of the politician or indeed of poetry as of any other art. (Aristotle in Whalley 1997: 154)

Whilst critiquing some aspects of Plato's theory, Aristotle holds onto some fundamental definitions from the *Republic*, namely that art is essentially imitative. Where he does venture further than Plato is to assert that art, by not simply being a mirror, a copy of reality, can be said to embody its own rules and conventions. He also argues that rather than being the enemy of reason, art gives us philosophical

insight into the condition of man, and he goes further in drawing analogies between the internal laws of poetry and the laws of the natural world (Aristotle 1997: 67).

Aristotle's love of the tragic theatrical form is clear in much of his *Poetics*, and many of his arguments rest on the power of this medium to explore the human condition beyond even the limits of his own medium, that of rational thought. Matthew Potalsky interprets this as "The fictional distance" that "allows a glimpse into the universal qualities of human life that are revealed by particular actions and characteristics" (Potalsky 2006: 37). One can say that what Aristotle advances in his model of mimesis is that art is not only a mirror to the world but also a mirror to the spectator.

This Platonic-Aristotelian conception of mimesis and art has persevered in countless variations and forms into the modern age. The term has not only served as a key concept in artistic discourse but also in much philosophical writing about the dichotomy between the represented and the representation, between nature and culture. The model of the simulacrum is one such manifestation in post-structuralist philosophy, where the basic tenet of Plato, that art is an imitation of something real, is undermined with the Deleuzian concept of an image without resemblance: "The copy is an image endowed with resemblance, the simulacrum is an image without resemblance" (Deleuze 1990: 257).

According to another post-structuralist philosopher, Jean Baudrillard, the simulacrum is not a copy of the real, but a "truth in its own right" (Smith 2010: 102). Whereas Baudrillard uses this as a negative critique, philosophers Gilles Deleuze and Jacques Derrida use simulacra as a way to deconstruct or challenge any accepted system. Since a work of art is contingent to the culture and conditioning of the viewer, it has already broken free from any single conception of an original. Deleuze uses the example of Andy Warhol's image of the Campbell soup can to demonstrate the independence of the simulacrum from the original. Derrida takes as an example a poem by Stephane Mallarmé, *Mimique* (1886), to show that when a text is referring to a book about a performance of a mime act (of Pierrot tickling his wife to death), complex ontological levels are brought into play, creating an ambiguity as to who the original author actually is, and making it very difficult to differentiate the simulacrum from the original. To both philosophers, the only way to escape the dominance of the Platonic paradigm of truth and mimetic falsity is through the simulacrum of mimesis: "Any attempt to reverse mimetologism...would only amount to an inevitable and immediate fall back into its system" (Derrida 1981: 207).

The concept of the simulacrum, all pervasive in today's hyperreal culture of the network, was already fertile in the works of Theodor Adorno, Walter Benjamin, and Guy Debord who, like Baudrillard, return the concept to a Platonic foundation of sorts, by stressing the political consequences in their critique of Western culture,

discussing copy-culture and highlighting, (and sometimes celebrating) the loss of a sense of authenticity.

Musicians have always directly referred to the outside world in the 'mimicking' of sounds of nature, events, or voices, transcribing these sounds and their associations into forms reproduced by voices or musical instruments. Examples of this abound, from Australian aboriginals mimicking the sound of dog growls on the didgeridoo, to Olivier Messiaen's bird transcriptions in *Catalogue d'Oiseaux*. The relevance of the concept of mimesis in the discussion of music becomes more complex in the age of recording media, where the real and its representation, could be said to be intertwined.

The incorporation of the real into the art space has certainly undermined the carefully constructed illusion of the mimetic world. It is interesting to note that when recording media first began to be extensively used in the world of art music, notably in the work of the French 'musique concrète' artist Pierre Schaeffer, it came packaged in the concept of 'reduced listening', in which one is not supposed to hear the sound of a train as an actual representation of a train, but as an abstracted sonic event. The challenging of that idea by the composer Luc Ferrari (amongst many others), in his *Presque de Rien* series, where he simply recorded a sonic landscape with very little editing or manipulation and presented it as a composition, shifted the paradigm in terms of the blurring of the real and artificial.¹²

John Cage could be said to have had an influential role in this shift, with his philosophy of regarding the sounds in our environment as music. However, it is interesting to note that, for all the musical revolutions that Cage initiated, and for all the importance his music and ideas have had for late twentieth century art practice, his attitude to the idea of how sounds can function in a musical domain tends towards that of 'reduced listening'; that is to say, in Cage's terms we can hear sounds as having 'musical' potential in a rather abstract sense, rather than taking sound, along with all its causal, contextual, cultural and semantic baggage, to challenge the ontological space of the music. An interesting side-note is that Cage had a motto, which he ascribed to Indian philosopher Ananda Coomaraswamy: "The role of the artist is to imitate nature in her manner of operation"(Cage 1946: 17).¹³

Notions of realism, which have held a strong attraction in various phases of art history, are still relatively fresh in music; due to the current technologies available to artists, one is now able to create a very faithful sonic double of reality. In the past, music has perhaps lacked the realism available to visual media such as painting and

¹² A much earlier example of photographic sound art is *Weekend* by Walter Ruttmann (1930).

¹³ In actual fact it was Thomas Aquinas who was the original source of the saying: "Ars imitatur naturam in sua operatione" ("art imitates nature in its workings") (*Summa Theologiae*, 1a 117), who himself was paraphrasing Plato's concept of mimesis.

literature, and has tried to compensate for this in the sense of emotional realism and mimesis found in much of the theory of classical and Romantic music. One can see advancement of technologies, which signpost the development of the history of art as merely subsequent steps in creating a more 'realistic' mirror of the world. From painting to photography, to film, to virtual reality scapes; each medium raises the stakes when it comes to more absorbing and meticulous depiction of the real.¹⁴

It is difficult to define what realism truly is in art, beyond the reproduction in a sensorial mode of certain aspects of our experience of reality. We may even come across a philosophical paradox, in that the 'real' is often equated with the 'truth', and that according to some philosophical standpoints, 'truth' is exactly that which cannot be known through sensory experience.¹⁵ The 'real' is often used to describe the 'truth' beyond outward manifestation of reality, and the 'truth' is used often as the justification for an artist's idea of holding a Platonic mirror to reality.

It is interesting, therefore, to consider whether there exists a discrepancy between the intentions of an artist to depict a 'reality', and the viewer's perception of that representation. When a sound artist plays back a recording of a forest at dawn through speakers in a concert hall, the listener does not for a minute mistakenly think that they are actually in that forest, just as a viewer seeing a photograph or film of the space is under no illusion that what they are doing is sitting in a space, viewing some form of art. These artists do not set out to deceive the viewer on the level of distinguishing the difference between the real and the copy. Indeed, the artist engages in highlighting this distinction, bringing to the fore questions of representation; questioning our idea of reality and making us see an aspect of that reality from a different perspective: transformed, layered or simply experienced through another consciousness.

The distinction between a Platonic definition of mimesis, the mirroring of the world, and an Aristotelian one, a way of understanding the structure of the world through convention, is crucial in differentiating between the reality, the object of art and the experience of it; as well as the residue of the experience: what remains, what is learned through this experience, and how we are changed through it. Philosopher Roland Barthes' understanding of the function of 'realistic detail' of a work of art, formulated in his essay "The Reality Effect", brings together these two concepts of mirror and convention, examining the use of descriptive detail in the writing of Gustave Flaubert:

'Concrete detail' is constituted by the direct collusion of a referent and a signifier; the signified is expelled from the sign, and with it, of course, the possibility of

¹⁴ Art historian Oliver Grau charts this trajectory in his 2003 book *From Illusion to Immersion*.

¹⁵ One of the central concerns of 'Rationalism vs. Empiricism' (Markie 2017).

developing a form of the signified, i.e. the narrative structure itself... The truth of this illusion is this: eliminated from the realist speech-act as a signified of denotation, the 'real' returns to it as a signified of connotation; for just when these details are reputed to denote the real directly, all that they do, without saying it, is signify it... *we are the real...* it is the category of 'the real'... which is then signified... (Barthes 1986: 147)

In other words, the realistic, insignificant detail in a work of art could be said to mimic reality by its conventional play of insignificance, by its very avoidance of meaning in the narrative scheme of the work. By circumventing metaphorical connections, we come close to the real.

Barthes goes further in *S/Z*, challenging culture's infatuation with realistic representation, by deconstructing Aristotelian conventions of the 'real'. Showing the deceptive quality of assuming the conventions of reality to be the 'real', he points to the insincerity of realism. Reading culture as comprised of an interlinked set of codes, he criticises realism for trying to tie itself to just one referent:

the (realistic) discourse adheres mythically to an expressive function: it pretends to believe in the prior existence of a referent (a reality) that it must register, copy, communicate ... (Barthes 1974: 465)

Barthes' conception of the various codes at play in a work of art – semantic, symbolic, hermeneutic etc. – comes close to the idea of ontological frames, which will be introduced later in the thesis. The reason this is relevant to a discussion of the perception of multimedia work is that it highlights the play of conventions or forms in a work of art which creates distinct viewpoints. 'Reality', however it is approached or even observed, might be one of these, but its mediation by the other codes, and negotiation or unravelling (as Barthes might say) by the viewer/listener, is what will essentially generate meaning: *a* meaning not *the* meaning.

1.2 Cognitive Immersion

Various concepts of immersion have comparable resonances with the ideas of Platonic mimesis, particularly the metaphor of the cave and the way reality is mirrored or said to be replaced by a 'hermetically sealed space of illusion' (Grau 2003: 5). If we were to suggest that diegesis, through the engendering of narrative viewpoints, creates critical distance, then we could also suggest that its opposite, mimesis, embeds the viewer, emotionally and sensorially, at the heart of the simulated space. The observed relation between figure and ground in a narrative situation, is what generates meaning. In immersion, the observer becomes the figure, and the distance becomes harder to observe.

The practice of immersive art has been around for centuries, if not millennia, if we consider cave art and the supposed rituals around it.¹⁶ From panoramic fresco paintings to virtual reality headsets, the sometimes less respected, arguably more populist tendency in art history has been to create illusion so bewitching that its strategy can be said to consist of replacing one reality with another. It has been shown under experimental conditions that increasing the strength of immersion – showing a film in virtual reality as opposed to 3D or a normal cinema environment – markedly increases the emotional response in a viewer (Visch, Tan & Molenaar 2010: 1439).

In much current new media art practice, there is a partiality towards creating compelling sensory spaces in their deployment of sound and visuals, which tend to be large in scale with strong mutual coherence between media. These works, it can be argued, reduce the critical distance of the artwork in favour of immersion. Grau, in his comparative historical analysis of immersion, suggest an interesting middle ground:

Obviously, there is not a simple relationship of "either-or" between critical distance and immersion; the relations are multifaceted, closely intertwined, dialectical, in part contradictory, and certainly highly dependent on the disposition of the observer. Immersion can be an intellectually stimulating process; however, in the present as in the past, in most cases immersion is mentally absorbing and a process, a change, a passage from one mental state to another. It is characterised by diminishing critical distance to what is shown and increasing emotional involvement in what is happening. (Grau 2003: 13)

In her essay "Immersed in Reflection" (2015), art historian Katja Kwastek develops the idea that immersion does not necessarily have to exclude critical distance. She specifically focuses on interactive art, which by nature involves becoming aware, sometimes overtly so, of one's own actions and emotional response within an immersive environment. The discrepancy between experience and contemplation at the heart of most forms of artistic expression is more acute when dealing with interactive art, where the audience is active in not only a cognitive sense but a physical sense also. Perhaps it is no coincidence that our understanding of the 'immersive' in art has developed from a simplistic sense of the illusion of the real, through the construction of hyperrealities, to art forms engaging the audience as an integral factor in the work. This is manifest in much interactive media art, interactive performances based on relational aesthetics¹⁷ and immersive theatre, where the audience is at the very centre of the action.

¹⁶ Writer Georges Bataille's study of Lascaux includes intriguing ideas about the immersive nature prehistoric rituals: Georges Bataille. 1979. *Oeuvres Complètes: Lascaux: La Naissance de l'Art*.

¹⁷ A movement in art defined by art critic Nicolas Bourriaud (2007).

Some art or entertainment forms have a greater tendency towards immersive experience: virtual reality media, 3D or IMAX cinema and first person video games. One could even describe different levels of immersion, according to the level of physical and/or cognitive involvement, or what kind of emotional experience is being engendered. What kind of immersion is relevant when dealing with music-text-film, where the text engages the audience on a cognitive level that is sometimes in contest with the auditive elements of music or sound? The text can be said to be the cause of a certain critical distance by creating a level of narration towards the music, a fixed perspective, though at the same time the inner voicing of the words synchronised with the music, places the audience directly inside that very narration. Kwastek's term 'cognitive immersion' is useful in describing this particular dichotomy:

This tension is not restricted to the realm of interactive art but is accentuated here by the merging of action and experience. In this category of art, not only the relationship between aesthetic experience and knowledge but also that between aesthetic experience and action must be reconceived. The embodied action of the participant is indispensable for the fulfilment of the artistic concept, which is intended to be experienced and reflected upon while being unfolded. (Kwastek 2015: 71)

There is something akin to an in-between space created in immersive art, where the body and mind are in two places simultaneously. One could even say that in music-text-film the ear is in three places; first, in the concert hall with its sonic architecture and all the intruding audience sounds accompanying it, second, in the diegetic space constructed by the music, and third, in the space of one's own voice, the resonance of the words subvocalized in our minds. In the essay "Neither Here nor There: The Paradoxes of Immersion" (Liptay & Dogramaci 2015) Fabienne Liptay references Barthes' fascination with the experience of doubleness when at the cinema:

by the image and by its surroundings - as if I had two bodies at the same time: a narcissistic body which gazes, lost, into the engulfing mirror, and a perverse body, ready to fetishize not the image but precisely what exceeds it: the texture of the sound, the hall, the darkness, the obscure mass of the other bodies, the rays of light entering the theatre, leaving the hall; in short, in order to distance, in order to "take-off", I complicate a "relation" by a "situation". What I use to distance myself from the image - that, ultimately, is what fascinates me: I am hypnotised by a distance; and this distance is not critical (intellectual), it is, one might say, an amorous distance... (Barthes 1995: 421)

As well as suggesting the impossibility of total immersion, because there will always be an anchor in the real world, what is interesting to highlight about Barthes' reflection on his cinematic experience, is that the duality of critical distance and immersion might in fact be quite a normal occurrence. To feel inside a diegetic space,

a constructed world, and at the same time to be looking at it from outside, is not a contradiction. But do these experiences happen simultaneously or does the mind oscillate between these states? It is probably quite difficult to know for sure the exact movement or overlap of these states, though we have all experienced at some point the feeling of being totally absorbed in a music or theatre performance, only to be taken out of this state of 'flow'¹⁸ by our own thoughts reflecting on the quality of the performance, or from reflection on the form of the composition or dramaturgy.

In experiencing music-text-film, I would suggest, there is a constant state of micro-fluctuation between cognitive processing and immersion in the imagined diegetic space, conjured by the text and supported by the sound. The strength of the absorption or detachment varies according to the nature of the imagined diegesis, the power of the sound world to draw the listener in, and of course the subjective experience of each audience member. The dynamic shifts between these cognitive states are, in my view, part of the excitement of multimedia work, and this relates to ideas of narration, which will be discussed later. I describe this state of immersion as 'cognitive immersion', because the immersion is perhaps never fully physical, as in interactive art, video games or virtual reality, but there is an engagement on the cognitive level, inviting the audience to participate with their inner voice; to place their voice at the centre of the artwork. While the audience is not necessarily physically present in the space of the work, their imagined voices are.

There are two issues related to the immersion experience which I would like to unravel in the following sections: how music itself creates an embodied and absorbing experience on the level of mimesis, and what exactly happens on a mental level when we subvocalise.

1.3 Vocal Embodiment

One of the ways in which immersion manifests itself on the musical level, can be explained through the concept of 'embodiment', a term that initially appears in the writings of Edmund Husserl. The philosophical branch of phenomenology, initiated by Husserl and later taken up by Martin Heidegger and Maurice Merleau-Ponty amongst others, places the body at the centre of perceptual experience; put another way, 'embodied aspects of experience permeate perception' (Gallagher 2014: 10).

The more recent philosophical branch of 'embodied cognition', as argued in the work of Mark Johnson (1987), suggests that many cognitive processes stem from bodily

¹⁸ Mihály Csíkszentmihályi's term for the mental state of deep absorption in an activity in *Flow: The Psychology of Optimal Experience* (1990).

experience. A way of understanding the world through physical senses initially acquired in childhood is transferred to abstract thought in later life. These take the form of 'image schemata', structures of physical interaction, that become established as patterns of cognition. Some examples are:

Containment, Path, Source-Path-Goal, Blockage, Centre-Periphery, Cycle, Compulsion, Attraction, Link, Balance, Contact, Surface, Full-Empty, Merging, Matching, Near-Far, Part-Whole, Superimposition, Process, Collection. (Johnson 1987: 126)

The act of the imagination, which uses these 'image schemata' to generate meaning, is also useful in understanding our experience of music. In a later work, *The Meaning of the Body* (2007), Johnson argues that it is these embodied schemata, rather than any sense of language, that create meaning in music and artworks (Johnson 2007: 208). He goes on to suggest some metaphors crucial to our understanding of music, namely those of music as movement, music as landscape and music as moving force. He, like many who advocate the importance of the body in cognition, underlines subjective experience as crucial in the formation of abstract ideas. As Sean Gallagher explains:

Sense of ownership is directly tied to the phenomenological idea of pre-reflective self-awareness, i.e. when we consciously think, or perceive, or act, we are pre-reflectively aware that we are doing so, and this pre-reflective awareness is something built into experience itself, part of the concurrent structure of any conscious process. (Gallagher 2014: 13)

According to phenomenologists, a 'sense of ownership' as well as its related 'sense of agency' are vital aspects of the conscious experience. The distinction between the sense of ownership and agency of an experience, is best illustrated by the example of an involuntary movement: if I am pushed to the ground by the hand of a random stranger, I might not be responsible for my movements, but I will still 'own' the experience (Gallagher 2014: 14).

Agency and ownership are interesting terms when they are applied to a musical situation. Taking the example of dance music: when we hear music that compels us to move our bodies, who has agency? Certainly the music could be said to be the cause of our compulsion to hit the dance floor, but it also transfers a sense of agency to us to move our bodies. We are the agents of our movement as well as giving us a sense of ownership. But do we feel ownership of the music? And how is that sense of agency transferred from the music to our bodies?

An interesting explanation of this experience is put forward by Cox in his essay: "Embodying Music: Principles of the Mimetic Hypothesis" (2001). In this paper he

sets out eighteen principles by which music constructs meaning through the induction of a sense of physical empathy through bodily motor imagery. Cox's use of the word 'mimetic' is in some way related to, but also different to, classical Platonic mimesis, as set out earlier in this thesis. He wants to differentiate between the objective idea of "art imitating life" and "the perceptual and cognitive processes, whereby music gets into the flesh, blood, and minds of listeners"(Cox 2001: 6). In general, most of the principles he outlines involve some kind of mirroring of perceived musical gestures, understood primarily as physical gestures. When we watch a drummer performing, he argues, from his study of neurological literature, we overtly or covertly imitate their movement as a way of embodying the music is being generated:

When we overtly imitate someone or something, we represent the observed behaviour in our own skeletal-motor system and in associated neural activity and blood chemistry. When we covertly imitate someone or something, we represent the observed behaviour in roughly the same way, except that the executions of the motor actions are inhibited, and the changes in other systems are attenuated. (Cox 2001: 19)

That is to say that there is also a sense of agency at play, whether we choose to physically act upon the 'mimetic motor imagery' or whether we choose to inhibit those movements. According to Cox, one can say that the sense of ownership or embodiment of this mimetic instinct happens regardless of whether we act upon it or not.¹⁹ He cites three variables of mimetic comprehension – how intentional, how conscious, and how overt it is – and adds that in adults it is mostly unintentional, unconscious, and covert, and that these variables are often shaped not only by individual but also by the cultural context (Cox 2001: 31).

How this is relevant to our understanding of the process by which the audience feels immersed or embedded in the context of music-text-film lies in the examples Cox gives of cross-modal imitation in instrumental music, one of the three modalities where he sees motor imagery occurring. The three modalities he cites are: intra-modal (finger imitation of finger movement), cross-modal (subvocal imitation of instrumental sounds) and amodal (abdominal imitation of the exertion dynamic that is evident in sounds) (Cox 2001: 38). Specific for cross-modal imitation is imitation occurring between different sets of motor actions, such as singing a melody heard on a violin, or in a covert sense, mimicking the melody with our inner voice. This subvocalisation of a melodic impulse is often not just happening on a purely

¹⁹ The 'globus pallidus' is responsible for inhibiting activation of motor activity. In cases of 'echopraxia', damage to the frontal lobe can result in patients compulsively imitating actions in their environment (Cox 2001: 20).

cognitive level, but involves the transmission of signals to the vocal cords, where the impulse to sing is inhibited or not.²⁰

This idea of the inner voice tracing heard melodic contours during the act of listening is fascinating, but it is hard to find clear empirical evidence. Cox cites a non-scientific survey conducted on a group asked to recall a big theme from a famous orchestral work. The act of recall was in most cases (90%) accompanied by some form of conscious subvocalization of the melody (Cox 2001: 42). There are, though, other suggestive examples, such as the role of the conductor. A conductor shapes the contours of the music, not only for the sake of communicating expressive information to the musicians, but also in translating the complexity of the score in a physical, mimetic sense that can be more directly experienced, on an intra-modal level, by the audience.

Cox goes further in describing how different kinds of music 'invite' different kinds of mimetic engagement. This notion of the 'invitation' is an interesting metaphor for the kind of code that a composer communicates to the audience at the outset of a composition. How is the piece to be listened to? He suggests:

composers design and shape the mimetic invitation, intentionally or not, and they can compose music that amplifies or attenuates intra-modal, cross-modal, and/or amodal mimetic engagement. Whatever the intention may be, music that attenuates the mimetic invitation is more likely to motivate descriptions of the music as "cerebral" and/or "academic," fairly or not, and we can understand this in relation to attenuated mimetic participation. (Cox 2001: 53)

An 'attenuated' mimetic participation gives a different kind of listening pleasure, because it tends towards a third-person narration rather than an immersive first-person one. This is similar to the immersion versus critical distance positions of the previous section. Fully mimetically engaged listeners put themselves in an immersed first-person perspective, embodying the music. A listener who is not mimetically engaged (either because the music itself does not communicate clear mimetic motor imagery or because the listener is focused on other aspects of the composition) could be said to have a third-person perspective, a more objective and critical distance to the music.

It is possible, like it is in literature, to experience shifts in perspective during the course of a piece, just as it is impossible to be totally immersed or totally detached throughout. Mimetic embodiment of musical gestures is a powerful process, conscious and subconscious, that can bring the listener to some kind of state of

²⁰ Evidence of electrical signals sent to the tongue, lips, or vocal cords can be detected when subvocalizing (Parnin 2011).

immersion. The shifting of perspective through the juxtaposition of different layers of musical or multimedia discourse is a strategy for creating a more dynamic relation between the poles of immersion and critical distance.

Another theory which is useful in this discussion, and is closely related to Cox's "Mimetic Hypothesis", is Gibson's theory of 'affordances': "The *affordances* of the environment are what it *offers* the animal, what it *provides or furnishes*, either for good or ill" (Gibson 1977: 127). This theory has been applied in many areas, including in music theory by Eric Clarke, Luke Windsor, Mark Reybrouck, Ruben Lopez-Cano and others. The theory of affordance is sometimes used to discuss what part the body plays in the perception of music. Lopez-Cano makes a typology of music affordances, which is not dissimilar to Cox's, dividing it into two main groups: 'manifest motor activity', which includes dancing or mimicking of postures and playing of instruments, and 'covered motor activity', which concerns the role the imagination plays in projecting an 'ideomotor' sense of physicality. However, the theory is useful as a way of examining music not just in terms of what it is, but to ask the question: what does music 'afford' the listener?:

Each listener finds in each piece of music, or in each different style, certain affordances and not others. This gives rise to a number of queries. Are all affordances heard by a listener in a given piece of music the same as those heard by other listeners? (Lopez-Cano 2006: 8)

This approach to discussing the potential inherent in a musical experience, what it affords, rather than talking about a specific subjective experience of it (phenomenological), provides a less categorical assumption as to how a piece of music is experienced, what definitive effect it has on the listener - specifically when discussing multimedia work, which one assumes would have many more possible frames of perception. This makes it possible to accept very different experiences music engenders as simply part of the subset of perceptions of actions which music 'affords'. This is especially relevant when the controversial subject of musical embodiment is at stake.

Thus, the question to what is specifically occurring in music-text-film in relation to the relative embodiment of the music and text, can best be phrased: what kind of activation do the music-text-film pieces 'afford' the spectator? At least one of these affordances could be said to rest on the notion of the imagined voice; specifically the activation of the vocal apparatus. I argue that there are two specific mimetic processes that can be said to be occurring in the vocal domain. On the first level, a musical one, the voice traces the melodic and gestural contours of the music as described above (Cox 2001: 42). It is subvocalising the music. At the same time, on a second level, because text is being silently read, the voice also subvocalises the words. Together, an imagined voice is created on a third level, in the combination of

the heard sound of the music and the projected sound of the inner voice. These three levels of engagement are sometimes in harmony, sometimes in competition, depending on the way the sounds of the words converge or diverge with the music being heard in the moment. This results in the varying levels of immersion or critical distance, which could be labelled 'cognitive immersion', which are indeed comprised of the aforementioned three levels of subvocalisation.

1.4 Subvocalisation

To read then might be also to hear what lies somewhere between the words, inside the white blanks, or over and around the languages that were once scratched onto paper, as an emotional energy. (LaBelle 2010: 108)

One of the most evident forms of cross-modal imitation afforded by music-text-film, that manifests itself in subvocalisation, is what is known as 'silent reading'. When we read silently we often project an inner voice speaking the words; more so, it seems, when reading dialogue (Ferryhough 2016: 1272). This is often modulated to mimic the particular voice represented by the text. The fact written words produce a private auditory experience, is evidence of a cross-modal or cross-sensory imitation.

Before people learned to read in silence, it was assumed that most people read out loud. There are contesting theories as to whether the Greeks and the Romans vocalised or not when they read. In writer Alberto Manguel's *A History of Reading*, a famous moment is recalled when St. Augustine encounters Ambrose reading. St. Augustine's surprise at Ambrose's ability to read silently is taken as proof that this was not such a common occurrence:

his eyes scanned the page and his heart sought out this meaning, but his voice was silent and his tongue was still. Anyone could approach him freely and guests were not commonly announced, so that often, when we came to visit him, we found him reading like this in silence, for he never read aloud. (Manguel 1996: 42)

Friedrich Nietzsche, in *Beyond Good and Evil*, also hints at the idea that in ancient times reading was rarely silent. He laments the fact that reading developed away from the voice and only to the eye (as he puts it). He implies that prose becomes poorer when the inflections of voice are removed from the text, pointing to the fact that even musicians, who would be assumed to have a greater inclination to hear the text as voice, are also culpable:

How little the German style has to do with tones and with ears is shown by the fact that it is precisely our good musicians who write poorly. Germans do not read aloud,

they do not read for the ear but only with the eye, keeping their ears in a drawer in the meantime. When ancient people read, if they read at all (it happened seldom enough), it was aloud to themselves, and moreover in a loud voice. People were surprised by someone reading quietly, and secretly wondered why. In a loud voice: that means with all the swells, inflections, sudden changes in tone, and shifts in tempo that the ancient, *public* world took pleasure in. (Nietzsche 2002: 139)

Recent scholarship has suggested silent reading in ancient times was perhaps more common than previously thought. Citing passages in Aristophanes and Euripides, classicist Bernard Knox highlights certain situations in which characters reading secret letters, unvoiced to the chorus and audience, set in motion ensuing dramatic consequences (Knox 1968: 433).

Eye movement in reading is also a crucial factor in subvocalisation. There are usually four distinct eye movements made when reading at an average speed: the 'saccade', the jerky eye movement coined by French ophthalmologist Louis Émile Javal; 'fixation', the stops between the 'saccades' (lasting on average 250 ms for a mature reader); 'regressions' (right-to-left movement); and 'return sweeps' (Vitu 2011: 732). The subvocalisation of words occurs during a 'fixation', and the slower the reader the more fixations there are. Fast readers tend to have less 'fixations' per line, and are able to move more fluently across the page. Still, even in these cases, words here and there are vocalised, and it could be argued that a certain type of text information, certainly dialogue, would have to be vocalised to make any sense. Author Charles Fernyhough, in his book *The Voices Within*, quotes psychologist Edmund B. Huey:

although there is an occasional reader in whom inner speech is not very noticeable, and although it is a foreshortened and incomplete speech in most of us, yet it is perfectly certain that the inner hearing or pronouncing, or both, of what is read, is a constituent part of the reading of by far the most of people.... And while this inner speech is but an abbreviated and reduced form of the speech of everyday life, a shadowy copy as it were, it nevertheless retains the essential characteristics of the original. (Fernyhough 2016: 1184)

Another interesting facet of subvocalisation of written text very relevant to the music-text-film is the phenomenon of projecting either the hypothetical voice of the protagonist, or of the author (if the author's voice is known to the reader). In several surveys and experiments, 80% of people reported hearing a voice of some kind when they read?²¹ The most potent form of voice hearing results from the author's use of what is known as 'direct speech': dialogue in quotation marks. Bo Yao and colleagues at Glasgow University were able to locate specific activation in the brain when subjects were asked to read direct speech. There seemed to be greater activation in

²¹ Guardian newspaper poll with Charles Fernyhough and another survey by Ruwanne Vilhauer of Felician Collegem New Jersey (Fernyhough 2016: 1238).

the right auditory cortex, a part of the temporal lobe responsible for processing voices. This was compared to the findings with subjects reading 'reported speech', when someone explains the words of someone else, and which doesn't require the reader to imagine a voice, rather just to process its meaning (Ferryhough 2016: 1272).

What is interesting here in relation to music-text-film is the question: to what level the text being projected alongside a musical phrase acquires the voice of the music, rather than the voice of the protagonist? It is clear in most cases shown in the later examples (in Chapters 4 and 5), that the projected texts are not a 'translation' of another voice – one we hear, as is the case of surtitling in opera – nor is the music representing in sound the phonetic contours of the voice, as in some of the work of Peter Ablinger²², for example, where the projected text reinforces what is being heard. Instead, in my music-text-film there is a dynamic between leaving space for the voice of the text to be vocalised by the audience, and modulating it in various ways through the rhythm, pitch and timbre suggested by the music. In one sense, the voice, formed by the combination afforded by the music and projected text, becomes embedded in the audience's lips, as each person projects onto it their own inner voice. The absence of a spoken or sung voice that correlates to what is being projected on-screen is one of the cardinal features of my definition of these music-text-films. This concerns not only the idea of a redundancy of medium – a performative voice being doubled with our own voices reading the text – but also the idea of absence, and the invitation to the viewer/listener to find a surrogate voice within the framework of the music. I will discuss this in more detail and on a case by case basis in Chapter 4 and onwards.

Is there any evidence of subvocal activity being a physical rather than cognitive phenomenon? From 2004 until their funding was terminated, a team lead by Chuck Jorgensen at NASA conducted research on precisely that, the physicality of subvocal communication. By placing sensors on the throat muscles, to measure electrical nerve signals – a technique called electromyography – and by using pattern recognition software, they were able to detect words that were unvoiced. This confirms the theory that just by thinking a word, a signal is sent to the vocal cords to potentially voice it. But because not all vocal sounds are generated in the throat (the mouth is responsible for much vocal nuance), only a limited vocabulary was recognisable:

So there was some preliminary work done on that, and the answer was: Yes, we can pick up some of those vowels and consonants, but not all of them, because not everything that you're doing with the muscles reflects what goes on with speech. An example of that would be what they call aplosives, which are the popping type of sounds that you make by closing your lips and pressurising your mouth (Peter, Paul,

²² As in Peter Ablinger's *Letter From Schoenberg*, from the series *Quadraturen* 3.

Pickled Peppers, etc.). Those types of aplosive noises are not represented. We did some work also at Carnegie Mellon connecting it to a classical speech recognition engine, except the front end of it was now a subvocal pickup. I believe that work got up into the 100s to possibly 1000-2000 word capability. That was probably the most advanced work using that specific approach to subvocal speech. (Jorgensen 2013)

The detection of electrical impulses in the throat muscles was not just limited to deliberate or conscious subvocalisation. As far back as the 1940's, psychiatrist Louis Gould conducted experiments on schizophrenic patients who suffered from auditory hallucination, and found that when patients reported hearing voices, their electromyographic recordings showed greater muscle activation –their vocal muscles were contracting. With some patients he could even detect the imagined voice as an almost imperceptible whisper when a microphone was placed at the throat (Sternberg 2015: 153).

The fact that modern writing systems are largely a graphical encoding of verbal communication, of spoken phonemes, hints at the idea that the voice is pivotal not only in the communication of language but in its perception and comprehension. As children we are encouraged to develop both the 'reciting' voice that subvocalizes the text and the 'thinking' voice that is in conversation with it, that probes it for understanding. This secondary voice is what is known as 'inner speech'; this often appears in dialogue between competing voices and point of views in our minds. Many philosophers have pointed to this activity as an essential cognitive process. Philosopher Charles Peirce names these parts of the self the 'critical self' and 'present self': aspects of the ego which are negotiating different parts of time, past, present and future (Archer 2003: 71). According to psychologist George Herbert Mead, inner speech arises out of a dialogue between a 'socially-constructed' self and an 'internalised other', which adopts different attitudes towards what the self is doing (Ferryhough 2016: 563).

So far we have sketched two instances where the inner voice is said to be activated by sending electrical signals to the throat: in following musical phrases and in reading text. In the next section, a third, more elusive inner speech, that interacts with processes of memory and cognition, is analysed.

1.5 Inner Speech

The inner voice is never a single voice though. Rather, it appears through a variety of registers, in a variety of volumes, at times only a soft murmur while at others as a full articulation of words. (LaBelle 2014: 87)

When we think, words tend not to be very conspicuous. Sometimes they come to the surface because we find ourselves voicing an inner dialogue or repeating a phrase in

order to make sense of it, but mostly words stay hidden. According to philosopher Don Ihde, this is because they are being obstructed by the object they are referring to:

Words do not draw attention to themselves but to the intended things in referring. This extends ordinarily even to the form of embodiment in which the language is found. Thus in speaking, what is ordinarily focal is "what I am talking about" rather than the singing of the speech as a textured auditory appearance. (Ihde 2007: 138)

Ihde acknowledges that the inner voice is present, but mostly in the background, or 'nowhere' and 'everywhere' at the same time. The sense of ownership of this voice, when fleetingly glimpsed, is reinforced by the feeling that it is coming from our own body or mind; that it is indeed ourselves thinking. But how does one actually know it is one's own voice and not a voice coming from elsewhere? According to neurologist Eliezer Sternberg, our minds compare the sound of the inner voice with what we expect to hear of our outer voice, and if the voice fits with the prediction of what we should be hearing, the mind affirms the ownership. This is different in the cases of people who hear voices as if they are someone else.²³ The brain does not recognise the inner voice as coming from the hearer, but from elsewhere:

the unconscious matching system incorrectly identifies a mismatch (false negative) and prevents (the hearer) from consciously recognising that it is his own speech that he's experiencing. His brain is left to reconcile two seemingly contradictory pieces of information: on the one hand he hears a voice that isn't his own. On the other hand, there's nobody else in the room. (Sternberg 2015)

In pathological cases, the mind might conclude that since there is no one else in the room, the voice is coming from an invisible force: a deity, a secret 'controller'. Since the brain needs to construct narratives to explain our inner and outer realities, it creates the idea that the mind has been infiltrated by an outside power. According to psychologist Julian Jaynes, in his controversial yet influential book, *The Origin of Consciousness in the Breakdown of the Bicameral Mind* (1976), this was the normal state of consciousness in the mind of our ancestors up until about 3000 years ago. Jaynes' theory of 'bicameralism' describes the state in which experience and memory in one part of the brain are transmitted to another through auditory hallucination:

Consider the evolutionary problem: billions of nerve cells processing complex experience on one side and needing to send the results over to the other through the much smaller commissures. Some code would have to be used, some way of reducing very complicated processing into a form that could be transmitted through

²³ This is no longer a phenomenon categorised as pathological. Through the work of the 'Hearing Voices' movement (<http://www.hearing-voices.org/>) there is greater acceptance of the mainstream occurrence of voice hearing not linked to mental illness.

the fewer neurons particularly of the anterior commissures. And what better code has ever appeared in the evolution of animal nervous systems than human language? (Jaynes 1976: 105)

According to Jaynes, because the bicameral mind lacks a meta-consciousness or the ability to consciously retrospect, these inner voices would have manifested themselves as supernatural voices and be heard as emanating from outside oneself: a voice of a god giving advice or commands. This psychological state gradually gave way to the evolution of consciousness, partly through the use of metaphorical language (Jaynes 1976: 138), so that the imagined voice of deities became simply the voice of our ego. He cites this seismic shift in consciousness at exactly the juncture between Homer's *Iliad* and *Odyssey*. According to Jaynes, the heroes in the *Iliad* act exclusively at the behest of their Olympian masters, the cause and effect of human action being simply shadows and traces of their god's impulses. In contrast, Odysseus seems to be the first hero to act in part out of his own volition:

[The Odyssey] is a journey of deviousness. It is the very discovery of guile, its invention and celebration. It sings of indirections and disguises and subterfuges, transformations and recognitions, drugs and forgetfulness, of people in other people's places, of stories within stories, and men within men. The contrast with the *Iliad* is astonishing. Both in word and deed and character, the *Odyssey* describes a new and different world inhabited by new and different beings. The bicameral gods of the *Iliad*, in crossing over to the *Odyssey*, have become defensive and feeble.... The initiatives move from them, even against them, toward the work of the more conscious human characters. (Jaynes 1976: 273)

Jaynes' description of imaginary voices being used as an intermediary between different spheres of mental activity, linking the left and right hemispheres of the brain, has received critical response from the scientific community (FERNYHOUGH 2016:2069), in part because of the overtly simplistic allocation of functions to different parts of the brain. In spite of this criticism, it remains a powerful metaphor for how inner dialogue acts as a vital aspect of cognition, and how 'heard voices' could somehow be an evolutionary step to thinking with a silent inner voice. Charles Fernyhough argues that it is not inconceivable to trace a link between the development of inner dialogue in children, first formed in conversations with parents or carers, developed in vocalised conversations with imaginary others, and finally becoming internalised, or as he puts it: "going underground", and one of the basic vehicles of our cognitive process (FERNYHOUGH 2016: 247).

I do not wish to delve too deeply into the scientific developments of research into both voice hearing and inner voice, partly because it is far outside my field of expertise, and partly because it seems to be an incredibly difficult subject to explore in both empirical and neurological experimental studies. One of the most successful methods used to highlight the function of the inner voice seems to be Descriptive

Experience Sampling, as developed by psychologist Russell Hurlburt amongst others at the University of Nevada (Hurlburt & Akhter 2006). The method asks subjects to jot down their exact thoughts in everyday life at the sounding of a random beeper attached to their clothes. The difficulty of separating inner voice from general thought processes, images and memory shows how entangled our daily cognitive experience actually is. Despite these difficulties, the research to date, as outlined above, is helpful in articulating a sense of the multiple voices that are present when experiencing music-text-film work. Even the very complexity of locating the mechanism of these voices, and the difficulty of establishing exactly what happens, makes it particularly fertile territory for artistic exploration.

1.6 Silent Voices

Considering the notion of 'mimesis', as it appears in Plato's *Republic* and Aristotle's *Poetics*, is useful in helping to understand the way the voice is engaged in listening and reading my music-text-film. Rather than using it in the classic form, describing art copying the world, nature and the real, I use it as a way of explaining how an artwork communicates to the audience, transforming the spectator's reality itself into a 'mimesis' of the artwork. Taking the statement encountered earlier in the chapter: "art imitates nature", I would turn it into: "nature imitates art", and state that one of the ways of engaging with art, is simply to 'become' it. This follows into the discussion of ways in which the spectator can be part of an artwork, through the presentation of theories of immersion and specifically the polarity of immersion versus critical distance, which is so relevant in my music-text-films. The idea of 'cognitive immersion' is used to define a state of listening that music-text-film affords, where there is not necessarily full immersion in the artwork but an engagement on the embodied cognitive level, inviting the audience to participate with their inner voices.

The theory of affordance (Gibson 1977), helps me to ask the questions: "What does this music afford the listener?", "What possible actions does it enable?" Rather than definitely stating how these pieces should or are listened to, I discuss several theories of how the body 'could' respond to such stimuli of projected text with music. Under the general heading of the 'mimetic', I have shown how the inner voice of the spectator could follow both the melodic or gestural contours of the music in a process of mimetic embodiment, this is what I call '**silent singing**'. Secondly, the more well known phenomenon of '**silent reading**', subvocalises the words, as they are read in time with the music. Finally, the third, less conspicuous inner voice, which is trickier to distinguish, and which varies a great deal from person to person, is the elusive inner voice of thought, that I call '**silent discourse**'; the voice that is in

constant dialogue with different aspects of the self, ever active in the process of trying to comprehend and respond to situations on the conscious horizon.

In the next chapter I take the other branch of Plato's dualism of poetry, 'diegesis', to formulate an idea of 'narrative voice' which can be seen as a projection back into the music of a perspective of listening, a mirror of our own voices.

Chapter 2: *The Diegetic Voice*

The concept of narrative voice has been widely discussed in narrative theory throughout the last century, with the main focus being literature. Gérard Genette, one of the principal thinkers in the field of narratology, takes the Aristotelian concept of diegesis and develops it to ask the precise question about the location of voice (Genette 1988). In literature, when we read words on a page, whether the narration is 'intradiegetic', coming from a character in the story world, or 'extradiegetic', told from outside the narrated world, things are relatively straightforward. Even in music, where an actual voice is concerned, in song or in opera, one can easily locate the focal point of the narration. Matters begin to get complicated when the idea of narration is removed from a discernible voice, when there is no clear embodiment of a storyteller, or when the idea of voice becomes split between different media.

In this part of the thesis I will be focusing on the concept of narration, posing the question of who is speaking in a musical discourse. In my view, the very absence of physical voice can underline an awareness of a narration and problematize the idea of what is being communicated. What I would like to highlight is how this narration can shed light on the background process of mental activity occurring in the mind of the listener, and how with the aid of the mirroring or the mimetic process set out in the first chapter, it can state something about our perception, reflection, and emotional response to a piece of music.

The chapter begins with an examination of classic diegesis, as it appears in the *Poetics* of Aristotle. This is the main theoretical underpinning of narratology and the first time the notion of a narrating voice is encountered historically. It is also interesting to note that Aristotle develops Plato's conceptions of the diegetic and the mimetic not as mutually exclusive concepts but as an entangled set. For Aristotle, mimesis is the root of all art and diegesis is the manner it is communicated. He divides this into the triad of medium, object and mode, the latter being highly relevant in the discussion of what constitutes narrative voice. This clearly resonates with one of the key concepts of narrative theory as developed by Genette, amongst others: the discussion of narrative's different ontological levels, which he defines as a triad of story, narrative and narration (Genette 1980). Like Aristotle, Genette's focus on the notion of narrative voice is born from his understanding of the act of narration. He develops this through the notion of what is termed 'focalization', sometimes referred to as 'point of view'. These two aspects of narrative, the 'who speaks?' and the 'who sees?' (though sometimes they are difficult to tell apart), are the main agents of narration. At its essence, the idea of focalization can be defined as 'how a subject envisages an object' (Hühn 2009: 79), but the difficulty in applying this

idea to a musical situation leads perhaps to a confusing question about what is the subject or object in music. In this sense, whilst taking the useful notion from Genette that some agency in the material of the work 'focalizes' the perspective from one thing to another, I prefer to avoid the word 'focalization' when talking about music or music multimedia, and use the more neutral term 'focus'.

How, then, does this apply to music? Can one make an assumption about the object that is put into 'focus' and from whose perspective we are looking when it comes to the level of abstraction found in a piece of music? In this chapter I sketch how levels of narration occur within music and, most importantly, between music and other media discernible in music's various contexts, with a specific focus on the multimedia framework of my music-text-film works. I argue that within the narration of certain musical works, a multitude of voices can be discerned, and that these voices can create a sense of difference of perspective within the narration of a work. I term this a 'narration network'. What interests me is how, when a listener moves between what are perceived as different levels of media, a change of perspective occurs, which in turn generates different meanings. Meaning that arises from observing one level from the point of view of another, influences how the narrative voices are perceived. How something is framed becomes a vital aspect of the composition, especially in the context of how text information can influence the way we listen. In literature, Genette's concept of the 'paratext', and Derrida's closely related idea of 'parergon' are an essential to understanding that narration is not hermetically sealed inside a work, but a continuous flow between the reader and the world.

In this chapter I propose how concepts of narrative voice and perspective could be understood in the context of music. This does not simply translate to the authorial or performative actions in the piece, the idea that the musician is the one who *speaks*, but that many narrative voices are perceived by the listener within the fabric of the music. These also occur between music's many different ontological levels, and these voices are constantly acting in the process of 'focalising' our point of view. I take this to be one of the main narrative dynamics occurring in my music-text-films. Just as the projected text focuses the musical narrative, the music focuses the narration in the text.

Beyond literature or more overtly narrative forms, it is sometimes difficult to find the distinction between narrative voice and perspective, especially when it comes to forms utilising more than one medium, as these theories were generally not designed for application in non-literary form. The process of trying to pin-point exactly what can be designated as 'voice', when there is no physical presence of voice and no overt narrative gesture, is often tricky. Furthermore, many multimedia works in the experimental domain, have at their core a fluctuating and unstable character in the position the viewer/reader/listener takes, so even if a voice has been perceived,

the chances are that it is not the only voice in operation. After all it is exactly this perceptual instability, which makes multimedia music work compelling as an art form.

The chapter ends with a reference to psychologist and a cognitive scientist Lawrence Barsalou and his idea of recursive attribute-value structures, known as 'frames'. I use this as a basis to formulate a model of how meaning could be generated between different perceived layers in music and in music multimedia, that the idea of narration does not necessarily have to be embodied in what is conventionally understood to be one voice, but can exist in a polyphony of different senses, engendered by changes of perspective.

2.1 Narration

Plato's ideas of the polarity of mimesis and diegesis (quoted in the beginning of the previous chapter) were expanded by Aristotle in the *Poetics*. Aristotle's interpretation brings further nuance to the idea of narration, by differentiating between 'medium' (in-what), 'object' (of what) and 'mode' (how). In contrast to Plato, Aristotle defines these all within the concept of the mimetic process; the idea of narration being embedded within an act of imitation.

'Medium', or what George Whalley, in his translation of the *Poetics* prefers to call 'matter' (Whalley 1997: 46), is defined by Aristotle as embodying rhythm, melody, and speech. This is an interesting point and will form a vital aspect of the following discussion of how we define a 'medium', and whether music might indeed already be more than just one medium. In Aristotle's definition of medium, we come very close to what we would understand as 'parameters' in music; defining music as composed of multiple parameters must, in a way, concede its multi-medial nature.

'Object' defines what the 'subject' being represented by the artist is (Whalley 1997: 48); what they are writing, acting, singing or painting about. This parameter of mimesis seems to be very clear when dealing with the classical arts, Renaissance painting or even Romantic programme music, because the subject that is being dealt with often originates in archetypal or mythical sources. How the 'subject' defined the medium and mode, or form and style to be used was very much a convention of the time. Mythological subjects tended to fall into the genre of tragedy or comedy, which in turn defined how they were structured. Following from this is the argument of whether genre itself defines the subject/object of the artwork or vice versa.

'Mode', or as Whalley prefers to call it 'Method', seems to be the most open and ambiguous of the differentiae, as it refers directly to the way in which the artist

communicates. This is where the differences between the mimetic and diegetic modes open up into vital distinctions between voices, time, and points of view at the core of modern narrative theory. Aristotle builds on Plato's definition:

For it is possible to deal with the same matter and using the same subject, but using different methods: (a) by narrating at times and then at times becoming somebody different, the way Homer composes, or by one and the same person speaking with no change in point of view or of method, or (b) by all the people (*mimoumenoi*) who are doing the mimesis taking part in the action and working on it. (Aristotle in Whalley 1997: 49)

Here Aristotle sets out his interpretation of Plato's definition of the mimetic and the diegetic, by implying that diegesis is a subset or a 'mode' of mimesis, which in his view is the overriding concept of all art. The difference between Plato and Aristotle rests fundamentally on this definition of mimesis. It is also of interest that Plato ascribes three possible modes: mimetic, diegetic and 'mixed' (Homer can sometimes impersonate other characters as well as speak with his own voice), whereas Aristotle groups both the diegetic voice and the 'mixed' voice in one. This can lead to complications, especially in the light of how an artwork is perceived, rather than conceived, and that is why Whalley, in his translation of the *Poetics*, insists on translating the Greek '*tropos*' as 'method' rather than 'mode', putting stress on the way the poet works rather than the manner in which the work is seen (Whalley 1997:54). The philosopher Paul Ricoeur comments in his seminal work *Time and Narrative*:

Platonic mimesis distances the work of art twice over from the ideal model which is its ultimate basis. Aristotle's mimesis has just a single space wherein it is unfolded - human making, the arts of composition. (Ricoeur 1983: 34)

Although mimesis is defined by Plato and Aristotle in different ways, in different scales of reach, or as Ricoeur suggests, in different ways of folding, their idea of diegesis is very close: mimesis embodies, and diegesis narrates. They even imply a very different time scale: mimesis is wedded to the continuous present, diegesis to the past. This becomes a subtle yet important mode of differentiation when dealing with how music, especially in more complex multimedia forms, sets out multiple temporal levels between the polyphony of its voices.

In literature, Genette makes an interesting distinction between mimesis of action and mimesis of language. Mimesis of action can evoke both landscape, characters and their activity, as a way of drawing the reader into the 'reality' of the narrative space of the novel. Mimesis of language, on the other hand, is where direct speech is used to represent the actual words uttered by the characters, bringing into fore the idea of voice (Fludernik 2009: 65). This is an interesting point with respect to the examples

given above in the discussion of mimesis, in which 'concrete' or found sound can function against synthetic sounds in a musical composition. Genette further distinguishes between three differing diegetic levels. Within his definition of the narrative mode of 'voice' he sets out to differentiate between: the 'extradiegetic' level, the voice of the narrator who is not a character in the story; the 'diegetic' level, which is understood as the level of the characters, their thoughts and actions; and the 'metadiegetic' or 'hypodiegetic' level, which in simple terms can be expressed as a story within a story, when, for example, the character in a novel narrates a story. It is important to note here that in any work of art featuring a diegetic form of narration (where a narrator is presenting a world that is not their own and that they are not a part of), there will always be two or more ontological levels: the level of the narrator and the level of the narrated. These levels are not independent of each other, but are inextricably linked, as musicologist Karol Berger writes:

While the primary focus of the reader's attention is usually on the latter mediated world, structurally the narrated world is hierarchically subordinated to the world of the narrator.... the former is embedded in, or dependant on, the latter, since it is the existence of the narrator's voice that makes the existence of the narrated world possible, and not the reverse. (Berger 1994: 412)

When dealing with the mimetic (dramatic) mode of narrative, the idea of voice is more clear-cut; the performers are speaking (or singing) through the voice of a character. In the diegetic mode, the voice can be the narrator, and might embody one or many voices. However, it is often entangled in the world being presented, jumping between the various levels of time and narrative. One can delve into further entanglement when dealing with multimedia forms of art, in that the Aristotelian differentiae of medium, subject and mode, start to lose their marked definitions. Media embodying different modes of narration can be overlaid, juxtaposed, merged or opposed to each other. The subject matter of a work can lose its objectivity and coherence, when seen through the filters of different media. One such example is how a form like a documentary film can sometimes say more about the maker or narrator rather than about the object of study. Going a step further, in multimedia forms of music, where for instance drama, video or text are mediated through musical voices, the blurring of these Aristotelian boundaries can be profound.

Where this definition of diegesis as the fictional reality might begin to unravel, is when dealing with a filmic universe that is not fictive, such as in a documentary. This point aside, the film medium can be defined as an 'epic' form using 'mimetic' elements. In a technical sense, it is the position of the camera and the timing of the editing that becomes our narration. Thus the narrative space of film, referred to as the 'diegetic' space, encompasses all aspects of the story, and anything happening outside this story space, such as titles and incidental music, are referred to as 'non-diegetic'.

These spaces are not hermitically sealed. In much contemporary work, elements can cross the diegetic divide. This is something that occurs in my own music-text-film, where voices spoken or sung live, or heard in the soundtrack, might cross over to the text or film layer. A term used in recent film theory to describe this is 'trans-diegetic'. A tune that is whistled by a character in the film and then taken up in the orchestral score, or the moment in Francis Ford Coppola's *Apocalypse Now*, when a radio playing the Rolling Stone's *Satisfaction* first located on screen, swells to be taken up in the off-screen space (Taylor 2007). It is worth noting that the term trans-diegetic, is in some ways related to the concept of 'metalepsis', as coined by Genette to describe the 'paradoxical contamination between the world of the telling and the world of the told' (Pier 2010). Both terms have in common the idea of jumping between narrative ontological levels, between the world of the storyteller and the world of the story. This is a useful concept to use when discussing multimedia music. However, in music, and arguably in much media that is not based on clear storytelling, the difference between the world of 'telling' and the world of 'told' is sometimes hard to define. This makes 'metalepsis' a much less explicit mode of narration with regards to music, as opposed, to other often cited examples like Luigi Pirandello's play *Six Characters in Search of an Author*, or Woody Allen's film *The Purple Rose of Cairo*, where the concepts of character and author are turned inside out, and the different levels of narration are mixed, ruptured and reconfigured.

2.2 Paratext

Overt narrative elements have long been part of music's expressive arsenal, not only in work using text vocalised in the music, but due to the array of words, images and rituals that surround music and play a part in constructing its meaning. From epic poetry to troubadour song, from word painting to programme music, from opera to incidental music; words have shaped the way we listen to sounds and coax meaning out of them. Of all forms, the song seems to have emerged as the most popular and enduring of musical conventions.

The most primary use of written text that is projected or published alongside the musical work itself is in the title and name of the author. This is what Genette defines as part of what makes up the 'paratext' of a work:

a zone between text and off-text, a zone not only of transition but also of transaction: a privileged place of pragmatics and a strategy, of an influence on the public, an influence that... is at the service of a better reception for the text and a more pertinent reading of it. (Genette 1997: 2)

The paratext, which in music encompasses title, programme note, and any other contextual material around a work, can have a huge effect on how the music is understood. I am interested in the ways in which all the different forms of words and images surrounding and contained within an artwork affect our understanding of it. In my own music-text-films the non-musical material can sometimes inhabit the sidelines and function much like paratexts, ancillary information that effects the way we experience the music.

Whether this information appears in media preceding the musical event – for example, on a poster outside the concert hall, in a program book carried into the hall, on a record or CD sleeve, or as metadata in a music download – it colours our perception and cognition of the codes of the music itself. Whether a programme note is of help or hindrance to the audience's own subjective interpretation of the music, is an interesting point of discussion. A recent paper suggests, that most audiences find programme notes to classical music not particularly helpful or enriching to the listening experience (Margulis 2010: 285), but with newly composed pieces they can offer a useful 'guide' and 'direction' to both the listener and the performer (Blom, Bennett & Stevenson 2016).

The naming of the author, followed by the understanding of one work through previous knowledge of the author's other work, is one primary pattern of cognition (Genette 1997: 37). Biographical information associated with a name has an affect on whether an audience can identify with a cultural icon represented, be it Bach or Boyzone, and this can influence how the music is contextualised in the mind. We have the tendency, when listening to a piece, to ascribe the music's narration to the author, so that the narrative voice is somehow traced back to an idea about the author. The creator's presence in a work of art could be said to fluctuate in different eras and cultures. In classical music this reached its height in the Romantic era, where the cult of the artist aligned the value of the work to the perceived value of its author.

The title of the work provides an additional frame, which can resonate throughout the course of listening to and thinking about the music. A metaphor, a form, an image, or a statement expressed in a title cannot easily be ignored, and its meaning might thus be constantly correlated to significant moments arising in the work itself, as a way of constructing meaning. This is most obvious in the statement of a form; the naming of a sonata or a symphony already sets us up how we listen to a piece of music. Images used in titles, like in much programme music, can imply a basic narrative which the music can represent. This happens in the form of suggestion. If we hear a river in Bedřich Smetana's *Moldau*, it is not solely because the music evokes it, but because the title does so as well. Similarly, a title of a concept album in pop music can unify the ideas or themes individual songs, whether directly expressed in the lyrics or not. Pink Floyd's 1979 release *The Wall* is a classic example

of a concept album in which an overriding narrative, suggested through the album's title, weaves together the different tracks.

Layers of paratext that envelop a work, the title, the identity of the creators, the programme notes, the lines or pages or books of commentary that have been written about it, all contribute to the construction of meaning in a work, and eventually lead to our own very personal impression of it. They play such an influential role in our understanding of narrative in music, to the extent that it is almost impossible to imagine a situation whereby one experiences a piece of music unframed by this kind of information. Paratextual material influences our idea of who is narrating and how the narrative is focused, constructing an ever-present frame through which we experience the work. This is comparable to Derrida's conception of the '*parergon*' –the frame of a work of art, both literally and figuratively –neither fully outside nor fully inside a work of art:

Neither work (ergon) nor outside the work (hors d'oeuvre), neither inside nor outside, neither above nor below, it disconcerts any opposition but does not remain indeterminate and it gives rise to the work. (Derrida 1987: 9)

To Derrida context is everything. It is not simply a portal one walks through to reach a work, but an ever present framework when experiencing a work. Following this, one can argue that words or images surrounding a piece of music, whether intentional or unintentional, are a critical element in the formation of narrative and perspective, and one cannot simply regard the music alone as a hermetically sealed space.

2.3 Narrational Network

The narrative created between words and music, between paratext, context and music, is where arguably most of the meaning of a musical work is generated. It is, nonetheless interesting to explore the other levels of a musical work where meaning could be said to exist. Many writers in the field of narratology split narrative into different ontological levels, in order to discuss the function of each part separately. One could question to what extent structural ideas applied to literature could easily translate to the discussion of music, but some useful parallels can be borrowed. Literary theorist Shlomith Rimmon-Kenan's influential book *Narrative Fiction* (1983) takes the spirit of Genette's 'histoire', 'récit' and 'narration', splitting the narrative into three aspects: 'story', 'text', and 'narration' (Rimmon-Kenan 1983: 3). According to Rimmon-Kenan's definitions, 'text' – what is read or heard – is the only aspect directly available to the reader or listener. 'Story', the world that is conjured by the text, has to be inferred by the reader, and the 'narration' is how the story is told.

Taking this three-fold structure and applying it to music, 'text' could be analogous to the material condition of music (notes, instrumentation, sounds, rhythm, form etc.), 'story' could be the 'work' evolving through it, and 'narration' the way this is communicated by the performer or performance event.²⁴

Can one argue that there is meaning generated within each of these levels, or is meaning still more distinct when the thresholds between the levels are crossed? As a way of understanding how this functions, it is useful to re-introduce the idea of 'voice', not in a literal sense, but as used in narratology, to underline the idea of an act of communication. One could say that in composed music there is a level of authorial narration, one ever-present voice, an extra-diegetic voice. Within the music itself, the Aristotelian mode of presentation is also represented by the concept of 'voice', which in this case could be defined as an 'intra-diegetic voice'. Even if there is no actual singing involved, a melodic line, a phrase or a sonic gesture then takes on the function of 'voice', becoming a narrative act. Examining the musical equivalent of the 'text' Rimmon-Kenan speaks of, one could say that 'voices' become apparent in constructing the flow of narrative. This is most obvious when something akin to melody or melodic gesture is present, but it is not exclusive to this situation. Even in the case of the most extreme drone music, the more we listen, the more we might discern figure, pattern or voice in the texture. This has been highlighted in some experiments on aural pattern recognition (Rogers & Pullum 2011) or in the phenomenon of apophenia, known in neurological studies as 'audio pareidolia': the perception of patterns in random noise. Extreme cases aside, one can say that 'voice' is detectable within the fabric of much music.

Generally speaking, in much music there seems to exist a relation between 'figure' and 'ground', melody and accompaniment. Even in cases where the boundary of melody and accompaniment is blurred – "where a melodic motif can join the accompaniment and the reverse, an accompaniment motif can be melodically thematised" (Berger 1994: 412) – the idea of figure and ground persists as a perceptual model of how music narrates. One could furthermore stretch this analogy to encompass the idea of 'narrative voice' in much contemporary music, even where this music substitutes the melodic for the gestural in its reconfiguration of the anthropomorphic subject. Electronic music composer Denis Smalley's concept of spectromorphology describes the sound-making process in terms of the physical gesture:

Sound-making gesture is concerned with human, physical activity which has spectromorphological consequences: a chain of activity links a cause to a source. (Smalley 1997: 111)

²⁴ This is not meant as a definitive translation into music of a model in the realm of literary fiction. What I want to highlight is how one can think about music's different ontological levels in a similar way.

This is to underline the fact that, in the fabric of the music, for 'voice' to be perceived there already exists a separation into figure and ground, as well as, arguably, the segmentation into phrases and perceivable units that occurs in many forms of language processing (Bates 1995). The difficulty arises in trying to distinguish whether the music is presenting itself (showing) or being presented (telling). Here is where the parallels between literature and music become less evident. Even if we concede that there is in music a narrative act of some kind, is it useful to speak of a narrator? And beyond text, context and paratext, what is the music itself narrating?

This point could be answered by referring to the idea of musical representation discussed by many musicologists, including philosopher Charles Nussbaum (2007), who sets out modes of musical meaning such as extra musical form, or extra musical content used as scenarios by the listener in the virtual musical space. Nonetheless, the idea of a single narrative voice remains harder to pinpoint. Does it rest on the level of the sounds, the notes, the performers, the author, or on the cultural forces that determine a musical language? Certainly, the various temporalities implied by these levels are ever present, and as discussed in the previous section, form an ever-present aspect of how meaning is generated for the listener. We understand something on the level of the notes because of an understanding of the relationship between the music or musician and their milieu. At the same time, the relationship between performer and score comprises yet another level of narration that crosses a temporal separation. But even within one perceived temporality, is it absolutely clear what or who is narrating?

In my opinion, it is difficult to talk of one narrator, or even one type of narrative act or narration, in a given work. There are many intertwined voices and stories within the fabric of a single music, and many ways in which they are communicated and perceived. Thus, it is perhaps more useful to talk of a 'narrational network', straddling the different ontological levels of how a musical work is perceived, from the level of the 'music', where a multitude of voices, a polyphonic entanglement of different perspectives constructs narrative, to the way in which these are communicated by the musician(s) and how they are related to the musical identity or the space being constructed.

In narratology the term 'focalization' is used to denote the different perspectives that a narrative is presented (Genette 1980). How does this apply to music? One way of examining this, is to assign the presentation of narrative only to the performer. Artist/researcher Vincent Meelberg underlines the differences between a narrator and the one who focalizes, the point of view from which the elements are perceived (Meelberg 2006: 66). He draws a parallel between focalization and performance. Whereas literature can have many internal 'focalizers', different characters who colour the reader's understanding of the narrative, according to Meelberg music only

has one: the performer (Meelberg 2006: 66). Although I appreciate the performer plays an important focalizing role (in a sense that works with Genette's definition), I do not think it is useful in the context of this thesis to stick to a rigid definition of the term. That is why I prefer to use the term 'focus', more neutral and general and less tied to Genette's very specific narratological definitions. Furthermore, because I see music as a network of many voices, that are not just represented by a particular musician or instrument in a score, I believe that, like in literature, also in music there are many 'focusing' elements creating multiple perspectives within the fabric of music and between its different ontological levels. The important element to keep in mind here is that, unlike in literature, where there is a production chain of author -> text -> reader; in notated music there are sometimes extra elements in the chain: composer -> score -> performer -> listener. This complicates the idea of Genette's 'focalization', because it occurs between score and performer, as well as between performer and listener. A similar case can be made in theatre, noted by literary theorist Roland Weidle:

The narratological concept of focalization as a filter through which the act of narrating takes place is problematic when applied to the analysis of drama. Because of drama's physical and visual nature and the material presence of the actors, focalization in drama, or to be more precise: in "reading drama", appears to be less dependent on the mediating process. Of course, even in drama, focalization does not take place without narration (by means of a superordinate narrative system), but the relation between narration and focalizing seems to be less prescriptive and more flexible than in narrative fiction. (Weidle 2009: 239)

The 'narrational network' together with multiple points-of-view in the fabric of the work, (in Rimmon-Kenan's level of 'text'), create a space where there are already multiple ontological levels. In general the splitting of the work into multiple ontological levels, which is the principal characteristic of the diegetic function, creates the space for the construction of narrative time. How this sense of narrative time manifests itself in a musical context, is a question which will be discussed in the next subsection.

2.4 Temporality

The flow of time in music is one of the primary ways in which a sense of narrative is created. This temporal aspect of narration is discussed extensively by philosopher Paul Ricoeur in his magnum opus, *Time and Narrative*. In literature, narration involves a complicated relation of time and tense: what is being narrated, in most instances, must have happened in the past, but it unfolds in the time of the author's recounting and the reader's current imagination. The appropriation of the narrative

voice in music is most obvious when a quotation is used – something that has already been said – or by allusion to something outside the musical space. Thus the frame of time in music, which is in continual state of flux, becomes one of the most vital defining features of narration. As in literature, so in music, there can be a discrepancy between the lived time and the alluded time, resulting from the relation of the musical objects in the musical space: "the modes of folding by means of which the time of narrating is separated from narrated time" (Ricoeur 1984 vol.2: 78).

The semantic problem arises when we try to find an analogue to the linguistic concept of tense. Musicologist Carolyn Abbate, in discussing Ricoeur, concludes that music, as opposed to literature, does not have a past tense, precisely because it is a temporal art. She claims that music, like theatre and other temporal arts, is mimetic by nature and thus "traps the listener in present experience, and the beat of passing time, from which he or she cannot escape" (Abbate 1991: 52). She constructs an argument on the impossibility of temporal narrative signs in music. Where I believe this argument falls short is in its confusing of the work 'itself' with the perception of the listener. Music uses strategies such as repetition, quotation, relative speed fluctuations (both harmonic and rhythmic), and rate of information flow, in order to communicate to the listener a sense of a fluctuating time. It is in fact the listener who creates the time frame: it is a relative perceptual concept. Perception and reflection on an event whether musical, visual or literary, inevitably entails a temporal aspect, without which there would be no communication; thus temporality and tense cannot be so easily dismissed. A definition of how narrative is constructed through time, is proposed by linguist Susan Fleischman in her book *Tense and Narrativity*:

"tense" is defined as *the grammaticalization of location in time*. More particularly, tense involves the location of situations predicated in a sentence or discourse relative to a reference time... tense is relational in that it involves at least two moments in time (which may coincide wholly or in part). (Fleischman 1990: 15)

The idea of tense as being a construct of *location time* and *reference time* underlines the idea of how narrative or narrativisation emerges from two distinct temporal levels. The fact that in music one cannot find a direct equivalent of the grammatical tenses found in language does not disqualify complex temporal relations from it, nor evident narration between these states. An example of this would be the use of live-sampling or processing of audio signals in the time-domain, where the musical present and musical past can be continually blurred, or the use of pre-recorded audio in a 'mixed' electroacoustic work, which constantly refers to and creates a dialogue with another space and time.

Meelberg takes Fleischman's definition and finds an apt translation into the musical domain:

The musical past is location time, the time of the events told. It is the time in which the music is regarded as consisting of musical events. The musical present, on the other hand, is reference time, the time of the narrator. It is the moment in which music presents itself as a continuous stream of sounds. (Meelberg 2006: 135)

I understand Meelberg's 'location time' in music to be the temporal plane, where the image or the representation that is constructed by the musical events resides: in other words, music's diegetic space. From a personal intuitive point of view and from my own conception of time in my compositional practice, I do not see past tense or location time as having a very defined linear perspective, as it does in literature. Because of music's ability to repeat, blur, morph, transform, or rearrange distinct sonic events, my own conception of the musical past is more akin to a space or a landscape gradually being uncovered, rather than a sequence of events stretching back in a sequential order. The order in which musical events are placed might be relevant for the musical present, but as they fade into the musical past they might be said to blur into a space that is accessible to the listener (in reference time) though not necessarily placed in any significant order.

The complexity of relations occurring between location and reference time, is exactly how meaning is created through what could be described as discrete narrativization. Furthermore, regarding the conclusion from the previous section as to who or what is doing the narration, it is perhaps too simplistic to say that it is only the musical performer or the composer who is responsible for the construction of the reference time. Certainly, the embedded musical voices within the composition also act to narrate between these temporal levels. In effect, one could already point to several ontological planes in musical performance, where narrativity is created in relation to temporality: the level of the composition, of the performance, of musical events and most importantly, how their parameters are perceived by the listener.

Another way that music creates a sense of narrative differentiation on the temporal plane, differently to literature, is in rhythmic polyphony. The auditioning of material on different rhythmic plateaux is another way of inferring different time levels; of being able to perceive one rate of change over another. This, in my view, is something analogous (though not equivalent) to 'tense' in literature. The perception of voices moving at different speeds reinforces a sense of different ontological levels where narration can materialize.

However, temporality is not the only way that different ontological planes are created in a music multimedia work. One of the differences between literature and time-based arts is that even though time is essential in articulating an event, things can happen at the same time. This means that one can perceive ontological differences between levels of material that are not separated by time or tense. In the

following subsection I will introduce the idea of 'frames' as a way of understanding how this occurs.

2.5 Frames

In any musical space, the distinction between the physical sounds we hear and the imaginary world and meanings these set in motion is of cardinal importance, as those differing levels of ontology are exactly what creates the idea that a narration is taking place. For classical narration to exist, there needs to be a hierarchical relationship between the world of the narrator and the 'narrated' world, in which the latter is embedded in the former (Vervaeck and Herman 2001: 83). One can understand this in terms of the idea of *frames*, where a frame is created on one level to expose something within it on another level. The important point here is that the frame is created already within one diegetic space, using the attributes and expressive potential of that particular world. In this use of the concept of frames, borrowed from Barsalou's idea of recursive attribute-value structures (Barsalou 1992), a model can be constructed in order to show how different media interact with each other to create meaning.

In Barsalou's model, which defines frames as 'dynamic relational structures', he defines a frame as consisting of three components: attribute-value sets, structural invariants and constraints. An attribute is a concept defining some aspect of the larger whole. Hammer, note, key, string, black, octave and chromatic are all attributes of the class 'piano'. A value is a subordinate concept of an attribute, so the key C# is a value of key, and also pedal might be the value of mechanism. Structural invariants are the concepts that define the relation between frames, where there might be an element of causality. Hands play a piano, the hands belong to a musician who has had special training. The hands are moving as an effect of lines and dots on a page, which the pianist is reading. Those lines and dots have a specific order and meaning, defined as a composition. These cognitive frames have a causality linking them and are therefore seen as invariant in some way. Barsalou's third component, 'constraints', is also a relational concept, but one with a variable relation and dynamic attribute values. So as the hands that play the piano move to the right, the notes get higher.

Translating this into the question of how meaning is created in a multimedia work, we could say that it is in these variable values that something expressive emerges. To understand the conceptual framework within which this variance in notes occurs, we have to intuit the structural invariants – the piano and why it makes sound – and have some kind of concept of the attribute value of a musical instrument or the idea

of music. This is analogous to how meaning is created between various frames of multimedia. We have an understanding of the variants and invariants inherent within a certain medium or ontological frame, and can look through one frame onto another.

There is a recursive aspect at play when we start hunting down the roots or the primitives of these structures. Looking for frames within frames leads to an almost endless quasi fractal hierarchical structure. Barsalou intends this as the basis for a computational structure of knowledge formation. Nevertheless, it serves as an interesting model of how music, and all its related media, cultures and contexts, is understood.

We can take the example of *song* to explain this analogy. Initially, we might see this within the frame of *song* or a *singer*, but then at a further point in the performance our frame might split into the two frames of lyrics and music, or even lyrics, music, sound and the performance space. One can look through the musical frame to the text frame, and back from the text to the music, constructing relationships and therefore engendering meaning; just like meaning is constructed between the various frames of words and phonemes constituent in the class of text. The same applies to all the other attributes, variant or invariant, that we construct in our minds, consciously or unconsciously, when listening to music. This is naturally a dynamic structure: an understanding of the frame we are looking through is constantly being updated with new information.

The reason for using the model of frames in order to understand how narration and multi-media relationships occur, is that it conveys the sense that we always look through something onto something else. We need the understanding and sense of one concept to understand another. We need to be able to distinguish variant from invariant for that understanding to occur; the insight we gain manifests itself in the form of a frame, opening the possibility of perspective. In the case of music, because we can distinguish between notes and instruments, between words and voices, between rhythm and sound, we begin to differentiate between levels of ontology, worlds or frames embedded within each other. In this way, a multi-framed work can exist in a musical form, music within music, just as it can in a work of literature, story within story. This is the case in my music-text-films, where the relation between perspective and frames of view becomes complex. However, while the possible complexity is increased, the opportunity to become aware of these perceptual frames also becomes greater.

If we are to accept that narration takes place when a space is created between differing ontological frames, then we could state that it is nowhere more pronounced than in multimedia work. Here, there exists a dynamic and metaphorical exchange between levels of meaning more apparent and conscious to the perceiver, than in a

mono-media work (if there can ever be such a thing), because the difference between the media can be clearly demarcated and compared. The benefit of the concept of ontological frames in understanding how multimedia work narrates is outlined in the following chapter, where I examine what is in fact meant by the term 'multimedia'.

Chapter 3: The Multimodal Voice

The sense of narration existing between differing ontological frames is no more clearly demarcated than when more than one medium is present in a work of art. How does this narration take place? Can one say that 'voice' is a transferable entity, not restricted to one medium or point of view but travelling between layers of music, text and image? This would entail a broadening of the definition of voice, not only to encompass the narrative sense (Chapter 2) or the inner voice (Chapter 1), but to include a multimodal voice weaving through other media, as it exchanges not only the perspective but the language of its expression. This involves metaphorical transformations from one medium to the other, so that 'voice' can at one moment be represented as text, or at another as music or image, slipping between them as a changeable entity. In this chapter I would like to propose an idea of voice not fixed to one perspective, but moving according the 'focus' presented in the work and the subjectivity of the spectator.

How, then, is a difference of medium demarcated, and what is, after all, meant by 'medium'? In this chapter I set out different approaches to understanding and analysing multimedia relationships, not only to underline the way this metaphorical voice moves through these ontological shifts, but also to be able to simply discuss how different media function in audio-visual work, music theatre, sound installation, and other types of so-called hybrid art; works where music plays a dominant role, including the tryptic of elements in my own music-text-film work.

The fact that terms such as intermedia, mixed media, multimedia, trans-media and cross-media overlap and blur their definitions, proves how difficult it is to find easy categorisations for work bringing to the fore a multiplicity of perspectives. This difficulty also unmasks the very slipperiness in pinning down exactly what is meant by a *medium* or *media*. Is it a channel of communication? The materials used to make a work? A language? A discussion of this subject could begin or even end with Marshall McLuhan's famous catchphrase: "The medium is the message/massage²⁵", that it is the medium itself that determines the message rather than reflecting it (McLuhan 1964: 9). Following McLuhan, if we accept that music, film, dance, television, internet, a score, a novel or a painting are all some form of media, then we must also accept, as per McLuhan, that they do not just communicate meaning but also construct it.

²⁵ McLuhan made a variation in 1967 on his much quoted phrase for the book he co-created with designer Quentin Fiore, intended for a wider public.

The instance of the electric light may prove illuminating in this connection. The electric light is pure information. It is a medium without a message, as it were, unless it is used to spell out some verbal ad or name. This fact, characteristic of all media, means that the "content" of any medium is always another medium. The content of writing is speech, just as the written word is the content of print, and print is the content of the telegraph. (McLuhan 1964: 8)

This idea of one medium enfolding another medium in a process of genealogical categorisation, could be taken to an extreme point where at the end one is left with pure information. Even when one looks at examples of how art practice in one medium can incorporate attributes of many others (as considered below), it is not easy to mark when one medium becomes many.

In order not to get too lost in the elusive definition of what constitutes a medium, and because it provides a useful parallel with the media that are prominent in my music-text-film, I will yet again call on Aristotle's three primary elements of tragedy, defined in his *Poetics* as 'opsis', 'melos', and 'lexis' (image, music, text). Aristotle's triangulation of mimesis into opsis, melos and lexis became one of the ways of talking about media relations in the last century.²⁶ This is a starting point that I use to discuss the analysis of media relationships; specifically how, in different historical periods, one medium gains dominance over others, or how the attributes of different media merge or influence each other. It is also interesting to note how the two poles of what can be described as 'media transparency' can be seen to be represented on the one hand in Aristotle's bias towards the media serving mimetic integrity, depicting the real as faithfully as possible, so that the medium becomes invisible, and on the other hand in art critic Clement Greenberg, the high-priest of modernism, who argues for medium-specificity: the medium to be defined only by itself, rather than what it is depicting.

Continuing from there, I propose a system of analysis that works on six different aspects of the relationship between media: on what I call a 'sensory' level, in terms of space, synchronisation and scale, and on a 'semantic' level, in terms of story, sentiment and style. Significant in time-based works is how this changes over time, so I choose to describe these relationships as 'converging' or 'diverging', as a way of underlining their dynamic aspects.

The second model I use is based on Lakoff and Johnson's definition of a conceptual metaphor, and how a cognitive model based on this is useful to understand the hierarchy created between two media. Metaphor has been applied by many writers on multimedia art as a principal concept in the ways we can understand the function of different media. In short, a metaphor is a figure of speech that describes one set of

²⁶ <http://csmt.uchicago.edu/glossary2004/melosopsislexis.htm> Chang, Vanessa. 2004. Melos, Opsis, Lexis. Keywords Glossary. University of Chicago. Web. 10 July 2017.

ideas in terms of another (Lakoff & Johnson 1980: 5). The two parts of a metaphor have an unequal relationship: one borrows attributes from the other to highlight some meaning, and this inequality is underlined by their cross-domain correlation. It is precisely the idea of inequality, rather than balance or equivalence, that I find a fascinating key to understanding our perception of multimedia. With metaphor there is always a hierarchy and this hierarchy is exactly what creates a narrative 'focus' in the work.

Finally I explore these ways of looking at multimedia through a case study of a work of mine, *Subliminal: The Lucretian Picnic*, looking at how the different relationships between image, music and language intertwine, and how the media hierarchies shift throughout the course of the piece.

3.1 Opsis Melos Lexis

Since the actors do the *mimesis* by acting it [out], a first aspect of tragedy [-making] would have to be arrangements for the 'look' (*opsis*) [of the actors and stage]; then song making (*melos*) and the [devising of] speech (*lexis*), for these are the 'matter' (in-what) the *mimesis* is done in. (Aristotle in Whalley 1997: 69)

Opsis, melos and lexis, perhaps more suitably translated for this context as image, music, text,²⁷ are, according to Aristotle, three of the six basic media of drama.²⁸ The quotation above is from a chapter dealing with the six main elements of the poetics of tragedy, ordered in what he sees as the most important to least important: *Mythos* (story), *Ethos* (character), *Dianoia* (idea), *Lexis* (speech), *Opsis* (spectacle), *Melos* (music). Aristotle's table of hierarchy as regards the dramatic arts is relevant to this discussion because it sheds light on what he saw as the real mimetic power of art. As mentioned earlier, even though Aristotle had more sympathy than Plato for the arts in general, he felt that true power came from a closer depiction of the 'real', in both a natural and philosophical sense. One could infer from this that any medium is inherently problematic to the experience of true immediacy, and in this hierarchy he certainly seems to favour transparency of representation in art.

The history of aesthetics from Aristotle to the present is peppered with many attempts to reorder the primacy of these three media, not only in the specific art practices prioritised in a given era, but in delineating the borders in which they are drawn. As pointed out by art historian Simon Shaw-Miller, the early modernist

²⁷ The order that I have chosen to use, is as they appear in Aristotle, which also happens to be the title of Roland Barthes' 1977 collection of essays: *Image Music Text*.

²⁸ When talking about poetry, Aristotle names rhythm, language and harmony as the three main modes.

period brought a resurfacing of attention on the function of media (Shaw-Miller 2014: 48). Two iconic critics of the period, who could be said to be on opposing sides of the fence in terms of how they considered the opacity of the media, were literary critic Northrop Frye and Greenberg.

Frye's conception of the triad of media as constantly collapsing on themselves, or elements of one appearing in the guise of another (such as musical or visual elements that appear in poetry) seems highly relevant to this discussion:

Considered as a verbal structure, literature presents a lexis which combines two other elements: melos, an element analogous to or otherwise connected with music, and opsis, which has a similar connection with the plastic arts. The word lexis itself may be translated "diction" when we are thinking of it as a narrative sequence of sounds caught by the ear, and as "imagery" when we are thinking of it as forming a simultaneous pattern of meaning apprehended in an act of mental "vision."

(Frye 1957: 243)

Since Frye is dealing with lexis as the primary medium, he sees melos and opsis as functioning within it. One could extrapolate this idea to other art forms, and seeing the constituent role played by the secondary media within that: the role that lexis and opsis play in concert music, or how melos and lexis function within painting. In regard to the latter, Greenberg, modernist art's most influential voice, was not only a strong advocate of medium purity, but also called for the greater influence of melos rather than lexis in visual art. In his essay "Towards a Newer Laocoon",²⁹ Greenberg suggests that opsis should be tending towards the musical rather than the narrative, because of music's inherently abstract, non-representational nature. He considered that literature had peaked as the dominant art form in the age of enlightenment, and that subsequently music had begun to play a more important role, becoming the prototype of all art. The transition of one art's position in society to another also entailed the evolution of the qualities specific to that particular art and medium, and so, according to Greenberg, music's less valued role in the twentieth century also included the transference of its qualities to the visual medium:

The dominant art in turn tries itself to absorb the functions of the others. A confusion of the arts results, by which the subservient ones are perverted and distorted; they are forced to deny their own nature in an effort to attain the effects of the dominant art. However, the subservient arts can only be mishandled in this way when they have reached such a degree of technical facility as to enable them to pretend to conceal their medium. (Greenberg 1986: 22)

²⁹ This was a reference to both Gotthold Lessing's "Laocoon: An Essay on the Limits of Painting and Poetry" (1767) which argued for greater differentiation between painting, sculpture and poetry, and Irving Babbitt's 1910 essay, "Laokoon: An Essay on the Confusion of the Arts."

What Greenberg thought to be the dominant art of his era was the emerging abstract expressionism in painting, which had absorbed some of the attributes of music: what he considered to be the previous dominant art form. In spite of the somewhat crude analysis regarding which art constituted the most dominant in a given era, Greenberg's model of how the qualities and attributes of one medium are constantly being subsumed and evolved into another in the evolution of art is a compelling one. Fundamentally, Greenberg sees the separation of media and the aspiration towards medium purity, as the ideal of art:

Each art had to determine, through its own operations and works, the effects exclusive to itself. By doing so it would, to be sure, narrow its area of competence, but at the same time it would make its possession of that area all the more certain. It quickly emerged that the unique and proper area of competence of each art coincided with all that was unique in the nature of its medium. Thus would each art be rendered "pure," and in its "purity" find the guarantee of its standards of quality as well as of its independence. (Greenberg 1993: 86)

This is a position diametrically opposed to that of Aristotle. Whereas Aristotle sought transparency towards mimesis, or the copy of the real, as the ultimate goal of art, Greenberg sought opacity of the medium itself; the medium to be defined by itself rather than anything it was representing: so-called medium specificity.

One can see that there are various ways to arrange, prioritise or configure this triad of media, and it seems that each era, culture and ideology, might have its own preferences as to how art is served. What is interesting to extrapolate here is how one can speak of a medium in not only its material form, but in its various attributes, and how those are embodied within each other. Musical performance, for instance, no matter how 'absolute', will always involve an aspect of opsis, in how the spectacle of the concert is communicated – concert space, performance ritual, gestural actions of performers – and lexis – titles, concept of the work and other paratexts. The extent to which the secondary media play a role in influencing how one experiences the primary, comes to the fore when speaking about multimedia work, and it is perhaps useful to make an initial evaluation as to the weight and hierarchy put in each of these categories. Sometimes this is not possible, either because a form is finely balanced between two specific media (as for instance in ballet or sound poetry), or because the media use goes beyond traditional categories of text, sound, and image, to utilisation of other senses and structures (such as can be found in interactive, relational or participatory art.)

As I will explain in the next section, hierarchy of media becomes important when discussing how meaning is generated between the media in a metaphor model. In its most basic form, metaphor borrows meaning from the 'source' (the secondary medium) to understand something of the 'target' (the primary medium). This does

not have to be static, and in much multimedia work it is constantly shifting. But as a first evaluation it is useful to understand how social context, audience expectations or even the creator's intention, can affect the way the media are stacked up. In the first place, social context, or the dominant norms of the entrenched practice have a huge effect on how the hierarchy between the media are defined. Taking theatre as an example, various traditions in the 20th century have shuffled the order of hierarchy of the media. In the case of Robert Wilson's form of theatre, taking iconic works such as *Einstein on the Beach*, *The Black Rider*, or *A Dream Play*, even though lexis is important, one could argue that the theatre is primarily defined through other media, such as melos and opsis, to the extent that the latter adopts a more dominant role. The theatre of Berthold Brecht, on the other hand, could be said to prioritise lexis over everything else, while the conscious separation of the media, is used to enhance distance and alienation. In traditional opera, melos could be said to take the primary role, and lexis is subsumed into the act of singing or through the narrative meaning conveyed in the orchestral score.

The expectation created by these entrenched forms is embodied in the way particular audiences experience multimedia work. In my own art practice I have come across different audience expectations, depending on the context in which a particular work is shown. In general, one could say that spectators at a fine art exhibition tend to not prioritise music or sound, expecting instead to focus on opsis, the spectacle and the way meaning is generated in relation to lexis, the narrative or concept.³⁰ Sound is often here experienced as representational, standing for something other than itself, or perhaps used in a cinematic sense, subsumed totally into the opsis. These expectations are often entrenched in the very spaces where art work is shown, which is why the practice of taking art out of the gallery or the museum, music out of the concert hall or theatre to unusual locations, increases the possibility of destabilising the well-established hierarchy of media. The idea that I would like to reinforce here is how the context of a work, historical, social and economic, determines the hierarchy of media. The context in which we experience a work can be in harmony or dissonance with the media itself; it can amplify one medium and mask another.

The question of how one medium collapses into another is alluded in a more general sense by musicologist Kramer in his discussion of musical meaning (Kramer 1988), whether in song form sound and words collapse into a single *medium*, or whether they remain as distinct *media*. At what level of cognition does this matter, on the level of the work itself or on the cognitive or sensory level? Kramer touches on the idea of medium through the examination of the residual meaning around art. His concluding stance is that the musicological position taken in the last centuries, trying

³⁰ I give an example of this in Chapter 5.3.2, discussing my work *Disco Debris*, and the problems encountered in creating sound based work in a visual art context.

to divorce meaning or context from pure music, is both a futile and misguided venture. Music will always burst open into its constituent parts, which will include meaningful elements such as technology, social circumstance, or historically constructed musical parameters.

The ubiquity of the problem suggests that something is fundamentally wrong with the core assumption that musical autonomy equals absence of meaning. If so, identifying that something might open the possibility of a musical hermeneutics no longer burdened by the foregone conclusion of its own futility or its inferiority to the purely musical... For although music minus meaning can be placed in its cultural context, it necessarily remains inert there; since meaning resides in the context alone, the music can at best be a symptom or token of some contextual element. (Kramer 1988: 14)

Another possible definition of what constitutes a medium, is given by philosopher Jerrold Levinson, in his essay *Hybrid Art Forms*, where he sets out conditions for hybridity as art forms, which arise from combinations of earlier art forms. For Levinson, the historical definition rather than material, is at the root of medium.

A medium is a developed way of using materials or dimensions with certain entrenched properties, practices and possibilities. (Levinson 1984)

How a 'historically defined art form' retains its status is a tricky assumption, since this is only temporarily assured; there is no universality to speak of here. By the same token, the way two historically defined media such as music and drama become subsumed into the medium called 'opera', whilst still retaining their constituent identities, was a centuries long trajectory with many pitfalls. Nevertheless, it is an interesting reminder of how historical use of a form of cultural practice is partially responsible for how a message is communicated, or how it is understood.

The triad *opsis*, *melos*, *lexis*, (mirrored in Barthes' *Image Music Text*), belongs in a broader category of definitions, trying to understand the function of media through a tripartite model. This triangulation also serves the purpose of highlighting a sense of mediation itself, and of breaking a too binary or dialectic approach to discussing multimedia art. Nevertheless, in the following chapter I will propose a more useful system of analysing the correlation between media in terms of a binary comparison of different parameters. This does not give us the complete picture, but it suggests a way of examining how different media compare or form metaphorical relationships. Eventually, utilizing these perspectives, one can form a more complex and complete image of the dynamic of interactions in multimedia work.

3.2 Media Correlation

Throughout the last century, various methods have been proposed for analysing the relations different media have with one another, whether in theatre, film, ballet or other forms with strongly entrenched association. Early examples focused on what we might now call 'synaesthesia' forms, where theories of the relation between sound and light, or perhaps tone and colour, tended towards a one-on-one correspondence, often walking hand in hand with ideas about the universal connectivity of things. Goethe's colour theory was one major influence on these aesthetics. Traces of this influence are clearly evident in the works of Alexander Scriabin, in the writings of Vassily Kandinsky (*On the Spiritual in Art*) and the subsequent theoretical underpinnings of the Bauhaus movement.

The crux of Goethe's idea, later borrowed and developed by Kandinsky, was that sound and light are connected not directly to one another other but through a higher spiritual order. This is explained in terms of a river metaphor, where the mountain is the higher formula:

Colour and Sound do not admit of being directly compared together in any way, but both are referable to a higher formula, both are derivable, although each for itself, from this higher law. They are like two rivers which have their source in one and the same mountain, but subsequently pursue their way under totally different conditions in two totally different regions, so that throughout the whole course of both no two points can be compared. (Goethe 1970: 748)

This idea gave succour to the development of theories at the end of the nineteenth century, in step with the spiritualist vogue of the time, which seemed to explain the connectivity of sound and light, not in the subjective sense that they might be experienced by a real synaesthete, different in every case, but in a universal, pseudo-scientific sense. These connections were often underpinned by a symbolic meaning, and we find this in the original Goethe text as we do in the works of Scriabin, Kandinsky and Arnold Schoenberg. Kandinsky and Schoenberg were friends for some time and naturally exchanged ideas about this. This interest gave birth to two early ground-breaking multi-media stage works, Kandinsky's *Der Gelbe Klang* and Schoenberg's *Die Glückliche Hand*. What many of these works had in common was the conviction that because there is a deep connection between harmony (Scriabin) or tone-colour (Schoenberg) and the visual spectrum, some deep spiritual meaning could potentially be unleashed through their joint expressive powers.

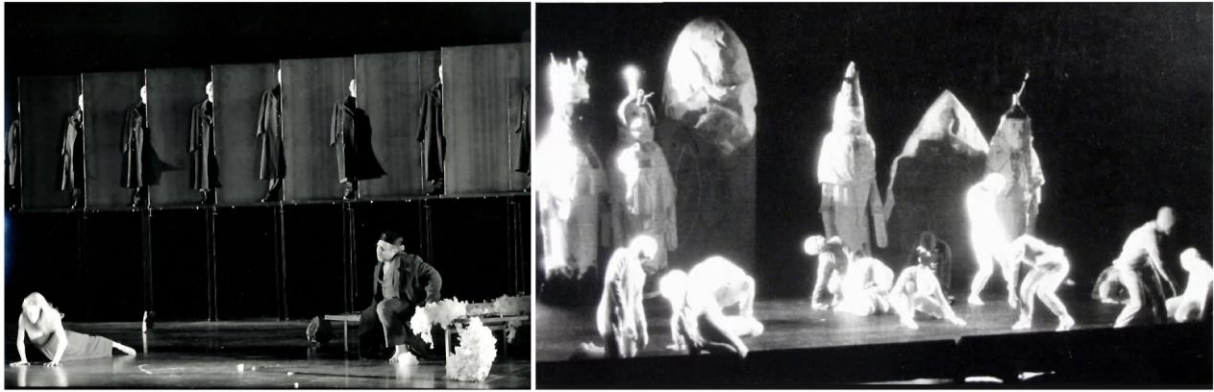


Figure 3 Photographs from stage performances of *Die Glückliche Hand*, Dutch National Opera (left) and *Der Gelbe Klang*, Guggenheim production, photo by Marilyn Mazur (right).

This idea of amplifying one medium with another was at the time (and some would say still is) the dominant approach to working with different media. This, I believe, had much to do with the entrenched forms of multimedia which had developed in opera at the end of the nineteenth century. The amplification of meaning has everything to do with the 'upsizing' of culture's scale of transmission. The ever expanding opera houses popping up around Europe and the new world, catering for the new bourgeoisie, brought with them ever grander themes expressed in an increasingly ostentatious manner. Part of this arsenal of grandeur was the effective deployment of new media to its end.

What all these early multimedia forms had in common was some kind of equivalence between forms; sound, light, text or action, either related directly to each other or through some higher bridging concept. These types of constructions justified the use of new media as well as amplifying the effect of the old media. According to musicologist Nicholas Cook (1998: 39), this tends to lead to a redundancy of one medium: the light organ in Scriabin's *Prometheus* does not express anything independently, because it is enslaved to the logic of the harmony of the musical score and thus only 'illuminates' it. Cook's point about redundancy is broadly relevant in theory, but I do not fully concur with him in practice, especially as regards *Prometheus*.³¹ Even a direct translation of one medium into another will always yield a difference in perspective; there is a hierarchy in how this is experienced, but this does not discount the possibility of perspective shifting multiple times through merely one listening. Moreover, while the light and the harmony are conceived and scored together, and theoretically the light is to a large degree enslaved to the music, in practice, when one experiences a performance of this piece, it is almost impossible to predict or read one medium through the other. The light seems to alter the way one listens to the score, both in its immersive effect and the way it subdivides the structure of the piece into larger phrases. One

³¹ An analysis at the end of this subsection will also illustrate this point.

experiences a different type of listening because of the effect of the light on the music. Subtle harmonic shifts are brought to the fore, while the macro-structure is reinforced through the changes of light. Conversely the light could be said to be problematised by the music, as if the slower shifts of colour come about as a result of an inner logic in the music. There is also a conceptual level at play where the idea of man-made light is itself the subject of the Promethean legend of the piece, so that light in some sense becomes the protagonist of the musical drama; it underlines a grand narrative, which is explicit both in the score and in the conception of the piece. In this sense, light becomes a 'voice' in the work.



Figure 4 Stills from Peter Struycken's 1998 visual interpretation of Scriabin's *Prometheus*.

It is important to note that a relation between media works on many different levels. For the sake of clarity, and the facility to analyse relations in more detail, I first divide this into two basic categories, that concern the relative immediacy by which the connections are experienced. These can be defined as 'sensory' –something that is largely perceptual rather than cognitive, experienced through the senses rather than deliberated over, versus the 'semantic', which deals with the various forms whereby meaning, or significance is generated.³² In the example above, I would describe the level at which we experience the music in relation to light as sensory, and the level where the music, combined with the light, narrates the myth of Prometheus as semantic.

Cook proposes a form of analysing multimedia relations through a flow-chart structure where one first questions the similarity of the content: is it consistent or coherent? If it is 'consistent' then there is a relationship of 'conformance'. If it is 'coherent', then another test (this time of difference) will determine if it is 'contrary', leading to 'complementation' or 'contradictory', which in turn leads to a relationship

³² I do not want to speculate that direct 'phenomenological' experience does not involve any cognitive reflection, since this is a highly debatable point, but by 'sensory' I want to convey the idea of the immediacy of sensory experience, along with its perception.

of 'contest'. This 'C'-friendly flow chart might be a little confusing as it is at times too complex to categorise works very clearly into one class or the other. In fact, most multimedia works one would put through this chart would end up in the 'complementation' category – the middle one – so although these are valid questions to ask regarding the interaction of different media, the classification system is insufficiently articulate.

However, as I suggested earlier, relations between media in a multimedia work function on many different levels, and change over time, so that something that coheres on the sensory level, for instance in synchronicity, might be contradictory on a semantic level. Moreover, something that might seem contradictory in the beginning of a work, might seem fitting by its end, once the logic of the relationship has made an impression on the viewer.

Building on Cook's idea of testing the coherence of the different layers making up a work of multimedia, I would propose a system that separately analyses different aspects of each medium, and compares their convergence in a correlation chart. I identify six aspects to be analysed, falling into the two basic categories mentioned above: the sensory, and the semantic.

In the **sensory** category, I define three main modes to be analysed: 'synchronisation', 'space' and 'scale'. In each mode, the aim is to evaluate the degree of convergence between the media. The choice of these parameters is based on primary questions about the media: 'do they occur at the same *time*?', 'do they occur in the same *space*?', 'what is the *scale* of their appearance'? These are questions that do not depend on specific parameters relevant only to a specific medium, such as 'colour', 'pitch', or 'symmetry'. Other possible parameters such as 'rate of information' would be discussed under the temporal category of 'sync', or 'proximity', which would fall under 'space' or 'scale'. Naturally, not all categories are relevant for all possible examples, and certainly some pieces might blur the distinction between scale, time and space, but this system of analysis is meant as a way of establishing a conversation about media correlation, rather than creating a fixed specification.

High level of convergence in the 'sync' mode results from extreme temporal unity. In the case of film, this means a soundtrack representing sounds exactly synchronised to movement or edits in the image. Low convergence (one could also say divergence) results from asynchronous sound; sound and image having a different temporality or rhythm. As I mentioned earlier, any temporal convergence of two media effects a perceptual amplification, while temporal divergence leads to the experience of the media as separate.



Figure 5 Still from a Tom and Jerry cartoon and a stage performances of Cunningham & Cage's *Walkaround Time* (1973), examples of convergence and divergence on the temporal and spatial mode (sync and space).

In the 'space' mode, convergence means that sound and image (in cinema) emerge from the same diegetic space. In the case of an audiovisual installation, this could mean that sound and image are emanating from the same room or object; that they share the same physical space. An example of divergence here is when sound and image imply different spaces; where the sound or music seems to exist in another space from the image.³³ This is in fact very often found in examples of film music where there might be strong convergence on the temporal dimension, but the music is clearly coming from an orchestra that does not exist in the same diegetic space as the actors (as with 'Tom and Jerry' cartoons, for example, where some sounds that could be thought of as existing in the space have been abstracted to musical instruments). An example of the opposite – temporal divergence with spatial convergence – is a John Cage & Merce Cunningham music and dance collaboration, where both actions are happening in the same space, so that a relation is being imposed one on the other; but because no temporal synchronicity has been negotiated the movement of the dancers and the sound of the musicians seem to retain their own independence.

³³ One could argue that the awareness of space is more semantic than sensory. I propose here that spatial awareness belongs to a fundamental human perception, related to basic functions of hearing and seeing; and as I mentioned earlier, I do not discount some level of cognitive processing in this awareness.

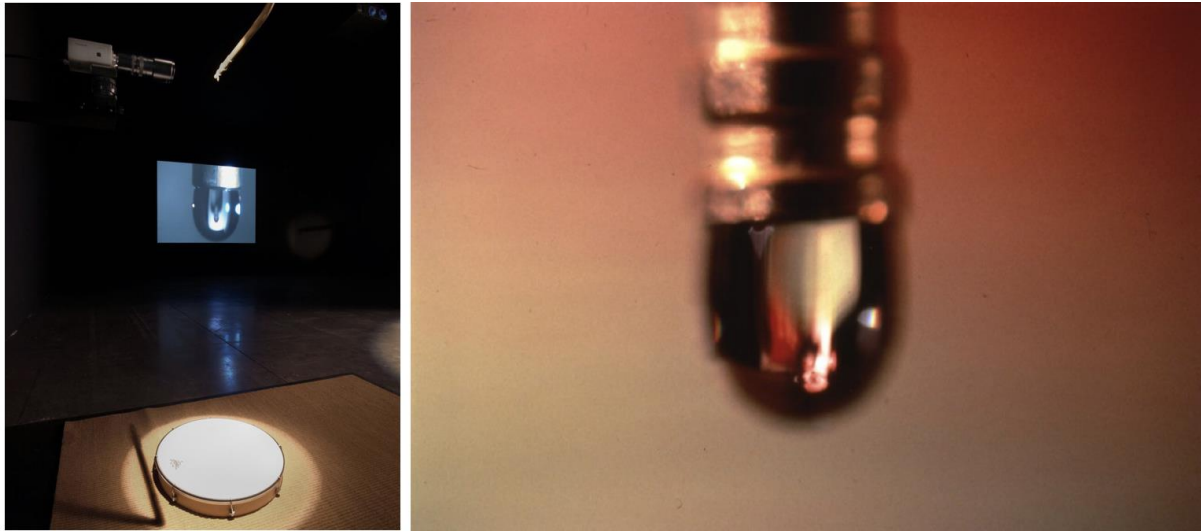


Figure 6 Bill Viola's *He Weeps for You*, an example of how the 'scale' mode can be utilised to extraordinary effect.

The '**scale**' mode considers whether the dimensions of the media are comparable. Setting extremely loud sound to tiny images, or the opposite – a huge immersive projection together with a quiet distant sound – are examples of weak 'scale' convergence. In conventional multi-media forms, scale is often used to express something about the distance of the viewer to the object: the narrative perspective. The cinematic trope of using sudden bursts of sound to imply the closeness of an object is one conventional example of 'scale' convergence. In some contemporary audiovisual work, the discrepancy and coherence between the scale of the media is highlighted for specific effect. An example of this is an early Bill Viola work, *He Weeps for You*. In this work, a drop of water coming out of a tap is magnified and projected live onto a large screen, where one can even see the viewer inside the water drop. As the water drop falls, it hits an amplified frame drum and is heard as a massive sound with enhanced low frequencies. This work clearly deals with scale on many levels. The amplification of the drop is in coherence with the scale of the image, but both are clearly a distortion of the scale of the actual drop and drum in the room. In other works of Bill Viola, the scale of sound or 'undersound', as he refers to it (Viola 2002: 91) – an almost inaudible ever-present sound that is amplified and brought into conscious perception – is used to enhance the sensual experience of the installation.

In the '**semantic**' category, one could analyse the meaning of what is being expressed and look to see if the media converge or diverge. In order to clarify the idea of what is being expressed, I divide this category into three modes: 'style' 'sentiment' and 'story'. Again, similarly to how the 'sensory' category is divided, I look at three questions as to how meaning is conveyed in each media: 'are the same cultural references or *style* used?', 'is the same *story* being told?', 'is the *sentiment* or emotion that is conveyed comparable?' There is naturally an overlap between some of these

parameters and like the 'sensory' category, these will not necessarily be relevant in all examples.

'Style' refers to the language used, the set of entrenched patterns and the associations it brings with it. Meaning arises not only from narrative patterns but from the cultural associations resulting from the way things are expressed. When we analyse a multimedia work that makes use of two distinct artistic practices with their own social and historic development, we may ask: does what we hear and what we see cohere in terms of cultural frame of reference? 'Style' coherence is in part concerned with examining whether the media arise from the same culture or era. A common divergence here is where classical opera is staged in a contemporary setting, as with Jean-Luc Godard's staging of Jean Baptiste Lully's *Armide* (1987) in a bodybuilder's gym or, conversely, when classical music is used in a film sequence that bears no cultural connection with what the images shown.

A good example of this is Werner Herzog's use of classical music in his post-apocalyptic film *Lessons of Darkness* (1992), where shots of an alienating landscape of burning oil-fields are accompanied by the music of Puccini and Verdi. Martin Scorsese's use of Mozart in the boxing sequences in *Raging Bull* works similarly. This is not an uncommon trope, because classical music can represent a sense of shared culture heritage, and certainly the epic nature of the music can influence how we see the images. This is what I would classify as a narrative effect. Nevertheless what 'style' divergence brings to the fore is the problematisation and creation of distance between the media. The juxtaposition opens up a gap in meaning, a space for us to ponder. It also places more semantic weight on the narrative aspect of the separate media. One could even argue that divergence creates more meaning because it provokes more questions.



Figure 7 Stills from Werner Herzog's *Lessons of Darkness* (1992) and Jean-Luc Godard's setting of *Armide* in the anthology film *Aria* (1987), showing divergence of image and music in the 'style' mode.

The 'story' mode is the concept communicated in the established use of the media; the idea that is being expressed. The question here is whether (staying within the example of film) there is a different story being told by the music and by the images.

This is sometimes trickier to answer than simply posing the question. The reason for this is, as I explained in Chapter 2, that a dominant meaning arises from the combination of the media, rather than the media in isolation, and also because we tend to focus on what they have in common rather than on how they differ (the metaphor instinct). Resultant 'meaning' aside, be this narrative intention or the intellectual ideas of each medium, what meaning is each medium engendering by itself? Is it comparable?

Finally, '**sentiment**', which can be but is not always closely related to 'story', relates to the emotion or affect that is being created in the different media. Again, like 'story,' the combination of media creates the dominant feeling that is being communicated, but sometimes this is not as convergent as we imagine. Also, convergence in 'sentiment' does not necessarily have to concur with convergence in 'story'. An example of this can be found in Michael Snow's seminal experimental film *Wavelength* (1967). In this film, over 45 minutes the camera zooms from the back of a room to a photograph of waves on the other side of the room. This is accompanied by a non-diegetic sine wave tracking of the zoom, the pitch moving steadily higher through several octaves. The 'story' of this relationship could be said to be very convergent. The sine waves are taken as a sonic equivalent of what the camera is doing: zooming. But the emotional affect is very different: there is an increase of tension, resulting from the sine waves' ascending frequency, which does not necessarily cohere with the sentiment conveyed by the steadily nearing close up of the still waves in the photograph. (If the sine waves were descending, the emotional result would be different.) There is, therefore, an increasing divergence between concept and emotional result, which is part of the reason why this is so compelling. In another part of the film, the emotional information between sound and image is reversed: two women are sitting by the window listening to the radio (to the song *Strawberry Fields Forever*). There is a divergent sense of emotion, since the relaxed, dreamy feeling conveyed in the song is at odds with an almost disturbing sense of transience that the scene depicts, in its constant change of light and extreme objectivity.



Figure 8 Two Stills from Michael Snow's *Wavelength* (1967) showing the two women listening to the radio, near the beginning and the final close-up of the photograph of waves at the far end of the room.

Table 1 The six modes of media correlation, using the example of sound and image.

	Convergence:	Divergence:
Sync	Sense of synchronisation. An image or action is reinforced by an analogous rhythm in sound.	Media are not in sync and have a different tempo or rhythm.
Space	Shared space. Real space and diegetic space. Sound and image seem to be coming from the same place.	Sound and image imply different spaces. Music seems to exist in another space than image.
Scale	Dimensions of media are correlated. For instance, quiet sound with small projection.	Inequality in dimension of media: one part is larger, louder, closer or stronger than the other.
Style	The vocabulary or style of narration is similar. The cultural association is shared. The languages of expression are closely intertwined.	The languages of the different media derive from separate cultural or historical periods. They differ in their social context.
Story	The concept or initial meaning expressed in each medium is analogous.	The idea conveyed in each medium is antithetical. A deliberate inversion of meaning.
Sentiment	The intended emotion expressed in each medium is similar.	The sentiment conveyed is unequal.

Below, I analyse the media correlation in the two iconic examples of work previously discussed: Alexander Scriabin's *Prometheus, The Poem of Fire, Op. 60*; and the final sequence of Werner Herzog's *Lessons of Darkness*. I assign a number between 1 and 5 (a low number signifying poor convergence (or divergence), 3 being neutral, and a high number indicating high convergence), assessing the extent of the media

convergence in each example, and giving a short explanation. These numbers are plotted on a radar chart. The larger the surface area of the resulting shape, the more correlated the modes of interaction are between the media; the smaller the surface area, the less correlated. This does not necessarily mean that a higher convergence is richer, and lower is poorer, when it comes to media correlation: this is not a judgement of value, but rather allows one to analyse the relations operating in multi-media work and the nature of the medial interaction. Furthermore, because the 'sensory' and 'semantic' categories are spatially assigned to the top and bottom of the chart respectively, one can verify, at a glance, in which general category there is greater correlation.

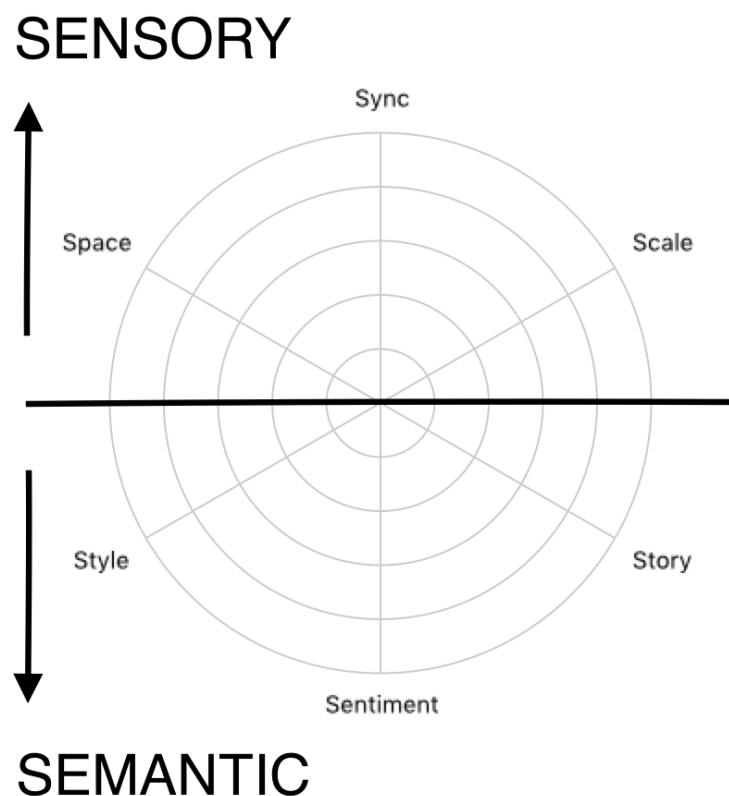


Figure 9 Media correlation radar chart.

Scriabin's *Prometheus, The Poem of Fire*, is often cited as a classical example of coherence of media (Cook 1998: 39), though it was not conceived in terms of a one-to-one relation between sound and light. Taking a hypothetical, contemporary performance of the work, which might be performed in a large concert hall, I will analyse the correlation between the music and the two aspects of the Luce part, usually played as a light object or video for the top part, and as lighting in the space

for the slower lower part. What one can witness below is that there is, in general, a greater degree of sensory than semantic correlation. This has to do with the relative difficulty in conveying a clear sense of narrative concept through two relatively abstract media (sound and light), but also because the technology used for the light as opposed to the music, tends to differ in cultural terms.

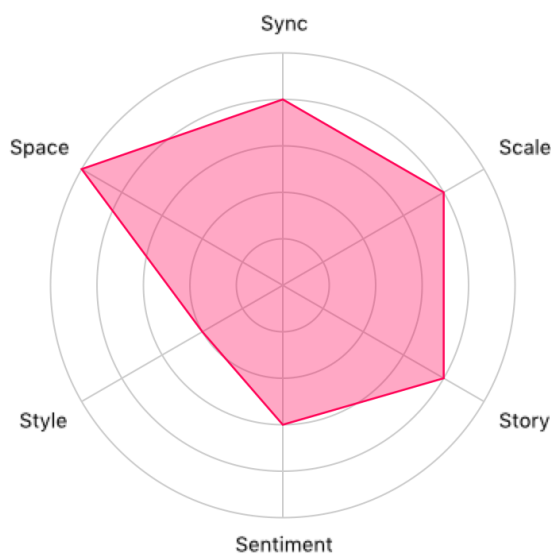


Figure 10 Correlation between music and light in Scriabin's *Prometheus*.

Sync: 4

Both 'Luce' parts have a clear correlation with the music, but not everything in the music is represented in the light changes.

Space: 5

Live performances often tend to include a spatial experience of the light. It is clear that both media are emanating from the same space, and have their origin in the same score.

Scale: 4

There is a hierarchy of scale implied in both 'Luce' parts, highlighted throughout the course of the music. However, the musical dynamics can sometimes exceed the dynamic range of the light: there is sometimes a discrepancy of scale.

Style: 2

The technology used to realise the lights is often a contemporary one, even showcasing the newest visual effects. This can sometimes be in stark contrast to the seemingly old-fashioned profile of the orchestra.

Story: 4

Even though the light cannot in itself communicate a clear narrative, it has been composed with a strong metaphoric intention. For those that can follow the symbolic narrative, it can communicate the 'programme' of the score (the myth of Prometheus) in a way that strongly correlates with the music.

Sentiment: 3

Similar to 'story', the emotional significance of the colour of the light is symbolically related to the particular tonality used. The question whether this communicates the experience of the emotion in a similar sense is debatable. The relation of colour to emotion remains subjective.

In contrast to a relatively convergent correlation in *Prometheus*, *Poem of Fire*, another fire related work, *Lessons of Darkness*, could serve as an example of Werner Herzog's particular documentary style where the distance of the perspective of the viewer is constantly being reinforced. The film was shot in the weeks after the first Gulf War in 1992 over Kuwait, and it tries to emphasise (some might say aestheticise) the devastating landscape and the human and environmental tragedy left behind.

The soundtrack of the film is largely composed of classical music, which for the reasons demonstrated below, enhances the effect of alienation, while at the same time giving an epic quality to the images shown. The scene I will analyse is the final one, where one sees oil workers, having stopped the cataclysmic fires, reignite the flow of oil. Accompanying this is Franz Schubert's *Notturmo in E flat major*, Op. 148, D. 897. From the resulting chart, one can infer that there is more convergence in a semantic sense than a sensory one. In general there is more divergence than in *Prometheus*, as is to be expected. What is interesting to note is that for the metaphorical relation between sound and image to work, there needs to be a seed of clear correlation in the meaning conveyed by images and music separately: the sense of accomplishment that is conveyed in the music and in the faces of the oil workers.

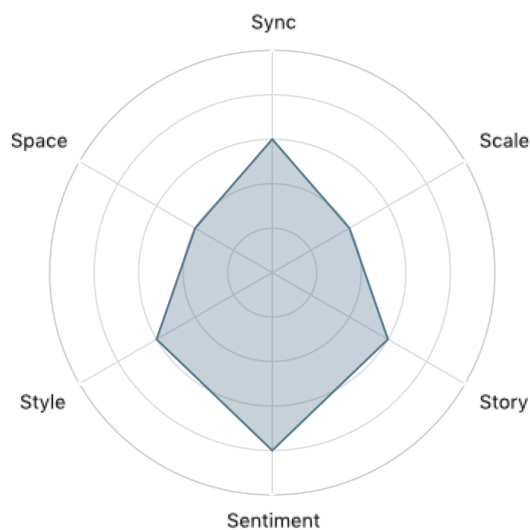


Figure 11 Correlation between music and image in Herzog's *Lessons in Darkness*.

Sync: 3

There seems to be an attempt at matching the pace of the music with that of the images, in that Herzog uses slow motion to impose a gracefulness in the movement of the workers. When the image moves at normal speed, it is emphasised by the diegetic sounds.

Space: 2

If we take only the parts in this scene where Schubert's music is used, one can say that music and image come from entirely different spaces. The contrast is highlighted with the contrasting diegetic sounds.

Scale: 2

The sense of scale that is expressed is somewhat divergent. The massive scale of what is depicted is in stark contrast to the intimate nature of the piano trio. Again, a sense of contrast is given with the diegetic sounds, which are loud and noisy.

Style: 3

The style of each medium, music and image, seem to be at odds with each other. The context of the music does not match the culture of the images being shown. One cannot imagine that this is the type of music that the oil workers listen to as they go about their business, nor that landscape of post Gulf War Kuwait has anything to do with the culture suggested by Schubert's music. On the other hand, within the context of the cinematic experience, a classical score is not so unusual.

Story: 3

The dominant narrative of the image is that of struggle to put out the fires. In the music, the narrative conveyed is one of contemplation, poise and stasis. There is therefore dissonance between the two narratives. This does not mean that they can't work together; on the contrary, this might result in a stronger metaphorical relation.

Sentiment: 4

In my view, there is an above average convergence on an emotional level. The music expresses a sense of effortless grace, perhaps of quiet contentment, a sense of accomplishment. This latter feeling is what communicates in the faces of the oil workers, a sense of doing good and satisfaction, probably the dominant mood that communicates through the combination of the media.

I do not claim complete objectivity in this method; it must inevitably rely on a level of interpretation of the user. Nor is it meant as a judgement as to whether a piece 'works' or not in relation to the correlation of different media. A poorer surface area of the radar chart does not mean that the piece is less valid, or less interesting. On the contrary, one could suggest that a more varied or contradictory shape might imply that the relations are less predictable, and therefore point to specific areas in the work that might say something relevant in terms of the artistic intention and effect. Furthermore, as is probably clear in discussing the Herzog example, even in one short clip one can witness huge dynamic shifts in the correlation of media. Therefore this method should be seen as a means to the beginning of a discussion, not the final verdict.

In the following subsection, I will discuss metaphor hierarchies as a method for analysing multimedia, and will then return to using this correlation method in a more detailed analysis of one of my works: *Subliminal: The Lucretian Picnic*, looking specifically at the differences between the correlations of the three dominant media.

3.3 Metaphor Hierarchy

In this part I expand on how the notion of metaphor – specifically what Lakoff and Johnson define as a conceptual metaphor (Lakoff & Johnson 1980: 8) – can be used to understand how we experience relations between different media. I begin with recent research into cross-modal perception – how the brain processes two or more different sensory modalities – and then borrow from Cook's idea of 'enabling similarity' (Cook 1998: 71), to try to understand the dynamic metaphor relationship between media and how it modulates through time.

On a neurological level, conceptual metaphors have been shown to have some kind of correlation to neural mappings in the brain (Feldman and Narayanan 2004:385). In 2001, neuroscientists Vilayanur S. Ramachandran and Edward Hubbard conducted an experiment (originally devised by Wolfgang Köhler in 1929) where they showed two shapes, one jagged and one rounded, to two groups (one American and one Tamil) and asked them "which of these shapes is 'bouba' and which is 'kiki'?". Both groups answered almost unanimously that the 'kiki' was the jagged one and 'bouba' the rounded. The researchers concluded that this shows 'synaesthesia-like mappings'

in the brain, where the visual and auditory cortices exhibit some kind of connectivity (Ramachandran and Hubbard 2001: 3). This was deemed to be evidence of the neurological basis of sound symbolism, an idea that vocal sounds are meaningful in themselves, that the brain creates cross-modal associations that are the basis for metaphor, and that synaesthesia and metaphor creation are linked.

More recently, neuroscientist Danko Nikolić has suggested that this experiment can be better explained by the concept of 'ideasthesia' rather than 'synaesthesia' (Nikolić: 2009). Nikolić defines ideasthesia as the phenomenon where concepts evoke perception-like experiences. Implicit in the idea of synaesthesia is the association of two sensory elements with little connection at the cognitive level; ideasthesia, on the other hand, puts emphasis on the cognitive aspect of cross-modal interaction, rather than the perceptual. Ideasthesia conceptualises the metaphorical connections of rational abstractions and the semantic links underpinning them. To clarify further, in the classic view our mind captures information, sounds, colours, smells, textures and tastes, and classifies them respectively as cicadas, blue, burning, rough or bitter. More recently, neuroscientists have concluded that these perceptions of the outside world are shaped by our conceptual understanding, that there is a constant feedback to our senses about what is being perceived, and that without the cognitive function there can be no perception. This implies a rich semantic network of both ideas formed by perception, and perception formed by ideas that underpin our understanding of metaphor, and suggests we do not need to be a synaesthete to have synaesthesia-like experiences, in other words, to have strong cross-modal metaphoric intuitions.

The importance of synaesthesia/ideasthesia as models of the experience of sound and image has been well documented and formed a basis of much multimedia work in the beginning of the twentieth century. In his book *Analysing Musical Multimedia*, Cook suggests that metaphor is a way of understanding how music and image interact, how we take ideas or emotions from one domain and map them onto another. Cook builds on the 'congruence-associationist' model of psychologists Sandra Marshall and Annabel Cohen (1988), which describes how meaning is ascribed to a film through music. In this model, attention is directed to the areas where music and film overlap, thus referential meaning associated with the music is ascribed to the overlapped audio-visual components. What is important here is that there is unidirectional transfer of meaning from one medium to the other, where there is perceived overlap of attributes. In Cohen's later study, "Congruence-associationist framework for understanding film-music communication" (2001: 259) she broadens the idea to a multi-stage model, involving 'bottom-up processing' (features derived from perceptions), 'cross-modal congruence' involving both semantic and syntactic features, 'top-down processing' (experiences in long term memory), and its interaction with consciousness and short-term memory.

Cook takes the idea of the 'congruence-associationist' model and proposes that it operates in a similar way to metaphor. He coins the term 'enabling similarity' to describe the way in which attributes common to two domains – the parts where they overlap – provide a basis for transfer of many other attributes not necessarily held in common:

The meaning of metaphor does not lie in the enabling similarity: it lies in what the similarity enables, which is to say, the transfer of attributes from one term of the metaphor to the other'. (Cook 1998: 71)

Once a link is created between concept A and concept B, through some idea of shared attributes, then there is the possibility of all of B's attributes being applied to A. Without oversimplifying what is a hugely complicated cognitive process, I would like to suggest that this unidirectional transfer of meaning, is one way we can understand what might be happening under the hood of consciousness when we are exposed to information on multiple sensory and cognitive levels. The fact that it is unidirectional does not undermine the idea that we might be experiencing this in terms of a constantly updated feedback loop of cognition, but it does underline the principle that a source and a target in metaphorical terms are never equal. For example, in the metaphor 'Time is Money', attributes of money are mapped onto our understanding of time, not the other way around: we do not see money in temporal terms (in this case), we see time in financial terms (Lakoff & Johnson 1980: 8).

In film music, is the music the source or the target? Are attributes of music used to understand the image or vice-versa? One would have to say that in the classic film-music relationship, where the narrative has the predominant role, the image is the target and the music is the source. Emotional attributes of the music are used to further understand the image. In a hypothetical scene where a young boy is mourning the death of his pet dog, the attributes of the music with which we associate the emotions of the boy give us an insight into his feelings of loss. We understand the meaning of the situation through the music. What is not the case, at least initially, is that we come to an understanding of the meaning of the music through the situation of the boy's mourning. This may be because the meaning of the music is not seen as paramount to the narrative; it is not our primary concern. The plight of the dog and the boy are.

Does this mean music will always adopt the role of 'source' to the 'target' of image or text? Not necessarily, because these kind of metaphoric relationships are dynamic and are constantly being updated. However, I would argue that they are never equal or balanced. Furthermore, they are often highly subjective, in that a viewer at any one moment, might be looking for one meaning through the image and another meaning through the music.

There is always a hierarchy in how source and target (in the terminology of metaphor) are defined.³⁴ In language, this is created by the syntax: the order of the words. But how does this manifest in a non-linguistic idea of metaphor?

The hierarchy of relationships between different media depends on many factors. Primarily, the entrenched historical form of a given genre is what mostly determines our point of view. These conventions play a large part in determining our position in the experience of multimedia. In much new audio-visual art the relationships between the media are not as entrenched as, for example in mainstream cinema, theatre, or even mainstream music video, so there can exist a greater dynamic range in how the hierarchies are constructed.

In the music-text-film pieces cited in this paper, in which at least three distinct media interact in our minds, the metaphoric hierarchy of these relationships – what is the source and target – play an important role in how the meaning of the work is eventually constructed. This is specifically what creates 'focus' in the work, affecting the perspective of the spectator. We might understand the music through an idea in the text, or the text through an idea in the music, or the music through a visual idea etc. Furthermore, these relationships are always in a state of flux. I have observed several ways in which these metaphorical hierarchies are constructed and modulated in the context of my music-text-film pieces, which can be applied to audio-visual forms in general:

1. Context: As suggested in the subsection 'Melos Lexis Opsis', context plays an important role in shaping the expectations and focus of the audience. The social or historical context pre-determines the hierarchy of media to a certain extent. As mentioned before, in an exhibition the spectator will much more likely focus on visual elements and in a concert on sonic or musical aspects of the work.

2. Structure: A medium might establish itself as primary when it gives a clear signal about its formal consistency. A distinct form or structure can make a medium seem self-sufficient, and thus to exist in its own inner logic. A countdown, an alphabetical series, a scale or other perceptible large-scale forms will always underline a strong inner order giving meaning to itself. An example of this is given in the following chapter: Hollis Frampton's *Zorn's Lemma*, where the very clear alphabetical order the piece is structured in lends weight to the hierarchy of words over images.

3. Rate: The rate and density of information of a given medium often determines the focus it will draw. An active visual field or a dense musical score might tip the scales, and determine it as a source or target of a metaphoric relation, it will draw our

³⁴ "The conceptual domain from which we draw metaphorical expressions to understand another conceptual domain is called the *source domain*, while the conceptual domain that is understood this way is the *target domain*." (Kövecses 2002: 4)

attention up to a point. If the data rate is too fast, going beyond a certain limit, we might not be able to process the information anymore, and that layer might recede to the background again.

4. Scale: The dominance of a medium could simply be a matter of scale. That which is bigger, closer, or louder tends to occupy more perceptual space in our consciousness, and therefore naturally finds its way into the foreground.

5. Voice: A physically heard voice will often become 'target' rather than 'source'. This may be connected with how we prioritise vocal or linguistic communication above other types of information (Latinus & Belin 2011: 143). Voice, similar to a body in figurative art, is prone to creating a figure and ground perspective (as outlined earlier) because of its anthropomorphism. This reinforces the hierarchies by which the metaphor relationships are established.

6. Concreteness: When a medium references or includes clearly recognisable elements, they will dominate over more abstract textures and structures: concrete sounds in music or visual objects from the 'real' world will inevitably draw attention (Cohen 2014:17).³⁵

7. Order of Appearance: If a medium establishes its presence before another is added, then it will continue to be regarded as the dominant medium for a while after the new medium appears. This might eventually change, for some of the reasons stated above.

8. Junctures: These are points at which hierarchies change, where one medium that has been dominant is removed or replaced by something else: moments where there is an eruption of sound or a disappearance of sound, a text which suddenly appears out of nowhere, a visual close-up, a synchronisation point, a black out, and so on. Anything that disturbs or changes the previous order will create a juncture point, where a change in point of view comes under negotiation, and therefore the relation of target to source may change. My composition cited in the next subsection, *Subliminal: The Lucretian Picnic*, is full of these points; where the hierarchy between text, image, soundtrack, voice and ensemble is constantly changing through dropping out, addition, modulation of speed, and emphasis of the media.

These attributes are some of the ways in which hierarchies can be established and then shift during the course of a piece. Ultimately there is a large element of subjectivity at play; there is no hard and fast rule for the way we experience music or any audio-visual work, but understanding some of the ways in which music, language and visuals interact with each other can indeed give clues as to the ways in

³⁵ Thanks to Marko Ciciliani for suggesting this.

which we can approach creating a space that is dynamic, open, and charged with signification.

3.4 Asymmetrical Balance

This chapter dealt with the idea of 'voice' as a transferable entity, not being confined to one entrenched medium, but travelling between levels of music, text and image, depending on both the perspective that is presented and the subjectivity of the spectator. The reference to Aristotle, and the discussion of ideological standpoints of media relationships, underlines the idea of how historically and culturally dependent the construction of media hierarchies is. The Aristotelian *differentiae* of *melos*, *lexis*, and *opsis* initially suggests the idea of transparency of medium. How much is the medium responsible for the construction of what is being shown? This is highlighted by the levels in which the different media are subsumed into each other, share their properties, or aspire to the attributes of another. These definitions, as I have shown, are culturally dependant, their technologies and weight shifting from era to era, as do the artistic currency of their modulating hierarchies. Thus, in the first place, it is important to look at the relevance of each medium's expressive power in light of the artistic context surrounding each multimedia work. In which artistic context is something shown? What is the dominant medium in terms of the artist's background, the artistic milieu that the work is created in, and the intended audience? How is the work seen/received? Is this result markedly different from the intentions of the artist?

Analysing a work in terms of the convergence and divergence of media correlation (Chapter 3.2) is useful in understanding the dynamics of multimedia interaction on a more detailed level. I have shown how multimedia work can have high degrees of convergence on one mode while being divergent on another, and how these can shift throughout the course of a piece. The aesthetic preferences for convergence on certain modes as opposed to others are sometimes culturally dependant: for instance strong temporal convergence occurs in much audio-visual work inspired by synaesthesia, whereas one might see divergence in that mode in work by experimental film makers wanting to break the conventional synchronisation of sound and image.

Finally, examining the material in terms of the 'metaphor hierarchy' model, one could ask, on both a macro and micro level, how meaning is transferred from one medium to another. Taking the asymmetrical form of the conceptual metaphor model, it is useful to observe exactly how the properties of one medium come to serve the understanding of another. My assertion is that there is never a state of total balance at any given time between media. One always looks through one medium

onto the other: the 'source' in the metaphor couple is always used to understand something about the 'target'. The way hierarchies can be defined and shift throughout the course of a work results from a number of factors, which I have defined in eight categories. These asymmetries in relation to media are also exactly what creates points of narrative focus that affords the spectator certain perspectives.

The 'media correlation' model and the 'metaphor hierarchy' model can both be used to analyse multimedia works. In the following subsection I undertake an in-depth analysis of *Subliminal: The Lucretian Picnic* using these techniques.

3.5 Case Study: *Subliminal: The Lucretian Picnic*

In the next chapter, I present examples of works (not my own) that use projected text, thus outlining a historical practice of this particular multimedia form. Instead of organising this discussion chronologically or by genre, I use some of the ideas outlined above to structure the discussion in terms of metaphoric relations. These examples are described in terms of dialogic relations stemming from the dominant metaphorical hierarchies. Before this, I conclude this chapter by looking at a case study from my own work, highlighting the complexity of relations between three rather than two media. I will apply the analysis method of examining how the media correlate in different parts of the piece as a whole, in order to understand how attention shifts between each realignment of media at the start of the many sections. Because of the constant shifts in hierarchy, the points made in the previous sections are highly relevant.

The work in question is *Subliminal: The Lucretian Picnic*. Composed in 2003 for a 15-strong ASKO Ensemble, this is a 32 minute work making use of electronics and a 10x1 video projection across the front of the stage³⁶. *Subliminal: The Lucretian Picnic* deals with the type and quality of information we can process at the 'subliminal' level. This exploration takes the psychological idea of 'priming'³⁷, how an indirect exposure to a stimulus influences an emotional response, and attempts to use it as a metaphor of how meaning is negotiated in a multimedia work. The initial intention behind *Subliminal* was to explore the powerful connection between images, text, sound and music lying on the threshold of perception: how they affect each other's expressive potential and resultant meaning, and how our perception changes

³⁶ Link to live recording: <https://www.youtube.com/watch?v=ZJDC-gBb9rQ>

³⁷ Priming is a term used in psychology to explain how one stimulus affects the response to another stimulus. This was shown to work most clearly within one modality in the experiments of Meyer & Schvaneveldt (1972). In the context of this subsection, I use the idea in a cross-modal sense, to explain how a projected text can subconsciously influence our understanding of something in another medium, such as music.

depending on the type and amount of information it is fed. The piece takes the form of a 'dream' essay, where fragments of film, soundtrack, music and 'psychoacoustic' electronic sounds are combined in constantly shifting polyphonic textures, to produce an overall effect of disorientation.

My fascination with what I have termed music-text-film originally sprang from an idea about subliminal messaging, perhaps even out of a misconception of it: how split-second flashes of hardly perceptible text could interfere with the audience's experience of the music. Notwithstanding, the level of subliminal effect that I had first imagined was very difficult to achieve with the standard computer-video-beamer set-up; one would need a 'tachistoscope', a projector with the capability of opening and closing at shutter speeds of about 1/100th of a second or shorter. Furthermore, there was the lingering artistic issue of whether something lying so deep under the radar of consciousness could significantly influence our perceptions. In practice, I found it more interesting to work on the conscious level of visual text perception, dealing with the subconscious level through the combination of the visual and aural.

Accounts of the first use of subliminal techniques in advertising are part of the folklore of modern media history. In 1957 James Vicary, a market researcher, claimed that over a six-week period 45,699 patrons at a movie theatre in Fort Lee, New Jersey, were shown two advertising messages: 'Eat Popcorn' and 'Drink Coca-Cola' while watching the film *Picnic* (directed by Joshua Logan in 1955). According to Vicary, a message was flashed for 3/1000 of a second once every five seconds. The duration of the messages was so short that they were never consciously perceived. Despite the fact that the customers were not aware of perceiving the messages, Vicary claimed that over the six-week period the sales of popcorn in the theatre rose by 57.7% and the sales of Coca-Cola by 18.1% (O'Barr 2005: 4).

These claims later turned out to be false, when Vicary himself admitted that the figures had been invented as a marketing ploy. Nevertheless, the idea of the techniques he supposedly used caused nationwide unrest at the time. People became fearful of being susceptible to this kind of subliminal manipulation, especially with regards to the new media, such as television, that was beginning to permeate the social landscape. This fed into the general climate of paranoia, in part stoked by the political reality of the time, but also because of the social and technological changes that were taking place within society (Packard 1957).

The composition *Subliminal: The Lucretian Picnic* is split between four basic levels of media: two visual – text and image – and two aural – live ensemble and electronic soundtrack. The visual, cinematic material consists of samples from the film *Picnic*. This was a commercially and critically successful romantic drama starring William

Holden and Kim Novak, about a drifter arriving at a mid-western town on Labour Day and falling for a girl destined to marry a less charming local.

The manipulation of the source film material in *Subliminal* is three-fold. Firstly, in order to disorientate the conventional narrative reading, disengaging the diegetic causality, the narrative order is reversed so that it does not dominate. In the second scene, for instance, we see the hero descending from a moving train and running backwards to meet a woman whose hands he grips passionately. In fact, this is the penultimate scene of the film, where we see the lovers parting and the hero running to jump on the moving train. The second form of visual manipulation is the rhythmic editing marking the temporal flow of the footage. Frame rates are manipulated and made to oscillate between positions on the timeline, creating a visual rhythmic polyphony. The third and final form of manipulation, and the one possessing the most important visual impact, is the dimension in which the film is projected. The film is cropped to a ratio of 10:1, and is subsequently projected onto a 10m x 1m screen hanging in front of the ensemble, so that what is eventually visible is predominantly a framing of hands and feet, a narrative which is restricted to bodily appendages. The idea behind this was to undermine the completeness of the original material, in order to bring it into dialogue with the music and the text, while making a reference to the visual manifestation of subtitling or surtitling.

The 'Lucretian' reference in the title is to the text used throughout the work, taken from the Rolfe Humphries' translation of *De Rerum Natura* by the first century Roman poet and philosopher Lucretius. His text is used as commentary on what we see and hear. It alludes to sensory experience – how the world is perceived, the nature of dreams, emotions and thoughts – with a strong emphasis on an Epicurean philosophical ideology. This becomes a fitting answer to the sense of hysteria conjured by Vicary's experiment. At times the text addresses this head-on, with notions about collective fear, the nature of dreams, reality and identity:



Figure 12 Consecutive stills from *Subliminal* (at 23:20).

The type of text animation used in *Subliminal* varies enormously. There was a certain exploratory approach to the many rhythms and layouts used. Nevertheless, the central premise was the use of short fast-blinking text, which appears a number of times during the piece; for example in the first visual scene of the train and bus, where a text about perception appears, blinking at an on/off rate of 1:12 at the speed of 120 BPM:

In a single time, no longer than it takes to blink, our mouth to utter half a syllable, below this instant, this split second, lie times almost infinite, which reason knows as presences; and in each presence dwells its own peculiar image, all of them so tenuous that no mind is sharp enough to see them all. (Lucretius 1968: 141)

Later in the piece, two or three consecutive texts might appear at the same time, superimposed, in parallel, or scrolling quickly, with the intention to put stress or weight on the mental ability to process information in one medium against another. This was one of the principal strategies I explored in the piece: experimenting with the perceptual borders of text, image and sound, in order to investigate how much it is possible to assimilate these at any given moment when the perspective and interrelation between the media is always shifting. The viewer has to constantly reassess the perspective towards the media, changing focus in each new section.

In this respect, the text narration is an extra-diegetic commentary, with no explicit connection to the visual or aural information in the work: it is there to suggest a possible interpretation of the other media. In other words, we know that the text does not directly refer to the images or the music, but it is up to the viewer to make the connections and draw the conclusions. Because it is never clear where the narrative voice is located in the narrative – who is speaking – the audience is perhaps persuaded to switch back and forth between the media in search of it.

The music itself is divided into multiple layers, which come across in differing hierarchies: 1. instrumental music, which contains quotations from the film score; 2. various electronic and instrumental pulses, which are there to emphasise or disorientate visual rhythms; 3. quasi-subliminal use of voices from the film; 4. resonances of these appearing within the ensemble or through the electronic soundtrack.

The music develops in polyphonic blocks, switching between different points of view and only occasionally finding a sense of momentum, most notably towards the final parts of the piece. This formal approach, which is intended to frustrate the audience's ability to be immersed in the narrative, is akin to a '*Verfremdungseffekt*', the distancing effect associated with Bertold Brecht and found in many of the films of Jean Luc Godard, where text, image and sound are constantly obstructing one another (Monaco 2004: 136). In this case, there is no explicit voiced narration, and

what text there is functions only intermittently. We are constantly going back to the music or the visual in search of possible meaning.



Figure 13 Consecutive stills from *Subliminal* (at 17:36).

Looking at the work in terms of media hierarchy, one would have to conclude that melos is the primary intended medium. The work was created for a music festival and took place in a concert hall, and there is sufficient weight on the music and soundtrack as opposed to the cropped form of the projection. However, the metaphor hierarchies change throughout the piece, specifically because 'juncture' points are used throughout to realign our perception. There are some formal structures audible in the music at times – scales, pulses, drones and loops – which reinforce the 'target-ness' of the metaphor relation, but it is not always fully autonomous. At the points where one detects the music is changing almost at the same time as the image, and that the music is directly related to the image (for instance in the dance scene at 14:18), the hierarchy shifts to the image. Both narrating 'voice' and the physically heard voice are active in shifting the hierarchy, in the form of the projected texts, but also when one hears traces of voice, sampled from the film (4:50). The 'rate of information' also works on the level of text (2:08), image (26:08)

and sound (24:38), pushing each medium to the foreground at different points. Scale of information is also a feature, specifically as the piece explores the idea of 'subliminal' levels of perception: the media sometimes lie hidden within each other.

The question of media correlation in this work is complicated, because there are changing interactions between the three main media. To begin, we can draw up three radar charts, in the manner of the examples given previously, with the general overview of the three main interactions.

Music & Text Correlation:

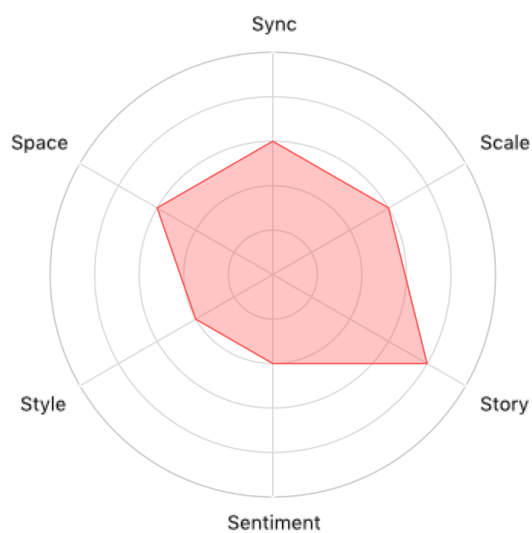


Figure 14 *Subliminal*: music & text correlation chart.

Sync: 3

Ranges from relative sync with the music (0:00), to relative rhythmic autonomy (12:44). Never extremely synchronised, but because of the shifting block form, only really correlates with structural changes in the music.

Space: 3

There are two distinct musical spaces: live musicians and quad soundtrack. The text seems to exist in a space between these as it is projected at the front of the stage.

Scale: 3

Text size relative to music volume doesn't change that much. There is not much fluctuation in this, and in general one could say that the presence of the projected text matches the presence and scale of the music and sound.

Style: 2

As opposed to the cinematic reference in the image, the text derives from a classical source. This re-contextualisation of the text to the modern electro-acoustic sound world is not an obvious match, though, because the language contains no reference to antiquity. There is also no jarring stylistic dissonance.

Story: 4

The general narrative of the text revolves around a commentary on the senses and the nature of perception, from an Epicurean point of view. One could say that this correlates to the reductive nature of the music; the way the music is sometimes stripped down to pulses and drones. Throughout the course of the piece, changes in music to illustrate the text are consciously composed.

Sentiment: 2

In terms of sentiment, and beyond the Epicurean philosophy, there are moments when the music leans towards the emotion evoked in the text: fear, dreaminess or awe. But in general the text remains at a distance to the music, creating a critical perspective.

Text & Image Correlation:

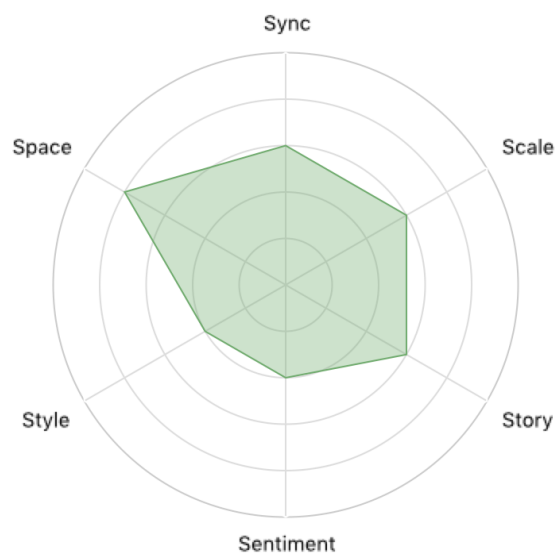


Figure 15 *Subliminal*: text & image correlation chart.

Sync: 3

There is generally no overt sync between text and image except for the structural change, though at times the rhythm of the image oscillations correlates to the text rhythm across different sections.

Space: 4

The projection space of the text and image is shared, though the text does not originate from the cinematic material.

Scale: 3

Scale does not vary enormously between text and image.

Style: 2

The style of the text here seems to be in starker contrast to the image than the more abstract nature of the music. The type of text subtitling does not cohere so strongly with the images represented.

Story: 3

In terms of narrative, there is a stronger need to understand the images in relation to the text. Even though one can glean that the text does not explain the narrative of the film, it offers a coherent commentary on it. Some parts of the text speak directly to the images shown. For instance, at 26:55, when we read 'no man starts to act before the mind foresees its will', in the following image we see a man walking backwards. This does not explain the image, but interferes with our understanding of it.

Sentiment: 2

The sentiment presented or invoked in the images seeping through the movie narrative revolve around desire, love, jealousy and primary emotions of joy and sadness,. This is not met in the same way with text, which remains detached throughout.

Image & Music Correlation:

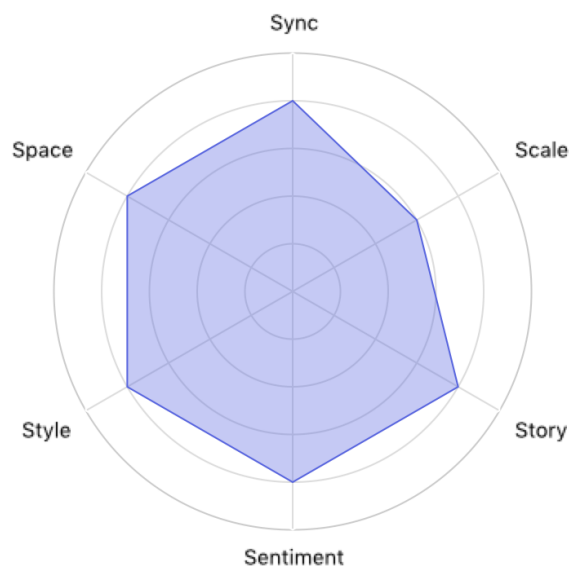


Figure 16 *Subliminal*: image & music correlation chart.

Sync: 4

Synchronisation in structural changes and often the timing of the oscillations in the image are in deliberate rhythmic relation to the sound.

Space: 4

The space of music and image cohere more than music and text. The music and sound make extensive use of samples and references to the original soundtrack of the film, pulling the sound into the diegetic space of the film.

Scale: 3

There is a greater variety of scale implied in the image and the sound as in relation to the text. Image can vary from close up of cropped body parts to crowd scenes. Likewise in the music, symphonic scale textures are heard next to whispered voices and close-miked single instruments. There is certainly contrast here between sections, but no overall bias.

Style: 4

Style here is more correlated than in the other relations: the soundtrack of the film is very present in the fabric of the music in the form of samples and melodic motifs played by the ensemble.

Story: 4

The narrative of the images has generally something of the same character that is shared in the music or the soundtrack. Partly because the sound is sometimes used directly from the film for that particular scene or the atmosphere of the scene matches the music in some way. The overall narrative implied by the image and sound is more complicated to perceive, because the structure is so fragmented, but within that fragmentation they correlate.

Sentiment: 4

There is more overt sentiment shared between image and music than with text, because they both contain emotive elements. There is a sense of playfulness in many scenes with the occasional shift to a cinematic relation between music and image, where music reinforces an emotional effect produced visually..

Together the three charts would line up in this way:

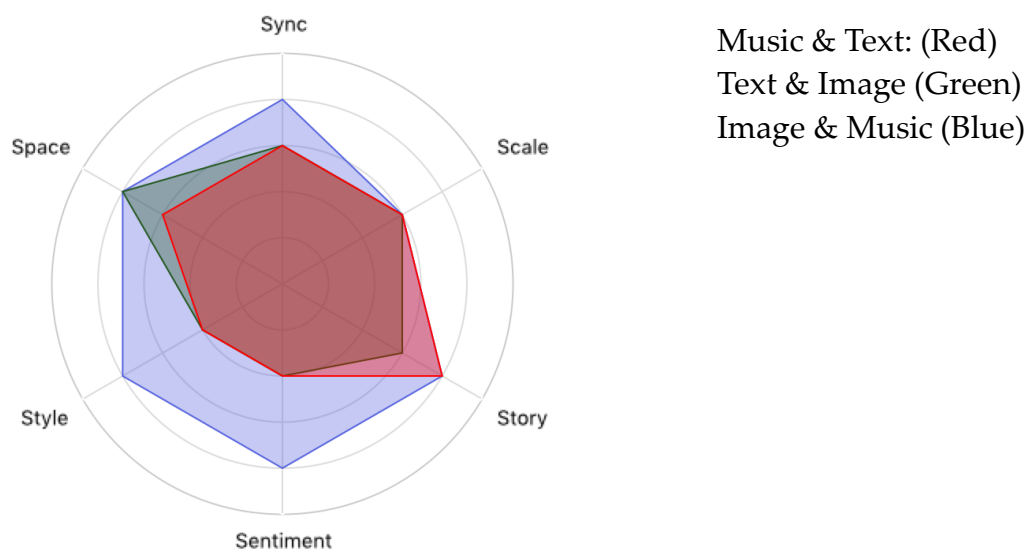


Figure 17 *Subliminal*: Three correlation charts overlaid.

What, then, can one say about the result of these three correlation tests? We might note that the relation between image and music is much stronger here than the relation between the text and other media. This is somewhat an exception in my oeuvre, as many of my other music-text-film pieces presented in this thesis correlate much more on the level of text to music than with image.³⁸ The strength of the music to image correlation comes from the fact that much of the music is drawn from the soundtrack of the film, either as samples or references to the orchestral soundtrack. The other reason is that the narrative voice of the text is very much 'extradiegetic', in Genette's terms: it is an outside-the-story narration, a detached commenting voice, which is not strongly reinforced in either music or image. In much of my work dealing with first person narratives, the relation of text to music is stronger, as one can read the narrational voice in the text through voices in the music, but not in this case. Although there are moments in which the music supports the text, framing it within an audio-visual context, mostly the narrative voice stays outside the diegesis of music and image. It acts as a distancing device, reinforcing the space between the other media. What is also interesting to note is which mode each relation favours. For instance, music-text correlation is more biased towards 'story', or concept, in this case, whereas in the text-image correlation, 'space' is more convergent, because text and image appear on the same surface.

³⁸ This is one of the reasons why I chose this piece for in-depth analysis, because the tripartite presence of media is very clear.

Another useful way of analysing this is by looking at each of the 64 sections and seeing which media are present and how the hierarchy shifts through the sections. Below is a table where the structure of the first fifteen parts is given in terms of media hierarchy, along with my justification of why one medium is targeted over another:

Table 2 Analysis of *Subliminal* in terms of hierarchy of media:

1	0:00	"no man..."	text - music	Text is primary, but there is enough space for the music to register because of the slowness of text refresh:- voices in the background take a while to gain attention.
2	1:00	(music only)	music	Speed of background voices from #1 are taken over in the ensemble: oscillation theme in music is later translated into the oscillations in image.
3	1:18	"in sleep when..."	music - text	Having established itself as primary medium, the music stays dominant (for a while at least): the end text references to 'hearing' and 'conversing' help establish the metaphor of how we can understand the music.
4	1:44	(music only)	music	Interchange between ensemble and electronic sounds.
5	2:08	fast text	text - music	Text takes over attention because of the 'rate of information' rule: harp punctuates and helps chunk the text.
6	2:28	train / bus aerial shot	image - text - music	Train siren at the beginning helps reinforce the primacy of the image. Dual function of music: electronic sound underpins image while ensemble sound connects with the speed of the text. The focus on the text becomes more prominent as the section evolves because of the need to understand the meaning under the difficult circumstances of intermittent blinking.
7	2:58	"all must focus..."	text - music	Text takes prominence over the music – especially because the pulse in the music seems to support the rhythm of the text changes – and the text animation has a more complex expressive quality.
8	3:14	train arriving	image - music	Space opens up in the music for the image to take over: the image is not immediately intelligible so that it takes more attention to process. Image oscillation speeds up to next cue.
9	3:54	static on porch	image - music	Although the image is almost static and the ensemble has taken over the oscillations - the image remains dominant - partly because we are still trying to process what has happened before and partly because it remains an enigmatic image.
10	4:06	"in our dreams..."	music - text	Continuity of the music helps establish its hierarchy, along with the slowness of the text fades. Switch between left and right columns of text reinforced in music.
11	4:34	(music with voices)	music	Voices emerge from the music and take our complete attention.

12	5:14	porch / "children fearing"	image - text - music	Much like #6, image starts as the dominant medium but our attention sways towards the text, again because of the effort of reading. The music reinforces both the rhythm of the text and video oscillations. The reference to "children fearing" in the text influences how we understand the image.
13	5:38	cleaning boots	image - music	The image though starting as the primary medium might be said to give way to the music because its fast rhythmic oscillations are experienced as an aspect of the music. Slow chord fades help shift towards music.
14	6:18	(music only)	music	Music is prominent but there are multiple layers going on at different speeds: sine drone, pulse, harp/piano chords, background sampled voices. Attention shifts between these.
15	6:50	police/ "sweet it is.."	image - text - music	Rhythmic oscillations in similar speeds between image and text take the attention. Harp chords reinforce structure of syntax.

This reading of the piece is by default highly subjective, as it betrays my own intentions as a composer, as well as what I imagine the effect on the spectator might be. Nevertheless, it gives an insight into how I see the attention shifting within the sections. It should be noted that the described states are not static, as attention sometimes moves between the layers of media repeatedly within one section. What is interesting in this work is that, because of the extreme structural changes, there does not appear to be one constant dominant medium. The focus is ever shifting between image, text and music, and regardless of an individual's subjective viewpoint, different metaphoric relations have the potential to develop.

Chapter 4: Historical Perspective

The use of projected text in cinema, art, music and various hybrid forms of literature makes an intriguing footnote to twentieth century art practice. This is often associated with an art that is, in the first place, non-absolute, narrative, and exhibits hybrid tendencies that encompass both experimental and commercial cultures. It transgresses the purity of Greenberg's medium-specific modernism that has dominated much artistic practice in the last century, and about which there has been fierce debate in cinema, music, theatre and literature. In cinema, where projected text first appeared in the form of 'intertitles', the debate raged in the mid 1910s and continued into the 1920s, as to whether text should have a place in film, or whether narration should be left solely to the power of the image. The development of sound film eventually put an end to this discussion, as hearing words took over the role of reading words. On the other hand, in twentieth century visual art, language has played a vital role, from inscriptions on paintings of the impressionists, Dada and conceptual art, to new media and hypertext, with each movement redefining its relation to the written word.

In the following chapter I attempt to trace a history of text-films or text-films touching on some of the main art practices they are found in. The sections are organised in terms of metaphoric relations between the two dominant media. This I found useful in examining how various practices and conceptions of media can lead to such different outcomes, though I do not want to make it seem like this is an absolute form of categorisation. The beauty of most of this work, is that it does not easily fall into one particular established art practice, and as such pushes on the boundaries of a given medium. Because it is organised utilizing the point of view of the most dominant medium; the context that the art is made, as discussed in the previous chapter, plays an important defining role. In this way, where image is the dominant medium, the work tends to be created by artists operating in a visual or cinematic context. Where music is the dominant medium and where language or text are the source of the meaning by which the music is understood, one finds the most examples that fall into contemporary music practice, as sound is here prioritised over language and image.

I begin with the section 'Image as Language', because it is this relation, which in my opinion defines the earliest and most iconic example of text-film, Marcel Duchamp's *Anemic Cinema*. Within this category one could also discuss the practice of film titling and the many ingenious examples of how text is used to construct image. Because these latter examples tend not to be autonomous entities, but are overshadowed by the qualities of the film they are coupled to, I thought it unnecessary to single any out.

In the section of 'Music as Language', there are many fascinating examples to choose from, but I focus on Dick Raaijmakers' *Ballade Erbkönig* because of its juxtaposition of seemingly divergent elements.

In 'Language as Music' I discuss two influential films created by the founders of the Lettrist avant-garde movement, Isidore Isou and Maurice Lemaître. The work of the Lettrists is far ranging, and has been highly influential on subsequent waves of art experimentalists, from the Situationists to American abstract films of the 60's. What is interesting in this context of their work, is the theory of *cinéma disrépant*, the divergence and independence of all the elements of cinema, and how initially the focus on the level of the letter and the removal of semantic elements of language led to sound poetry and their particular idiosyncratic music.

The Lettrists could just as well be discussed in the section on 'Language as Image', but instead, here I bring up two films that are personal favourites of mine, created by artists associated with the so-called Structural film movement, Hollis Frampton, and Michael Snow. In both these works, but in different ways, the word or language as a structure is used to understand something about the temporal framework of the cinematic medium.

The music video would be a natural subject to discuss in 'Music as Image', as image is used at the service of projecting or understanding something about the song it represents. Instead of taking one particular pop video as an example, I thought it would be interesting to analyse sections of film-maker Adam Curtis' *It Felt Like a Kiss*, because of the complex interaction between narrative, image, song and lyrics. Essentially here is a 'jukebox' film, a playlist of songs from the 60's, that are contextualised in a complex narrative, using stock footage from the BBC library, reminiscent in style to Guy Debord's film *La Société du Spectacle*. Unlike in most of Curtis' films, in *It Felt Like a Kiss*, there is no spoken narration, and songs and images are underscored with an on-screen text. Music here is very much in the foreground, as the narrative deals in part with the singers of the songs used.

Finally in the section 'Image as Music', I discuss two films that could be said to belong to an abstract film culture, Stan Brakhage's *I Dreaming* and Guy Sherwin's *Newsprint*. In the former, a song is used as the basis of a melancholic visual contemplation on memory and consciousness; in the latter, the image of newsprint is sonified, so that one understands the structure of the image through the pattern of what one hears.

4.1 Image as Language

The first paradigm concerns the use of words to construct an image. In this metaphoric relation between image and language, the attributes of language are used primarily in a visual representation. Sound or music is in this case relegated to a supporting or non-existent role. The visual medium takes dominance, and words are used primarily for their iconic status rather than the meaning they might convey. This is not to say that in the examples given, meaning is totally absent, rather there is a redundancy of meaning. A primary example of this is film title design. Here the text is informative but not essential in understanding the film. It might be useful to know who the principal creators of the film are, but it is often more of a vehicle for expressing the visual identity of the ensuing feature. The visual message conveyed has an impact on how the film as a whole is understood. Rather than analyse a specific case of film titling, I prefer to look at an example from the visual art avant-garde, that not only represents this approach to 'image as language', but is probably the father of all text-films, Marcel Duchamp's *Anemic Cinema* (1926).

Signed under his alter ego name of Rose Sélavy, *Anemic Cinema* (1926) is a 7 minute film which alternates sequences of spinning 'rotoreliefs' with sequences of erotic puns rendered in a similar spiralling form. The rotoreliefs were optical works that Duchamp had been developing since the early 1920's, initially with Man Ray.³⁹ They were essentially painted concentric circles on cardboard disks, that when spun on a turntable, gave the illusion of three dimensional movement. By combining and juxtaposing this illusion of movement and space with text, he reduces the elements of silent film to the play of movement and language.

4.1.1 Intertitles

To understand the context of words in film of the 1920's, it is useful to see it in the light of the use of intertitles. The alteration of rotoreliefs with spiralling texts exactly mimics the dual state of mainstream film of the time, that of image and word. Early cinema required narrative to make sense of images, since as a vehicle of mass-entertainment, the cinematic experience is essentially one of stories told in images. Before sound film, intertitles were used to convey both narrative context and any dialogue deemed necessary to the understanding of the storyline. In this sense, the intertitle was something of a transitional convention. Essentially, intertitles or 'title-cards' were printed texts edited into the film sequence, which would either carry dialogue, give some background narrative information related to the images, or just describe what was happening. 'Dialogue intertitles' would appear just after the

³⁹ A catalogue of the work inspired by their friendship is published by the Sean Kelly Gallery: *Marcel Duchamp/Man Ray: 50 Years of Alchemy*, 2005.

actor's lips were seen moving, while the 'expository intertitles', which would set the context of the narrative, would appear at the beginning of scenes. They would differ in appearance in the emphasis of the lettering, in the framing of the title card, or with the use of quotations. It is interesting to note, that with the advent of dialogue in sound films, and the logical disappearance of dialogue intertitles, the falling away of expository intertitles was not so expected. This seems to imply that films relied more heavily on signposting the narrative than was in fact necessary (Chisholm 1987: 137).

The general development of intertitle use can be seen as moving from externally driven narratives to internal driven ones of diegetic texts and character dialogues, the idea being, that a closer relation to the character's point of view, would lead to a more emotional experience of the narrative. Tension in the 1910's arose between the two schools of thought: one that advocated the intelligent use of intertitles for the sake of more complex narratives, and the other, brought forward by the 'purists', who resented the use of language in film:

[The spectator] demands that you give a play by pictured action only, and resents your impudence in offering him text, and the insult to his intelligence it implies; but the insult is really directed against your own technique – or lack of it – in having to resort to any medium of interpretation other than the pictures themselves.

(Dimick 1915: 17)

Eventually, the purists won the argument, even though this took a seismic shift in the technology. This same debate amongst avant-garde film makers of the time, seemed to manifest that same suspicion of the word against the new medium of moving image. As Frampton writes in his 1981 essay: "Film in the House of the Word", expelling Eisenstein's aversion to language:

Language was suspect as the defender of illusion, and both must be purged together, in the interest of a dematerialisation of a tradition besieged by the superior illusions of photography. (Frampton 1981: 83)

Filmmakers wishing to banish word from their works in the 20's included Germaine Dulac (*The Seashell and The Clergyman*), René Clair and Francis Picabia (*Entr'acte*). On the side of the artists, who embraced language in their work, stand two important works, which celebrate the use of text and undermine it at the same time: Marcel Duchamp's *Anemic Cinema* (1926), and Luis Buñuel and Salvador Dalí's *Un Chien Andalou* (1929). In *Un Chien Andalou* the filmmakers subvert one of the conventional functions of intertitles, i.e. to create a sense of temporality, to understand how a sequence of events fits into a chronology and to account for the passage of time. These are the film's intertitles: "Once upon a time - eight years later - around three in the morning - sixteen years ago - in spring."



Figure 18 Two intertitles from Luis Buñuel and Salvador Dalí's *Un Chien Andalou* (1929).

The mixing-up of chronology, the use of time markers that carry radically different scales of temporality, is part of the strategy of the film makers, used to highlight the irrational nature of the events shown, to break from reality and to enter the surreal logic of the dream.

4.1.2 Anemic Cinema

On the other extreme Marcel Duchamp's *Anemic Cinema* highlights and subverts the very aspects of the illusion of film. The idea of ready-made, as in much of his work, also permeates Duchamp's use of language. There is the sense that colloquialism and slang is taken from street context, and realigned on semantic and consonant levels, in order to create a fleeting sense of curious connotation.

Earlier, in 1916, Duchamp had already experimented with a set of postcards, *Rendezvous du Dimanche*, where the text was built on meaningful syntax with meaningless words. This deconstruction of language comes close to a Chomsky's "Colourless green ideas sleep furiously" (Chomsky 1957: 15), a sentence that is syntactically correct but makes no semantic sense. It is in this way that Duchamp moves towards the redundancy of language, as if he suggests, that language cannot be understood in a linear sense but as a visual spiral. He both celebrates language and at the same time points to its superfluity.

In *Anemic Cinema*, Duchamp elevates these language games and deconstructions to the level of pun. Because a pun highlights the phonetic aspect of language, one could argue that it questions its very semantic validity; there is an inherent sense, that meaning which is being conjured is accidental, and a by-product of the phonetic consonance of the words. According to art critic Katrina Martin, in her essay on Duchamp's use of language:

A pun itself is already a subtle comment on the function of language, where the consonant arrangement of the words suggests an infinity of potential meanings and at the same time mocks any conclusive definition. By dealing with language as an overall, concrete phonetic entity, the pun questions the value of language as an abstract metaphor. (Martin 1975: 53)

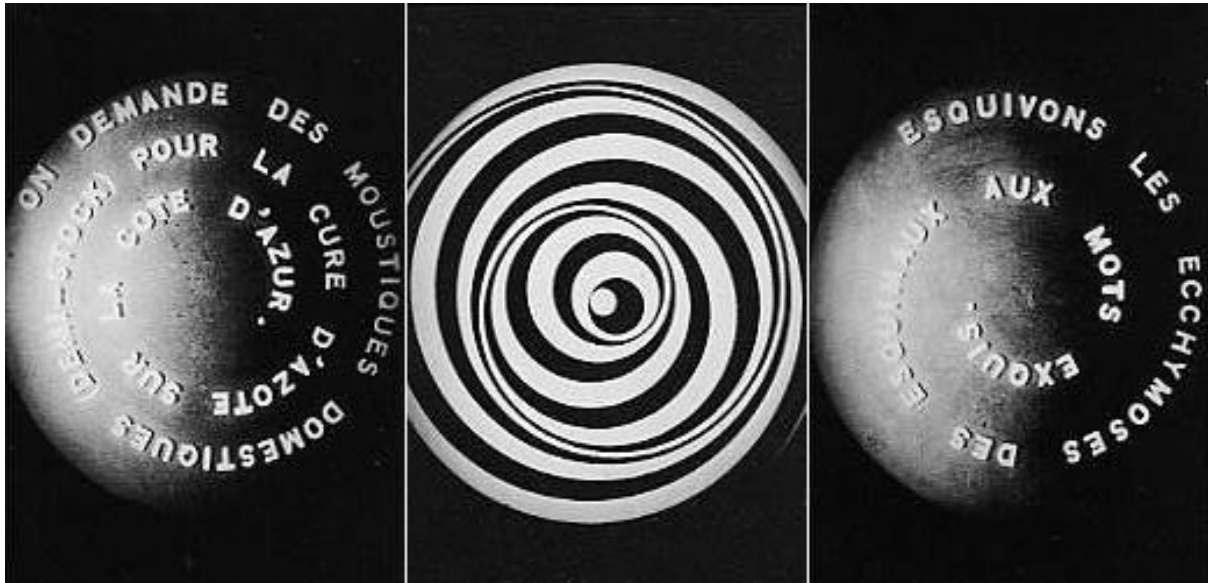


Figure 19 The juxtaposition of spiralling text and roto-reliefs in Duchamp's *Anemic Cinema* (1926).

Duchamp is here deliberately highlighting ambivalence and contradiction in the fabric of the piece. Just as there is a highly subjective experience of the roto-relief parts of the film, the perception and experience of the language part is largely dependant on the viewer's ability to decode the text; and not only is there insufficient time to do this through the course of the film, but the resulting proliferation of semantic associations cannot be pinned down to any fixed meaning. Taking what is probably the most straight-forward of the nine sentences of *Anemic Cinema*: "Si je te donne un sou, me donneras-tu une paire de ciseaux?" This is an image that found in his seminal work: *Le Grand Verre* (*Large Glass*), or *La mariée mise à nu par ses célibataires, même* (*The Bride Stripped Bare by Her Bachelors, Even*), where a pair of scissors and a bayonet could be alluding to the sexual act. The 'pair' of scissors, which is more of an Anglicisation of the singular French scissors (*un ciseau*), would according to Martin be referring to a pair of thighs, so the translation could read: "If I give you a penny, will you give me a fuck?" Nevertheless, even considering this relatively straight forward pun, there is still ambiguity to be found in the image itself. A pair of scissors can allude to a violent, almost castratory metaphor of sexuality, and along with the suggestion of a financial transaction (or is 'penny' an anglicised pun of 'penis'), the lasting image that remains in the viewer's imagination is both suggestive and perplexing as it is incomplete.

It is also fascinating about the use of text that, because it is constructed on puns and alliteration, it forces the viewer to not only read but to speak the text with their inner voice in order to understand the hidden meaning. In a sense, it is an abstraction on two levels, using as its material, references to a very corporeal sense of erotic imagery. This further highlights the sense of the absent mimetic representation, that has always been considered to be the norm of cinematic language. Film theorist Bart Testa, commenting on this effect that the film generates:

On another, let's call it a phenomenal level, the combined reading and viewing of silent films conventionally give rise to a third activity: our imaginative conjuration of a domain with all the space and furniture of a world. It is what film semioticians term diegesis. *Anemic Cinema* exposes, by its reduction, this third and paradoxically maximizing activity: our imaginary production of diegesis, which can still happen in *Anemic Cinema*. And the film does this, amazingly enough, by dismissing mimesis. (Testa 2002)

As there is hardly any time to process the meaning of the phrases, before the next rotorelief wipes away any semantic trace of the text, and plunges the viewer back into the visual vertigo, ultimately the language is subverted and surrendered to the hypnotic nature of the rotorelief image.

4.1.3 *Television Delivers People*

One can trace Duchamp's legacy in the use of on-screen text in the visual art world to artist Jenny Holzer and beyond. Specifically Holzer's famous use of text with LED media, which has an entrenched commercial rather than artistic use. The collision of pop art with sculptural minimalism, and the emergence of video saw many artists experimenting with on-screen language in the early 1970's. Another iconic example of the time, is Richard Serra's *Television Delivers People* (1973).

This is not a work where the visual carries weight in terms of beauty of design, but by the text's constant reference to the visual media's shortcomings. A gently scrolling text, in the style of television end credits, directly critiques mass media and the corporate control they exercise on society. This is accompanied by the lilting tones of 'muzak' in the background. There is disjunction between the content of the text and the image created by the medium, and the manner by which television is transmitting it. In this sense, the image of the TV is the one that is in the foreground, and it is being critiqued by the very medium that it is drawing meaning from - the scrolling words.

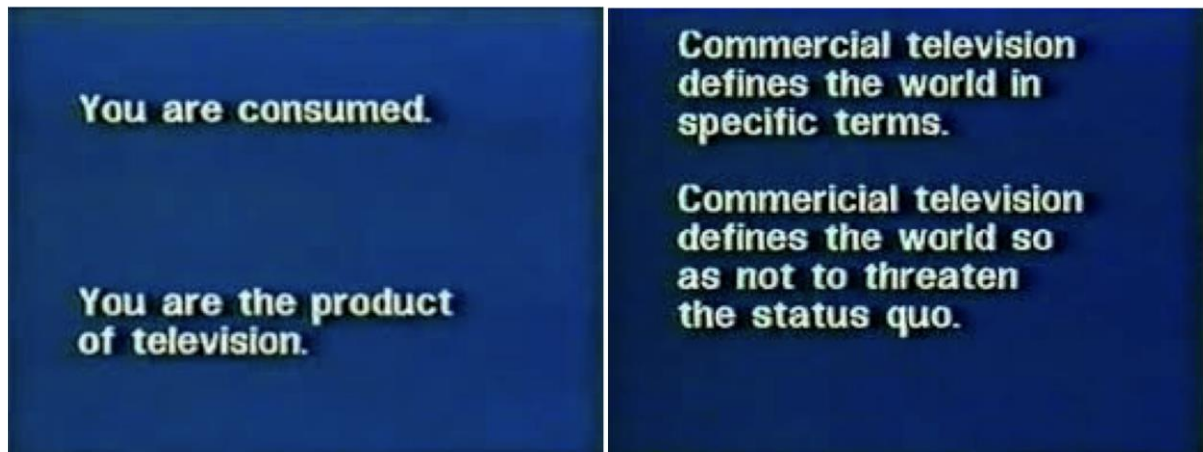


Figure 20 Stills from Richard Serra's *Television Delivers People* (1973).

4.2 Language as Image

Once we can read, and a word is put before us, we cannot not read it.
(Frampton in MacDonald 1988: 49)

Whereas the previous chapter dealt with examples of film and video, which use language or words in order to construct an image, where that image could be argued to take precedence over the meaning of the text, in this chapter I will examine two examples where the reverse dynamic is valid, i.e. where image, or on this case, the film medium, is used to shed light on an aspect of language. Both of these examples originate from what is known as the structural film movement in the US from the early 1970's. This movement, which was not unrelated to minimalism in art and music, includes artists such as Michael Snow, Hollis Frampton, Paul Sharits, Tony Conrad and Peter Kubelka. The movement is characterised by work dealing with the materiality of film. Peter Gidal defines the work in this way:

Structural/ Materialist film attempts to be non-illusionist. The process of the film's making deals with devices that result in demystification or attempted demystification of the film process. (Gidal 1976: 1)

Many film works associated with this movement make use of the material of film as the subject: the camera, film stock, light, sound, time. These make up some of the underlying concepts, which are examined through the form of the film. This results in many works using flickering, looping, mechanical camera movement and other such processes, which highlight the material nature of film. On the other hand, what is also apparent in this genre of work, is that the act of perception, as executed by the viewer, is of central concern. In the same essay, Gidal continues to suggest that:

The mental activation of the viewer is necessary for the procedure of the film's existence. Each film is not only structural but also structuring. ...The viewer is forming an equal and possibly more or less opposite 'film' in her/his head, constantly anticipating, correcting, re-correcting - constantly intervening in the arena of confrontation with the given reality. (Gidal 1976: 1)

This self-reflexiveness is something apparent in both the works which will be discussed: Hollis Frampton's *Zorn's Lemma* (1970) and Michael Snow's *So Is This* (1982). The relation of image to language in both these works is very different, yet they both locate it as a central concern in their films, and problematize its function in a cinematic or visual context. In the case of *Zorn's Lemma*, the idea of language itself is in the foreground, and image is used to systematically dissolve its signification. In the case of *So Is This*, a self-reflexive text highlights the experience of reading the very text that is projected.

4.2.1 *Zorn's Lemma*



Figure 21 A tapestry of images in alphabetic order from the second part of Frampton's *Zorn's Lemma*.

It is no coincidence that Frampton came to make films, which seem to be suspended inside a no-man's land between words and still images. His first passion was poetry. Early on, he struck up a friendship with Ezra Pound, when the ageing poet was hospitalised in Washington D.C. After his move to New York in the late 1950's, while rooming with his old school friends Frank Stella and Carl Andre, photography became his primary medium. This resulted in a series of *Word Pictures* (1962-63),

where the idea of photographing words in an urban environment first surfaced. The power of words used to infect the visual medium, was a theme in his filmic work throughout his career. From *Surface Tension* (1968), where in the third part, text to a hypothetical film is superimposed on an image of a goldfish tank and waves, through *Poetic Justice* (1972), again a hypothetical film script, this time filmed page by page next to a plant and a cup of coffee, to *Gloria* (1979), part of his epic, uncompleted *Magellan* cycle, which uses computer generated text in the form of sixteen propositions about his maternal grandmother. The fact that he came to be associated with the structural film movement, was something he did not entirely agree with. In a 1976 talk, he seems to have raised an objection to film historian P. Adams Sitney's term 'Structural Film', stating that classifications like that 'render the work invisible' (Windhausen 2004: 76). Certainly his use of text in film was motivated by ideas beyond the strict ideological standpoint of the structural film theorists. An eloquent writer himself, he argues in a 1981 essay "Film in the House of the Word", that cinema, ever since the invention of sound film, has developed a deep suspicion of the printed word:

Every artistic dialogue that concludes in a decision to ostracize the word is disingenuous to the degree that it succeeds in concealing from itself its fear of the word...and the source of that fear: that language, in every culture, and before it may become an arena of discourse, is, above all, an expanding arena of power, claiming for itself and its wielders all that it can seize, and relinquishing nothing. (Frampton 1983: 83)

Tension between word and image is exactly what is played out in *Zorn's Lemma*, and for all its formal construction, the piece reveals the very personal nature of this dilemma in Frampton's work. The work can be seen in two distinct ways, as a metaphor of language acquisition, and at the same time as a metaphor of film making itself. There is a gradual development in the film from the word to the image, and back again, which takes the viewer through a quasi recalibration of one's language capacity.

The film is constructed in three distinct parts, much like his previous work, *Surface Tension*, where the media relationships are differently configured in each one. In the first very short part, over a blank screen, a woman's voice reads from an alphabet primer used in schools in the early part of the 20th century:

In Adam's Fall / We sinned all.
Thy life to mend / God's Book attend (Bay State Primer).

This is to establish the authority of language, both in a strict Judaeo-Christian framework and to expose the idea of the alphabetical order which is used to structure the images in the subsequent part, the forty-minute main bulk of the film.

In this soundless part of the film, 1-second shots of words in an urban environment are presented in alphabetic groups of 24 (i/j and u/v are combined as in the Roman alphabet). The fact that in this part of the film, image is substituted for sound, underlines the subject of the alphabetic order, that was experienced only aurally in the first part. Interestingly, as this second part initially unfolds, it takes a minute or two before the structuring principle of the alphabetical order establishes itself in one's mind. At first it seems as if the image itself has dominance over our perception, we examine the space of where the words appear, understanding the context of the words becomes paramount: buildings, windows, store fronts, signs, fabrics, graffiti. The relation of word to context is the main source of meaning. Gradually as the relentless alphabetical order persists, environmental context seems to become less relevant. The viewer shifts attention to that which gives a clearer structure, the alphabetical order. This is a dynamic that relates to metaphor hierarchy (explained in the previous chapter), showing that when a part establishes its own inner structure, it becomes the primary focus (the target), and does not so readily lend its attributes to the other. Thus, the initial balance of image and word in the second part of *Zorn's Lemma* shifts from relative equilibrium to a state where the alphabetic order takes strong precedence in how we perceive the words, and image becomes less relevant.⁴⁰ Then another shift occurs. As the set of words in each letter subset are used up, Frampton substitutes each of these 1-second slots with a set of wordless images, actions, which seem to have another metaphorical meaning. The first letter to be used up is 'x'. This is replaced with the image of a bonfire, so that whenever the letter 'x' comes up in the alphabetical order, we no longer see a word beginning with 'x' but a continuation of the film of the bonfire. Gradually after more than 100 cycles of the alphabet, all word images are replaced with action images: turning pages, frying an egg, washing hands, grinding meat, painting a wall, digging a hole, hands tying shoes, changing a tire. This gradual erosion of the alphabetical order that was set up in the mind has a profound effect. Frampton describes it as a 'long dissolve', a dissolve not only in the filmic sense but from one mode of perception to another. In an autobiographical context he writes in his notes:

'my adolescence & early childhood were concerned primarily with words & verbal values. I fancied myself a poet; studied living & dead languages - hence my early contacts with, for instance, Ezra Pound... That 13 years in New York saw a gradual weaning away of my consciousness from verbal to visual interests. I saw this as both expansion & shift... That I began, during the making of the film, to think about leaving the city. Part III is prophetic, in that sense, by about 5 months.' (Hollis Frampton, handwritten notes reproduced in MacDonald 1995: 58)

⁴⁰ I have watched this film countless times, because I show it to students in my multimedia class. From my own experience and from comments of the class, I can vouch for this shift in perception, even after repeated viewings.

Finally the third part, alluded here by Frampton as 'prophetic', is a single shot of a man, a woman and a dog in a snow covered landscape, moving away from the camera to a forest in the distance. The authorial voice returns, combined with the 1-second rhythm transformed from a visual cut to a sonic cut, though this time multiple voices are hocketing a philosophical text under the direction of what seems to be a metronome at 60 beats per minute. Even under normal circumstance, the text would be very difficult to comprehend: A medieval text, *On Light, or the Ingression of Forms* by Robert Grosseteste, Bishop of Lincoln as translated by Frampton, is a treaty on the universe, that speaks of form, matter, composition and entirety. But in this context, it becomes even harder to make any semantic sense of it, due to its fragmentation between the six voices. The relationship between image and language (carried in speech now rather than in text) shifts again. There is a wish to understand what is being spoken, a re-igniting of the semantic power of words (over the logic in the alphabetical order), but since the meaning remains difficult to penetrate, what stays prominent is the hypnotic rhythm of the hocketing. This reinforces the idea that Frampton sets up in the second part, that in language, just as in film, it is the temporal structure that creates the space of our perception. And just as there is an opposition set up in the image space between the second and third part, between the urban and the pastoral, there is a contrast in the text used between two different authorial uses of the word. In a written interview between Frampton and Gidal from the *Structural Film Anthology*, Frampton explains:

The key line in the text is a sentence that says, 'In the beginning of time, light drew out matter along with itself into a mass as great as the fabric of the world.' Which I take it is a fairly apt description of film, as the total historical function of film, not as an art medium but as this great kind of time capsule, and so forth. It was thinking on that which led me later to posit the universe as a vast film archive which contains nothing in itself and presumably somewhere in the middle, the undiscoverable centre of the whole matrix of film thoughts, an unfindable viewing room in which the great presence sits through eternity screening the infinite footage. (Frampton in Gidal 1976: 67)

This idea of a Borges-like *Library of Babel* in film comes close to unravelling the concept of *Zorn's Lemma* as a whole. The title refers to the mathematical principal named after mathematician Max Zorn, who in 1935 proved that 'every partially ordered set contains a maximal fully ordered subset' (Campbell 1978:77). What Frampton was trying to show, in his own words, was that the abstract subsets found in his film - all shots containing certain colours, the list-able aspect of words, the subsets of the fictive elements - are in fact not the maximal subset:

What you see (consciously) most of all is the 1-second cut, or pulse. So that what I imply, is that the maximal fully-ordered subset of all film (which this film proposes to mime) is not the "shot", but the CUT - the deliberate act of articulation. Beyond that, there is the pulse of 24 FPS which is truly the maximal fully-ordered subset of

all films-and, obliquely, of our perceptions, since that is the threshold at which they FAIL us. (Hollis Frampton, handwritten notes reproduced in MacDonald 1995:56)

What remains fascinating about this film in the context of the language and image relationship, is how the shifts of perception between the media are handled in such a conscious and poetic sense. The gradual disintegration from one order to another, is certainly something that was a concern to much minimal art practice of the time, but the shifts between the media, which Frampton accomplishes, is something quite unique. Thus, it is perhaps misleading to say that this film represents the category 'Language as Image', because the metaphoric relation is certainly not a static one. The work should not be seen in the sense that image, a picture, gives us an understanding of the meaning of the words, but rather that film, as a medium, is the source for understanding the idea of language as a whole. After all, language, not words, is the subject of this film, just as the material of cinema and not image, is the means through which he expresses this.

4.2.2 *So Is This*

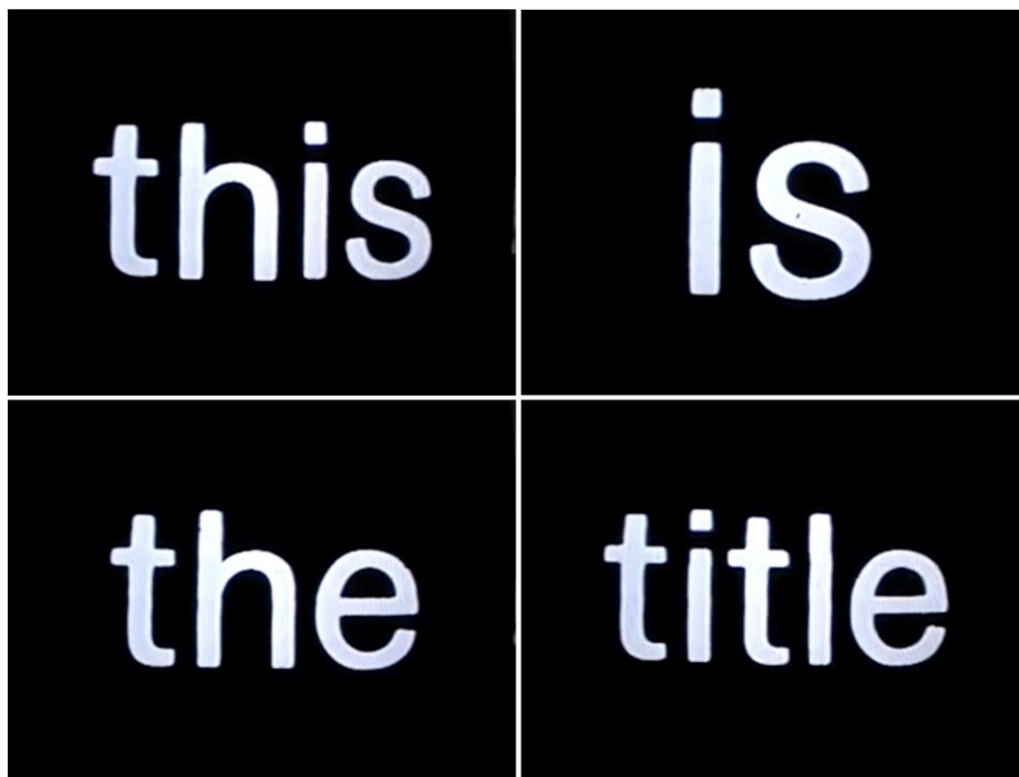


Figure 22 The first four words from Michael Snow's *So Is This*.

Michael Snow's *So Is this*, like *Zorn's Lemma*, uses the medium of cinema to reflect on the semantic power of language. In his film there are no images other than the text, projected one word at a time, centred on screen and with variable font sizes, so that

even short words fill the screen. Also similar to Frampton's film, time is here the crucial parameter, though in Snow's film it becomes the expressive vehicle rather than what is under scrutiny. The speed of delivery of each word is controlled in such a way as to manipulate the viewer's expectation of the text. Sometimes so slow that the reader starts to imagine different conclusions to the phrases; sometimes too fast to read, with even some supposedly single frame flashes of censor sensitive words. The way this reflects on the act of reading is reinforced by the fact, that the text itself is constantly referring to this. Snow takes the 'paratext' of the film (see Chapter 2.2) and makes it the central material of the work. It is a film that fully explains itself:

This is the title of the film. The rest of the film will look just like this. The film will consist of single words presented one after another to construct sentences and hopefully (this is where you come in) to convey meanings. This, as they say, is the signifier. (Opening text from *So Is This*)

Even though this is the first time Snow used text in a cinematic medium, it is not the first time he used text in a prominent position in his artwork. Originally a musician, he occasionally released records alongside his visual work. In 1975, his New York gallery released his record entitled *Musics for Piano, Whistling, Microphone and Tape Recorder*. On the cover and the other three sides of the gatefold sleeve, a text with a diminishing typeface, reflects on the conventions of album cover art.



Figure 23 Front and back record cover of Michael Snow's *Musics for Piano, Whistling, Microphone and Tape Recorder*.

The text of *So Is This* gently meanders between self-reflexive, factual, conversational, and provocative modes. After describing what it's going to be about, Snow suggests that the film will be two hours long, and then suggests he might be lying (the film is about 45 minutes long). He goes on to more factual information, something we

might find in programme notes of a concert, when the text was made, who the collaborators were, and what the title refers to. Subsequently he lists some artists who also made use of text in their work: Richard Serra, Tom Sherman, Su Friedrich, John Knight, and Paul Haines (interesting that he does not mention Marcel Duchamp or even Hollis Frampton and Paul Sharits). He suggests that the idea of using projected text is perhaps not an original one, but that it has much potential:

The author would like to have been first but it's too late. Priority is energy. In some respects this is the first. Obviously this is not the first time that this has been used for the first time. This belongs to everybody! This means this, you think this, we see this, they use this, this is a universe! So what is important is not this but how this is used.
(Text from *So Is This*)

The overt use of the word 'this', what is defined by linguists Roman Jakobson and Otto Jespersen as a 'shifter' word, a word which refers to itself, is not a coincidence. Because in this section of the film the words come at the viewer so slowly, Snow is playing with the constant postponement of meaning. One is constantly trying to refer backwards to what 'this' refers to, and forwards to the meaning it is heading to. Sentences are made and remade in the reader's mind, as each new word is revealed. The fact that smaller words tend to be shown in a larger typeface to fit the screen, also exaggerates their importance in the sentence, perhaps giving them more stress in the inner voice of the silent reading. Another aspect, which certainly effects that importance of some words over others, is the control of the duration of each word-slide - the rhythm:

The decision has been made to concentrate on the distinctive capacity of film to structure time: the word as the individual unit of writing, the frame as the smallest unit of film. In this film writing is the lighting. (Text from *So Is This*)

In the subsequent part of the film the rhythm of the words becomes more differentiated, especially in a hilarious passage about the role of censors, where he inserts the words: 'tits', 'ass', 'cock', and 'cunt' surreptitiously, as single frames, out of context in innocuous sentences. He highlights the sense of timing in a passage in which he repeats a sentence four times with different speeds, showing how subjective reading speed is, and how the film is actually controlling this for the viewer. The idea of finding a comfortable reading speed for everybody is mentioned in an ironic sense, because the film, for all its sense of familiarity it tries to foster with the viewer, deliberately employs irregular rhythms, both as semantic ruptures and as grammatical obstacles to problematize the act of reading on-screen text.

The most cinematic part of the film is the 'flashback' moment [31'15"- 35'15"]. A parody of the 'flashback' convention in film, where one sees or re-sees some material from the past, but from a distance. In this case, Snow re-films the slides in a smaller

frame and scans through them, sometimes at single frame speed, often pausing on the word 'this'. The colours are more saturated and sometimes over-exposed, adding to the sense of a 'tinted' past-tense. This is a very clear example of how the image or cinematic convention are used as a 'source' to understand the 'target' of text.

The film ends with a poignant quote from Plato, where he compares painting to text. This sums up the relation of the image to text, when it is but a reproduction of a lived experience.

You know Phaedrus that's the strange thing about writing which makes it truly analogous to painting. The painters' products stand before us as though they were alive, but if you question them they maintain a most majestic silence. (Plato quoted in *So Is This*)

4.3 Music as Language

Trying to find examples of text-film in the realm of music, where music is the primary focus, is not too difficult; there are plenty to choose from. This has to do with music's age-old connection with the word. In the many classic forms of text to music relationships: in the liturgical canon, in music theatre, in song, in so-called programme music, and many others; there is no stable hierarchy in how the music is understood through the word or vice versa, roles are constantly in flux. There are certainly examples where the power of the music and text is so perfectly poised, that calling one subservient to the other is misleading. Rather, in cases such as Franz Schubert's lieder settings of Goethe, Heine or Müller, it is exactly the movement between the perspective of poetry and music that is so compelling.

4.3.1 Surtitling & Music Video

When it comes to contemporary examples of projecting text with music, one cannot ignore the practice of supertitling (or 'surtitling') in opera, and the various off-shoots of this in performance practice. 'Surtitles' were first used (and the name was subsequently trademarked) by the Canadian Opera Company in their production of Richard Strauss' *Elektra* in 1983.⁴¹ Essentially, surtitling serves two possible functions. The first, to translate what is sung in a foreign language to the native language of the audience (rather than singing a translation of the text, as was the convention in some opera houses), functioning much like subtitles to a foreign film. The second, to reinforce what is being sung, in the same language, because what is being sung, for whatever reason, is not always easy to understand. Actually some of these reasons

⁴¹ <http://www.surtitles.com/intro.html> (accessed 12.01.2017)

have not only to do with the acoustics of the space, the technique of the singer, or the poor word setting of the composer; there are neurological studies on the recognition of sung text, that show how difficult it is to focus on both music and lyrics at the same time.⁴²



Figure 24 Stills from left: Bob Dylan's *Subterranean Blues* (video by D. A. Pennebaker) and right: Prince's *Sign o' the Times* (video by Bill Konersman).

In music videos the trope of projecting lyrics is not an uncommon practice. Some of the classic videos in this genre include Prince's *Sign o' the Times*, Bob Dylan's *Subterranean Blues*, or George Michael's *Praying for Time*. What these videos aim to do, and this general statement can also be applied to surtitles, is to reinforce the level of the lyric in a musical context, to highlight what is already there, rather than provide new narrative information. The difference between the music video examples and surtitling, is essentially that of the added value of visual expression. In all three examples, the way the text is presented adds something in terms of the rhythm, the style, the stress of certain words over others, or the discrepancies between what is sung and what is read. For example in Dylan's *Subterranean Blues* the timing of what is being shown by Dylan in the placards is at times out of sync with the text of the song. There are slight discrepancies between what is written and what is sung (the third verse starts to go increasingly out of sync with the placards).

In Prince's *Sign o' the Times* the animated text, even though it replicates exactly what is sung, manages to add weight to the words by underlining some over others; for instance, the refrain of 'time' is always accompanied by a slow floating animation of the word, as if to contrast the slow passage of time with the speed and madness of modern life. In contrast to these examples of text reinforcement, the practice of surtitling aims for neutrality, an inconspicuous presence. There is no visual expression whatsoever, in order not to draw attention away from the music or staging.

⁴² Neuroimaging and other neuropsychological evidence suggest that lyrics and melody in song are processed separately in the brain (Fritz 2013: 457).

4.3.2 The Cave



Figure 25 An image from a performance of Steve Reich and Beryl Corot's *The Cave* (NY Times 28.01.2005).

In contemporary music practice, there are many excellent examples, beyond surtitling, of this type of doubling of sung and projected text. Steve Reich and Beryl Corot's multimedia opera *The Cave* is an interesting case. This is a three act work based not on Plato's 'Allegory of the Cave', but on the The Cave of the Patriarchs, where the biblical Abraham is supposedly buried, a sacred site for Jews, Muslims and Christians. The piece interweaves religious texts with interviews of Israeli, Palestinian and American people about their view on the story of Abraham, and about fundamental questions of their faith. Although there is some independence of text in the 'typing music' sections, the projected text reinforces what is being heard, either sung, spoken or in musically transcribed speech. What makes this interesting is the polyphony of voices and texts weaving the fabric of the music. There is a space between what is spoken and what is sung, which Reich exploits in a very distinctive manner. On the one hand, the fragments of speech taken from the video interviews, are transcribed into the musical material. On the other hand, biblical text is quoted and rhythmically intoned by the chorus. These two worlds are constantly colliding with one another, the 'vox pops' gaining an epic quality by repetition and distillation into musical structure. The sacred text rendered and typed out as if it is a news item from a television network. One of the reasons why in my opinion, this works so effectively and has an emotional impact, is exactly because both class of text resonate and merge through our inner voices as they are projected on screen.

4.3.3 *A Letter from Schoenberg*

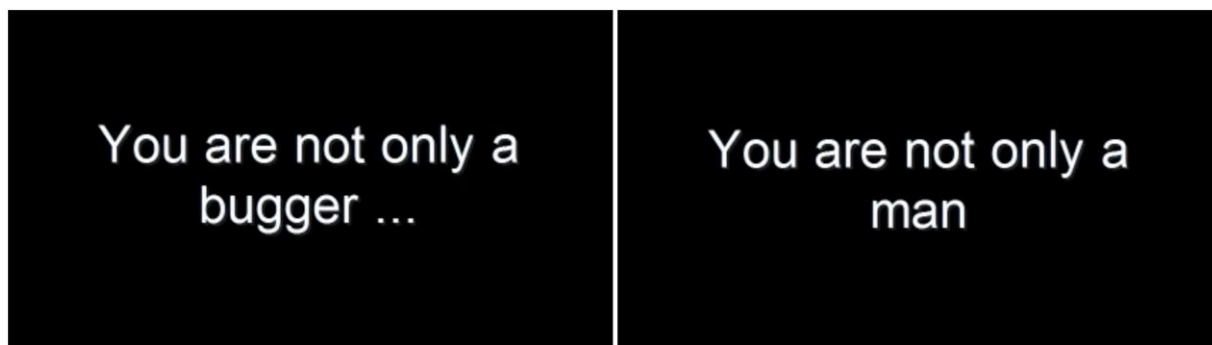


Figure 26 Screenshots from Peter Ablinger's *A Letter From Schoenberg*.

Another example is Peter Ablinger's *A Letter From Schoenberg*, which he calls a 'reading piece with player piano'. In this work, the voice of Arnold Schoenberg, dictating a letter of complaint, is spectrally transcribed onto a computer controlled piano, and played back at the same time that the text is projected. This is a relevant work in the context of this thesis, as it addresses the issue of voice, both in the transducing of speech to piano as well as in the content of the letter, Schoenberg's complaint, that the voice in his work *Ode to Napoleon* was re-transcribed in a performance from male to female. The question as to whether one can hear the voice in the spectral reworking to the player piano without the help of the projected text is an interesting one, because the piece underlines the limits of both media, trying to capture the essence of the original absent one, the voice recording itself. The voice is separated into its two constituent parts: the voice as pure semantic communication versus the unique identity and vocal expression of the individual. In my opinion, the two aspects of voice manage to remain in equilibrium, because neither medium is sufficient in conveying the original voice image alone. There is additionally an interesting irony that Schoenberg's complaint is based on a switch in gender of the voice, whereas the reproduction of the reading voice through the piano, supposedly Schoenberg himself, also remains ambiguous in terms of gender - the piano cannot really convey whether the text is spoken by a male or female reader.

4.3.4 Other Examples

Another remarkable example of text-film used in a musical context would be plunderphonics pioneer John Oswald's *Homonymy* (1998), dedicated to his friend and collaborator Michael Snow, as a musical homage to *So Is This*. This piece uses the concept of homonyms, same words that mean different things depending on context, merging a bilingual English-French vocabulary in a highly synchronised music to text composition. Yet another more recent example is Jennifer Walshe's *Everything is Important* (2016), for voice and string quartet, which uses imagery and text slogans

resonating with a post-internet art sensibility. Although music here plays a dominant role in bringing diverse elements together, reverberating through the immediacy of Walshe's own vocal performance, the relation of image to text and music is a fluctuating and dynamic one.

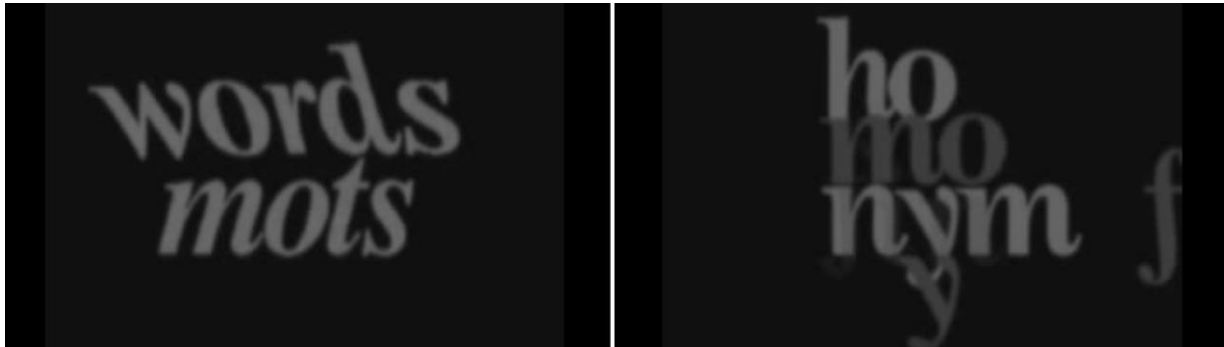


Figure 27 Screenshots from John Oswald's *Homonymy*.

4.3.5 *Ballade Erlkönig*

An example I would like to explore in more detail, partly because it predates many of the examples given above, and partly because of my own personal connection with the composer, is Dick Raaijmakers' 1967 work *Ballade Erlkönig*. This example of audio-visual work using text as the primary visual medium, alludes more to a tradition of programme music than to a reinforcement of text spoken or sung, even though the title and narrative refer to Franz Schubert's famous setting of Goethe's *Erlkönig*. The undercurrent of a voice not heard, which is at the centre of the poem, is exactly what the Raaijmakers version of the poem addresses.

Ballade Erlkönig consists of a collage of a huge variety of shortwave radio recordings, faint traces of transmissions from around the world, modulated voices, fragments of music, communication signals, interference, noise, which are layered and mixed into a dense sonic tapestry. Over an almost twenty-four minute structure slides of text are projected at specific time points (also indicated on the slide), like silent movie intertitles, that refer to the narrative of Goethe's *Erlkönig*. Specific points in the sonic landscape refer directly to the story of the father riding in the night with his sick child. The merging of these two narratives is made in an almost serendipitous manner, because it seems that the sound was mixed and structured before the idea of coupling it with the Goethe text came about. This gives both media a clear transparency. The visibility of the timing on the slides, reinforces Raaijmakers' wish to match up specific points of sound with narrative. In the liner notes of the recording of the work Raaijmakers writes:

After listening to this remarkable, spontaneous mixture of sound, the dramatic 'contour' revealed a remarkable similarity to Goethe's *Erlkönig*. Everything, including

the climax, all the verses, all the changes of states of mind, all the indications and interpretations of the father, the boy and the Erlking, all the natural sounds, in short everything which happens in the original ballad coincided with what was occurring on the tape. (Raaijmakers 1998)



Figure 28 The last 9 slides from Dick Raaijmakers' *Ballade Erlkönig*.

There is a certain mischievous tone in Raaijmakers' statement that "everything which happens [...] coincided with what was occurring on the tape", as the world of Goethe and the world of 1967 shortwave radio sounds are not natural companions. Analysing this from the aspect of media correlation which I outlined in subsection 3.2, one can point to the very weak convergence on all levels between the media. A forced narrative convergence is present because the listener is encouraged to understand the stream of audio in terms of the poem. A forced temporal synchronisation is also present, partly because the slides actually indicate the exact time of the tape.

While I discussed this with a long-time colleague and friend of Dick Raaijmakers, Gilius van Bergeijk, who knew him at the time this work was created, the latter suggested, that having created a mix of the piece throughout the summer of 1967, Raaijmakers was looking for ways to give the structure some meaning and

conceptual rigour, and stumbled upon the idea of coupling the sound with the narrative of the Goethe poem, almost as an afterthought.⁴³ As is often the case with the works of Raaijmakers, the structure is never suggested by the musical form alone, but by the concept that surrounds it. In this case, the work looks ahead to his music theatre pieces of the 90's, where the collision of different forms, narratives and media, give rise to an inimitable sense of a multidisciplinary composition.

Nevertheless, Raaijmakers has a justifiable argument in pointing to the congruence of the media, as the narrative of both sound and text coincide more deeply on various levels, the connection through the metaphor of communication resting at the heart of the material of both works. If there is one clear thing the radio signals in this work convey, is the desire to communicate over long distances, and the fragility of that communication, as we hear the signals fade in and out, become modulated beyond recognition, appear and disappear into a stream of noise and interference patterns. We hear the sound of the communication medium itself, the noise that both carries and destroys the message. Sometimes it is difficult to distinguish between signal and noise, as morse code fades into rhythmic noise patterns, voices become distorted into unrecognisable grains, and static resonances seem to move in perceptible melodic contours. This sense of imagined communication is a metaphor reinforced in the poem. The sick child believes that he hears the voice of the Erlking as they ride through the night:

"Mein Vater, mein Vater, und hörest du nicht, Was Erbkönig mir leise verspricht?" –
"Sei ruhig, bleibe ruhig, mein Kind; In dürren Blättern säuselt der Wind." ("My father, my father, and hearest you not, What the Erlking quietly promises me?"- "Be calm, stay calm, my child; Through dry leaves the wind is sighing.")

The discrepancy between the Father's and the son's reality is expressed through the act of listening. It is as if the son can detect the patterns in the nocturnal noise that the father cannot. It also suggests that some communication, which is openly audible to everyone, can only be understood by the person it is addressed to. For instance, at the moment in the piece [13'45] in which we read the slide: "'Willst, feiner Knabe, du mit mir gehn? ('Do you, fine boy, want to go with me?')", when the voice of the Erlking speaking to the child. We hear the voice of a German so-called 'number station', most likely a broadcast from the then East German secret service. 'Number stations' are agency-to-agent broadcasts of numbers, morse signals or melodic codes by government secret services, that can only be decoded using one-time random pads. Although we can clearly hear the communication, we have no idea of what is being said. Is the listener here being leveled with the perspective of the father, who hears the noise but cannot read any meaning into it? The son, in his near-death state,

⁴³ Personal conversation with Gilius van Bergeijk after a performance of *Ballade Erbkönig* at Dag in de Branding, Den Haag 13.12.2014.

can discern the patterns of communication in the sound of the wind, his fever-induced auditory pareidolia⁴⁴ acts as a window to the spirit world. Raaijmakers is perhaps showing us the same perspective, his slides project a possible reading of the shortwave radio noise-scape.

On a dramaturgical level, Raaijmakers underlines the shared "dramatic contour" of both media in the climax of the work. At the exact moment that the child says:

"Mein Vater, mein Vater, jetzt faßt er mich an! Erbkönig hat mir ein Leids getan!"
("My father, my father, he's touching me now! The Erlking has done me harm!")

explosive sounds of warfare and gunshots are heard; and immediately afterwards a voice speaking in Hebrew. This is most probably a transmission that is connected with what was one of the major political events of the Summer of 1967, the so called Six-Day War fought between Israel and an alliance of Arab states. Was there an underlying political message intended here, in aligning this text to this fragment of audio? Perhaps only a desire to expose a measure of contemporary reality, using current news broadcast in the media of the day, as a way of highlighting senseless loss of young life, rather than suggesting any partisan political reading of the work.

Another aspect conveyed in the form of the text slides and reflected in the sound world, is a sense of incompleteness. The slides do not contain the complete Goethe poem (see slides above). They refer to it, but like the half hidden shreds of the sound collage, they give birth to a sense of fragmentation. This is the overriding impression that the piece creates, the idea of incomplete or misunderstood communication appearing and disappearing into the aether. Raaijmakers gives a more programmatic interpretation of the narrative issues involved:

After the clouds of dust which the father has raised in the land without language and music have settled a little, after the Erlking's kingdom with its fairies, draperies, ghosts and illusions disintegrated before the eyes of the child, his ally, because of what we would now call neglected influenza, the singing in the ether continues, relentlessly confirming the grey hopelessness of life. (Raaijmakers 1998)

Like many of Dick Raaijmakers' works, a singular piece such as *Ballade Erbkönig*, is the result of an intuitive thought taken to the extreme. It is also a prime example of how text can be used to highlight the question of meaning in music. The fact that the text slides are exposed with such a deliberate reference to their timing in the music, invites the listener to make sense of the music in relation to that particular part of the narrative. Voice, in the guise of 'dialogue intertitles', makes up a large share of the text. Something that is being said by the child, the father or the Erlking, prompts us

⁴⁴ The ability to hear patterns inside noise, that is found amongst enthusiasts of EVP (electronic voice phenomena).

to search for that voice in the fabric of the sound; and because we do not immediately find it, nor can we detect anything that is a recognisable image of what is being suggested by the text, the music never falls into a state of illustrating the story, but remains as the primary medium within the hierarchy: music as language not language as music.

4.4 Language as Music

So we like extreme oratory because it is music, and the words seem to be only the vehicle for the music. Not a very original idea. But the words still work as words, so if we pay attention to what we are doing we can have the meaning of the words and the meaning of the music, merged. This way, neither seems to make a lot of sense, but the result makes you think, which is unusual⁴⁵.

Language or the word aspiring to music, is a paradigm that permeates many forms, from sound poetry to hip hop. The main difference between regarding language rather than music as the dominant medium, has to do with the context and the relative substance of the language, over the transparency of the music. The work of Robert Ashley, a composer who I have a huge admiration and love for, sits exactly on the cusp of language and music. Language and the voice play such a dominant role in the construction of his musical discourse, that the narrative role is largely taken up by the meaning of the text rather than the music. He deliberately drains the music of narrativity and expression, to let the words speak for themselves.

4.4.1 *Perfect Lives*

The ground-breaking video opera *Perfect Lives*, which provided a paradigm for much of his later work, contains some interesting use of on-screen text, though it is mostly tied to the text being aurally communicated. The strength of the visual element in *Perfect Lives*, realised by director John Sanborn, was made possible by Ashley's way of working with what he called 'templates':

Within the rules defined by the 'templates' the collaborators in all aspect of the work are free to interpret, 'improvise', invent and superimpose characteristics of their own artistic styles onto the texture of the work. In essence the collaborators become 'characters' in the opera at a deeper level than the illusionistic characters who appear on stage.⁴⁶

⁴⁵ Robert Ashley from a lecture entitled *Thinking About the Sound of Speech* (Ashley 2009: 500).

⁴⁶ Robert Ashley from the press release to the performance of *Perfect Lives* at The Kitchen, New York (1983) from Ashley (2009: 250).

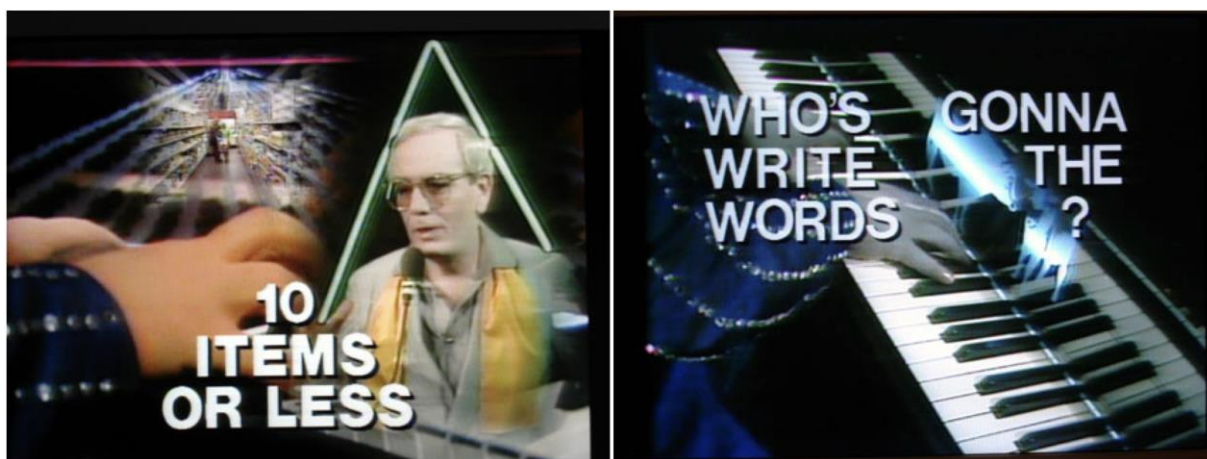


Figure 29 Stills from Robert Ashley's *Perfect Lives*.

This open approach to collaboration, which possibly originates from his experimentation with open score as a compositional practice in the 1960's, creates a strong independence of media, yet stays true to the structural principles of the whole. In this sense, the on-screen text in *Perfect Lives* not only serves the purpose of underlining a particular text heard, but also reinforces the visual character of the particular section of the piece.

4.4.2 *Ursonography*

Other examples one can cite of language as music within the context of text-film could be *Ursonography*, a performance of Kurt Schwitters' *Ursonate* by sound poet and vocalist Jaap Blonk, which uses live typography created by artist Golan Levin to visualise what is being vocalised. In this version of the *Ursonate* a live headshot of Jaap Blonk's performance is filtered and mixed with the text, which appears sometimes as a straightforward subtitle, but is mostly subjected to many forms of inventive typographic processing, affected by the incoming analysis of the live voice, thus highlighting the peculiarity of the vocalised text. It enables the audience to mimic and follow the vocalisations of Jaap Blonk on a more granular level, and similar to the example of the visual use of text in Ashley's *Perfect Lives*, it provides a visual characterisation reinforcing the structure of the work, delineating the sections and movements of the whole.

published by Gallimard, *Introduction à une nouvelle poésie et à une nouvelle musique* (Introduction to a New Poetry and a New Music), he clearly states his intentions:

Destruction of WORDS for LETTERS

ISIDORE ISOU	Believes in the potential elevation beyond WORDS; wants the development of transmissions where nothing is lost in the process; offers a verb equal to a shock. By the overload of expansion the forms leap up by themselves.
ISIDORE ISOU	Begins the destruction of words for letters.
ISIDORE ISOU	Wants letters to pull in among themselves all desires.
ISIDORE ISOU	Makes people stop using foregone conclusions, words.
ISIDORE ISOU	Shows another way out between WORDS and RENUNCIATION: LETTERS. He will create emotions against language, for the pleasure of the tongue.
	It consists of teaching that letters have a destination other than words.
ISOU	Will unmake words into their letters.
	Each poet will integrate everything into Everything
	Everything must be revealed by letters.
POETRY CAN NO LONGER BE REMADE.	
ISIDORE ISOU IS STARTING	
A NEW VEIN OF LYRICISM.	
Anyone who can not leave words behind can stay back with them!	

(Isou 1947, translated by David W. Seaman)

Together with this outspoken belief in the power of letter, came the idea that the letter needed to be liberated from the printed page, for the letter to come alive it must be uttered and not read. Hence the creation of the particular form of sound poetry that has come to be associated with this movement, the massed voices of Lettrist choirs intoning their remarkable brand of letter music. The relevance in including reference to their work in this thesis, is that they form part of a tradition of radical film utilizing on-screen text and original music in such an idiosyncratic way, that it has been hugely influential to subsequent generations of the avante-garde, even if their own films have remained less recognised in the mainstream.⁴⁸ Below I will touch on two Lettrist films: Isou's *Traité de bave et d'éternité* (*Treatise on Venom And Eternity*) and Lemaître's *Le Film est déjà commencé?* (*Has the film already started?*). I will also examine how the Lettrist's influence extended in one direction to the experimental film-scratching techniques of Stan Brakhage (see Section 4.6), and in another to the so-called 'discrepant' use of sound and image, that has travelled via

⁴⁸ In the introduction to the 1979 version of *Le Film est déjà commencé?* Maurice Lemaître claims it is a precursor to the work of Resnais, Godard, Marker and Duras. These film makers were certainly present at Lettrist screenings and though the claim of plagiarism which he later levelled at them would be too strong, the influence is undeniable.

Jean-Luc Godard and the films of Guy Debord, through to the contemporary documentary film maker Adam Curtis (see Section 4.6).



Figure 31 Two stills from Isidore Isou's *Traité de bave et d'éternité* (Treatise on Venom And Eternity).

Traité de bave et d'éternité is a two hour film, first shown (in part) at the Cannes Film Festival in 1951. The film can be seen in roughly three parts. Following an introduction of credits, dedications, an Isou bibliography, a disclaimer in the form of intertitles states:

'Dear spectators, you will see a discrepant film. No complaints will be accepted upon exit. The Management'.

The first part uses footage of Isou and other Lettrists walking the streets of Saint-Germain, Paris, the nerve centre of their movement. In the soundtrack we hear what sounds like a rowdy film club debate, and an extra-diegetic narrator explaining the position of the main protagonist Daniel in this. All the while a chorus of Lettrist poetry incantation is heard in the background. The extra-diegetic narrator in Isou's film refers to the hybrid multimedia nature of what is being seen and heard as:

...music within poetry,
painting within the novel,
and now the novel within the cinema.

This film is difficult to categorise, as almost every aspect of it is unconventional and must have seemed very iconoclastic at the time. In the first place, the film tries to eradicate the idea of narrative through image and replace it with voice. Speech is used as a form of 'détournement'⁴⁹ of the visual image (Cabañas 2014:15), while on the one hand it is used in a straightforward communicative fashion, as an on-going

⁴⁹ 'Détournement' is a term used by the Lettrists and later Situationists to describe a way of using existing artistic production or material from the mass media to change its original meaning or turn it against itself.

narration, both aural and visual, it is also disintegrated in the Lettrist sound-poetry of the soundtrack, and in the deconstruction of the letter on-screen. At a point near the beginning of the film, Daniel asserts:

I want to separate the ear from its cinematic master: the eye.

Isou's conception of *cinéma disrépant* and the underlined principles of *montage disrépant* lie behind the notion of separating the audio and visual components of film. He sees film history as splitting into two phases, the 'amplique' (amplic) and the 'ciselante' (chiseled). The amplic phase represents film at the service of narrative, where all the techniques of the medium are used in a coherent way, which amplify one another (Cabañas 2014: 8). On the other hand, with *Traité*, Isou wants to bring cinema history into the chiseled phase. Using the means of *montage disrépant*, narrative coherence is undermined and decentered. The word 'chiseled' is used here to imply a forceful rupture between the elements of the film medium. This film is an example of divergence (at some points in the film) on all modes of media correlation, as I set out in the previous chapter: sync, space, style, story, sentiment and scale. This divergence is a way of delivering his goal of total independence of the media: on the aural level - speech, sound poetry, field recording; on the visual level: intertitles, filmed persons, news-real footage, film scrapings, texts.

What is interesting to experience in this dislocated configuration of media, is that while elements are constantly being separated, our minds still make connections over the course of the film. When we hear a narrative concerning the protagonist's love life later in the film, we might connect it to images of young men walking the streets of Paris that we remember from earlier, as well as the newsreel images shown at that point. Or we might connect the powerful vocalisations of the Lettrist chorus with the explosive letters and scratches appearing on the surface of the film, even though they occur at different moments.

On the level of the materiality of the film, the so-called 'chiseling' also refers to the technique of manual interventions on the filmstrip itself, by scraping or painting each frame. This results in a restatement of the film's material status, as well as undermining the illusory power of the image. It also creates a meta-narrative of symbols and words, acting like a parallel world to the audio track. One could even say that the way language is deconstructed in the soundtrack (moving between the poles of coherent voice-over narrative to abstract sound poetry) is mirrored in the use of text in the visual domain, between coherent intertitles, credits, isolated words and their distortion into almost illegible script-like scrawl and graffiti.

4.4.4 *Le Film est déjà commencé?*

The person responsible for the chiseling of the images as well as some of the compositions heard on the soundtrack of the former film, was Maurice Lemaître. Very soon after the first screening of *Traité de bave et d'éternité*, Lemaître came up with his own cinematic rendering of the Lettrist ideals in *Le Film est déjà commencé?* (*Has the film already started?*). This film diverges even further from coherent narration towards a more extensive exposition of writing on film. Similar to Isou's work, text is implemented in three different ways here: as intertitles, as letters originating from fragments of filmed words (see below), and as almost illegible doodles. As well as contributing to the subversion of the cinematic illusion, each of the three layers of on-screen text have a different function within the film. The intertitles are constantly reminding the viewer of the conditions for watching the film. At the opening we read:

You are not going to see a film, but a film session, which must be composed of: 1. a special screen 2. a picture tape 3. a sound tape and 4. spectacular interventions in front of the movie house, in the entrance, and even in the audience.

This already suggests that the film is not simply contained within the frame of the screen but spills over to the street, to the auditorium, before, during and after the screening. Or questioning the audience's motivation to see the film:

Your stubbornness in seeing this film is incomprehensible.
But Lemaître's film has no value, neither aesthetic nor ideological.
It is a jumble of commonplaces of no interest.



Figure 32 Stills from Maurice Lemaître's *Le Film est déjà commencé?* (*Has the film already started?*).

The layer of filmed letters appearing around 20 minutes in is in part sampled from a Cinzano campaign, that was circulating in Paris at the time. Other sampled letters are reconfigured to spell 'IN-TO-LÉ-RA-CINÉ'. This is both reminiscent of the re-

appropriation of imagery, found in the public space prevalent in much of Dada art (for example the text image of "on a volé un collier de perles de 5 millions" in Fernand Léger's *Ballet Mécanique*), and also looks forward to the *détournement* techniques of the Situationist work of Guy Debord.

The third layer of on-screen text, the so-called 'hypergraphics', is text disintegrating into abstract signs and barely legible particles. One can see this as the primary Lettrist strategy of breaking down the conventional function of language into an abstract territory of signs. This is a strategy attempting to level the use of all the media in the film, whether it is found film processed as a negative, text particles, scratched photographs, or the mesmerizing and unintelligible incantations of the lettrist chorus. There is an attempt to collapse the boundaries between *opsis*, *lexis* and *melos* by a system of defamiliarization, where there is a constant replacement of one medium with the other. The goal here is not so much to create a synthesis of media, but to break it down into a form of abstraction, granting each medium the possibility to communicate on its own. Or in a semiotic sense: "the incorporation of the entire ensemble of communicative sign treats the totality of the symbolic not as signs but a signal constantly being re-encoded." (Charles 1989: 80)

Perhaps this work doesn't achieve the attention it deserves precisely because it successfully breaks through many of the conventional genre categorisations that it might belong to. *Le Film est déjà commencé?* is not just a film. Lemaître calls it a '*Séance de Cinéma*'. The published script of the film is split into three columns, 'son', 'image', 'salle', (sound, image, hall), where the live performance and actions meant to be occurring during the screening are meticulously notated. For example:

An extra will stand up in the theatre and shoot at the screen with a revolver. Another will tear the screen with knife-blows and will pass through the other side of the screen into the wings or will attack the wall. (Lemaître 1952: 164)

These type of performative actions recall in some way the spiritual forerunner of this film, the Dada spectacle *Relâche*, which featured René Clair's *Entr'acte* (with music by Erik Satie), in the way that the on-screen action must have been reflected in the on-stage actions, specifically in the example above, how Jean Börlin, the star and director of the show, bursts through the 'fin' screen and is then kicked back through it.

There are several important artists and movements which Lettrism influenced in the subsequent decades: for example Situationism, specifically the films of Debord, which develop his idiosyncratic essay-like narration over found footage, a dynamic relation between image, sound and narrated text, seeking to constantly disarrange the viewer's accepted sense of the 'spectacle'. I will not linger on Debord, fascinating as his ideas are, because his occasional use of on-screen text is relatively

conventional, consisting mostly of intertitles of chunks of text from his writings. What is perhaps interesting to note in the context of Debord's arrangement of media, is that he prioritises the word, in Aristotelian terms *lexis*, over the image and the sound, as he tries to create a strict separation of the three worlds, in modes that are divergent (see chapter 3.3). Thus he aims to establish an Aristotelian transparency of the three media, as a way of coming closer to the philosophically 'real'. One could even imagine that Plato would approve of his use of mimesis to deconstruct the illusion of the Cave. In the last section, I will discuss Curtis, whose films bear a very close connection to Debord, even though he himself might never admit to that connection.

4.5 Image as Music

On the other side of the Atlantic, Lettrism had one very noteworthy supporter, a young film maker by the name of Stan Brakhage. In 1962 Brakhage writes to Isou:

This is a long overdue letter to express grateful thanks for all you have given me. Some, almost, 10 years ago I saw what is here titled "Venom and Eternity". It immediately worked one of the most profound and lasting changes upon all my development as a film-maker...It is, and always has been, particularly marvellous to me that "'Venom and Eternity" does actually free me without, as so many other sources of so-called freedom, imposing its own new forms upon that freedom. (Letter to Isou from Brakhage, November 5, 1962, cited in Cabañas 2014: 135)

The impact Isou's film had on the godfather of American experimental film, Brakhage, is not so well known, but also not surprising to discover. One can imagine that the possibilities, carved out by the Lettrist's *cinéma dérépant*, became welcome armoury to a young film maker, wanting to detonate the foundations of conventional narrative film. Brakhage, who created more than 400 films in his 50 years of working life, worked almost entirely without sound. His films became known for their attempt, not only to liberate the visual from a too dominant narrative interpretation, but to create a new way of seeing, reaching beyond conventional perspective and composition:

Imagine a world alive with incomprehensible objects, and shimmering with an endless variety of movement and innumerable gradations of colour. Imagine a world before the 'beginning was the word'. (Brakhage 1963)

The quotation above appeared in *Metaphors on Vision*, a publication which established the basis of his theory about film and his attempt to banish the word and the way of seeing through language from the screen. One can detect shared

principles with the Lettrists' attempt to destabilise the word, though Brakhage and Isou go about it in very different ways. Rather than breaking narrative coherence through extreme 'chiseling' of the media (in Isou's case), or through the dissolving of the media into each other (in Lemaître's case), Brakhage takes the high-modernist route as advocated by Greenberg, to purify the media, in this case the visual, *opsis*, and to return it to a state of pure light and form, a state of depiction of the world that is non-representational. The influence of the abstract-expressionists is in this sense very clear in Brakhage's work, but the influence of the Lettrists is undoubtedly there as well, not only in ideology, but in some techniques, such as the direct scratching onto the film itself, which became a trademark of his.

4.5.1 *I...Dreaming*

There is one late film by Brakhage, which is interesting to discuss in the framework of this thesis, as it uncharacteristically brings back both word and music into his world. *I...Dreaming* (1988) is an 8 minute film using a soundtrack of cut-up Stephen Foster songs, with selected scratching of the lyrics onto the film stock. The images in the film shows Brakhage himself, as an old man, next to time-lapsed images of his grandchildren. The melancholic nature of the resulting film, is probably autobiographical, as it was made just after the separation from his first wife.



Figure 33 Stills from *I...Dreaming* (1988) by Stan Brakhage.

Beyond the contemplative mood of the film, the overriding sense communicated is the discontinuity of memory, consciousness and emotion. This is underlined in the choice of words scratched on the film, picked out from the song heard on the soundtrack, "void", "cold", "longing", "dreaming visions", "starlight in silence", "pleasure", "bear", "true".

What occurs on the level of word projection repeats in the soundtrack. The music is constantly skipping back and forth, repeating and isolating single words, as if we are

listening to a worn-out record. The words are amplified, both visually and acoustically, and attain almost the same quality, of being fleeting, worn-out, and emotive, in the way they are presented to both eye and ear. Analysing the film using 'media correlation', one could say that between projected text and music, there is a high degree of convergence in the semantic modes of the 'story' and the 'style'. There is less convergence in 'sync' and 'space', which is what gives the film a more distant and detached atmosphere. Analysing this from the point of view of the relation between the film and the sound editing, it is clear that the temporal mode plays an extremely important part. Although music and film are rarely synchronised in terms of their edit points, there is a rhythmic complementarity in how the discontinuity is managed in both media.

Even though this complementarity, and the amplification and focus of the song lyrics contain the ingredients to push the film into music video territory, the occasional dominance of the soundtrack is counterbalanced with a complexity of visual information drawing its attention from a strong sense of what is inferred, rather than revealed in the image. There is a clear subjectivity in the way Brakhage films himself and the space around him. Furthermore the scratched words are not overly easy to perceive. In this way, using the metaphor model of analysis, one can claim that the visual retains the cinematic sense of being the target rather than the source. If one could define the meaning of the narrative, it would result in something like: the emotional sense of consciousness rests more in the techniques of the materiality of the film, rather than only in the meaning conveyed by the images themselves.

4.5.2 *Newsprint*

Visual disorientation is also at play in Guy Sherwin's 1972 short film, *Newsprint*. Newspaper is glued onto 16mm film, so that it not only covers the part that is projected, but also the part of the film that is used as the sound track, the so called optical sound. That is not to say that sound plays the primary role, as I would argue that the visual is still at the fore. Nor does it provide the primary material for the work, for as the title suggests, the material or subject of the film is printed news, words and paper. What the sound does is act as a bridge between the material level of the newspaper and its projection as light. It underscores the rhythm of the translation between print and light. The words in this case are sonified, not in their meaning, but in the density of their print. The more dense the print on the optical track, the more dense the noise resulting. Closely pixelated images, as those found on newspaper photographs, result in some kind of harmony emerging out of the noise. Emptier frames provide a momentary respite from the incessant scratching of the breakneck flicker of printed words.

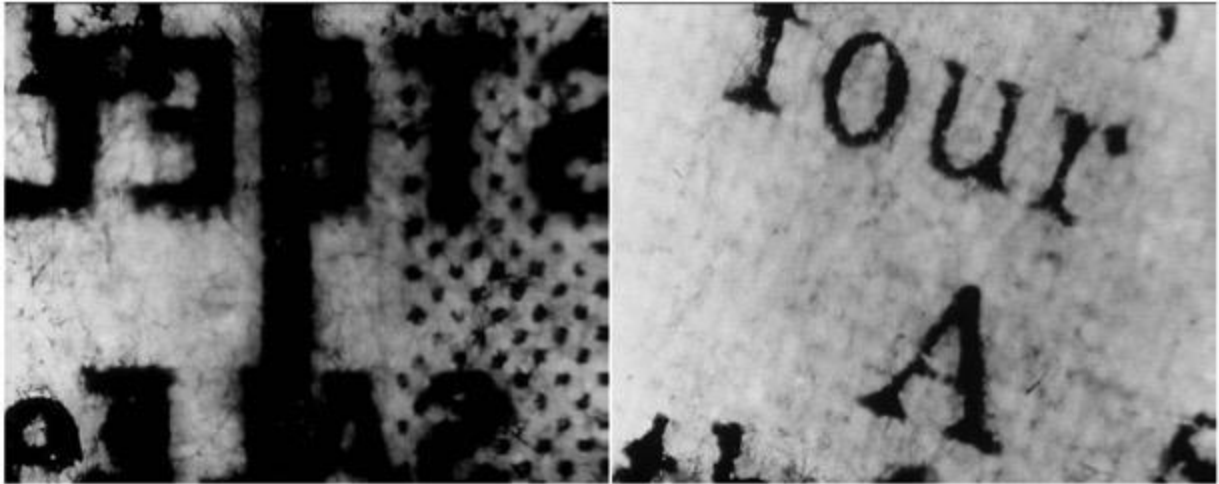


Figure 34 Stills from *Newsprint* by Guy Sherwin.

The act of reading is also under scrutiny in this film. That is why the language-to-image relationship is so crucial, but because it is mediated by sound, our ears help our eyes to see:

When reading text such as a newspaper we might assume that our eyes move smoothly over the words. In fact we scan text in a stop/start motion since we can only register words when they appear static. These eye movements, known as 'saccades' are similar to the intermittent motion of the film projector. In *Newsprint* the relationship of eye to text is reversed, for the newspaper is animated by the projector while our eyes are, relatively speaking, static. (Sherwin 2007: 26)

The eye is not only busy trying to recognise fragments of words and piecing together some kind of semantic sense. From my own experience of the work, I would say that one quickly gives up this idea. Rather, the experience becomes highly visual, as the viewer is plunged into a microscopic world of print, where the marks, smudges, and glitches left on the printed paper are magnified in size and are reinforced in the audio domain. As the dot screen patterns of print grow and recede in visibility, so does the accompanying sound change from harmonic coherence to intermittent noise. It is as if the projection machine, the solar cell that picks up the light through the optical track, is reading the text for us, like a machine-reading of the news. We hear the text as a voice, but can understand nothing of the language.

A live version of the film was adapted in 2003, *Newsprint #2*, where two projectionists use identical prints to project superimposed images onto one screen. The image size of each projection varies throughout the performance. They take turns to freeze images, so that some words in this version do become almost legible. The fact that the film here freezes, means that the sound also stops, creating an interesting rhythmic interplay between the two projectionists. Another important consequence of the superimposition of the two films, is that at times it is as if a three-dimensional

space is created between the two streams of newsprint. Sherwin also describes a 'rushing' effect that occurs:

At times, when the two projections are slightly out of phase and of a slightly different size, a rushing motion occurs between the two images and in the two soundtracks. The attempt to catch these unusual kinetic movements forms part of the projection performance. (Sherwin 2007: 29)

The performative nature of this version of *Newsprint* is again enhanced by the soundtrack, because one hears the rhythm of the freeze framing as it moves between each player. Sherwin also scores a general pause for both players throughout the performance, where both projectors switch off for a second or two, so that one experiences this hiatus both in light as well as sound.

In *Newsprint*, Sherwin tries to find a balance between image and sound media through a mechanical translation of the two. It is one of the few examples of text forming a part of his image world. Another such piece is a collaboration with Lynn Loo, *Vowels and Consonants* for 6 projectors, based on one of her films. Here language is deconstructed into letters floating and flickering across the screen. Again, like in the previous examples, one experiences the visual medium through the material of language. And again, because this material carries with it a strong phonetic impulse, the letters also become imprinted in our minds as sounds.

4.6 Music as Image

Music video is the genre par excellence for examining how image reinforces a meaning potent in the music. Primarily intended for marketing purposes, selling records and concerts, the hierarchy of media seems to be already economically defined. Namely, that the medium is predominantly about music, as the music is visualised, given context, narrative, and dressed up. Most of the music in this context are songs already equipped with voice and narrative, so it is difficult to separate the text from the music, as it was conceived as one entity. The video might underlay an aspect of this narrative or present an image world that converges or diverges in various ways, as outlined in the media correlation chapter, but what it aims to achieve, in most cases, is provide image material to lend meaning to the song rather than vice versa.

4.6.1 Three Music Videos

Example of music videos which have diverging correlations in the various parameters described previously, namely style, story, sync etc. could be: Björk's *All is Full of Love*, a love song of sorts in which she sings "You'll be given love / You'll be taken care of / You'll be given love / You have to trust it". The video was directed by Chris Cunningham as a cyborg romance, not wholeheartedly contradicting the narrative of the song but creating ambiguities in terms of how one understands the words 'you' and 'love'.

Another video toying with stylistic dissonance, imposing an 18th century narrative of duelling aristocrats onto a song, which is essentially British rap about personal and social identity, is *Too Cold* by Roots Manuva. The lyrics: "Don't you see that we some big broad bad man / Born and bred in this 'ere big broad bad land / Known all over the world as a mad man / Life is hard but it's just too bad man", get transposed away from the usual urban housing estate setting of British hip hop, to a Regency era stately manor. This context underlines some of the 'classical' sounds used in the music, as well as some of the concepts of class and identity alluded to in the lyrics. The video though, never simply becomes about Regency aristocratic life, it remains a metaphor to understand the lyrics of the song.



Figure 35 Still from Chris Cunningham's video for Björk's *All is Full of Love* (left) and Roots Manuva in the Regency fantasy video for the song *Too Cold*.

The ambiguity inherent in song lyrics, the way they can refer to something very specific, yet applicable universally, is something consciously emphasised and manipulated by video directors. Another example is *This is Hardcore* by Pulp. Here the metaphor inferred throughout the song is of seeing life as a movie, either a porn flick or an unrealistic Hollywood-esque fantasy.



Figure 36 Stills from the video to Pulp's *This is Hardcore*.

There are strong sexual metaphors in the lyrics, as well as in the title of the song, but the video concentrates predominantly on an ironic over-the-top portrayal of Hollywood golden age film business, and so underlines the idea of fantasy vs. reality. The sexual allusion in the lyrics: "You are hardcore, you make me hard/ You name the drama and I'll play the part/ It seems I saw you in some teenage wet dream/ I like your get-up/ if you know what I mean", is underplayed, so that the ambiguity of what is meant by "hardcore" prevails. Furthermore, the narrative in the video ironically compares the glamorisation of violence in the film industry with the unrealistic portrayal of sex that is alluded to in the lyrics. A footnote to the short discussion of this video, is the use of text in the momentary shots of leader tape, clapperboards, and intertitles. It happens too quickly to decipher the text, but it is used rhythmically to punctuate the change of scenes, and also to reinforce the sense of detachment and perspective towards what is being depicted.

Fixed video or VJ-ing for live performance tends to be less narrative than conventional music video, though sometimes the images overlap. There is a clear focus on the live performance, and the visuals take an even less prominent role in the hierarchy of media. What is an interesting trend here, is the use of animated text, not always in sync with the singer, but as a way of presenting the lyrics or some of the keywords used in the lyrics in a dynamic and graphic form. Examples of this can be found in some of Kraftwerk's recent live performances, or in big production hip hop shows like Stormzy (Glastonbury 2017).

4.6.2 *It Felt Like a Kiss*

An interesting and sophisticated example of the relation of image and song is *It Felt Like a Kiss* by documentary film maker Curtis. This film was part of an 'immersive' theatre production by Punchdrunk for the Manchester International Festival in 2009. The main theme of the film is the juxtaposition of American soft power and hard

power; the American dream, encapsulated in a series of pop songs is juxtaposed against the Vietnam war, the Cold War, assassinations and power grabs of the 60's and 70's. The film begins with the titles:

When a nation is powerful it tells the world confident stories about the future - the stories can be enchanting or frightening - but they make sense of the world - when that power begins to ebb - the stories fall apart - and all that is left are fragments which haunt you like half forgotten dreams.⁵⁰

Typical of Curtis' films, he constructs a narrative by piecing together a wide variety of archive materials, mostly from the BBC archives. His films are usually driven by interplay between images, music (mostly songs), text intertitles, and his extra-diegetic narration. In the case of *It Felt Like a Kiss*, his voice is absent and replaced by a more prominent use of on-screen text. What makes this film so special, and perhaps the reason why his voice is absent, is that the voices of the many songs featured seem to take over the narration of the film. There are several threads of narrative, which exchange prominence, some originating directly from the songs and the lives of the singers. Doris Day and Tina Turner are two of the singers featuring in this film, constructed of a playlist of over 30 songs.



Figure 37 Stills from the Adam Curtis' *It Felt Like a Kiss*.

The opening shot of the film features a man getting out of bed and lighting a cigarette, accompanied by Ruth Brown's song *Oh What a Dream*. We are led through enigmatic images of a belly dancing club, an airplane, a chimpanzee, what looks like a science laboratory, a painting of a nude, a boy running into the backdoor of a suburban house, a sunset, a parade and so on. These fast editing of images is typical

⁵⁰ Opening titles of *It Felt Like a Kiss* (2009) by Adam Curtis

of Curtis, not only posing the question 'what is the dream?' but 'how is a dream being constructed?' and 'who is dreaming?'. This is partly answered by the next set of on-screen text, overlaid onto an ongoing barrage of images: "This is the story of how America set out to remake the world"⁵¹

Fats Domino's *Let the Four Winds Blow*, with the lyrics: "Let the four winds blow/ Let 'em blow let 'em blow/ From the east to the west / I'll love you the best", accompanies the introduction of the central characters of Curtis' narrative: Doris Day, Rock Hudson, Saddam Hussein, Lee Harvey Oswald, Enos the Chimp, and "everyone above level 7 in the CIA". The formula here is typical of the relation between image, projected text and lyrics. The song provides a constant counterpoint to the narrative of the projected text. In this case, it can be interpreted as: how these characters, coming from all corners of the world, came to be loved (and wield so much power). Juxtaposition of different quality of images is also typical of his visual language. Mixing the seductive with the shocking, the fluffy with the tearful, is something that compares to his juxtaposition of pop songs with other more cinematic use of music in the film such as music by Britten and Shostakovich.

In another scene, *Pink Shoelaces* by Dodie Stevens underscores a montage of images and text relating to the attempted assassination of Fidel Castro by the CIA: "Now I've got a guy and his name is Dooley/ He's my guy and I love him truly/ He's not good lookin', heaven knows/ But I'm wild about his crazy clothes." At the point just before Stevens sings: "He takes me deep-sea fishing in a submarine", a text relates how the CIA tried to blow up Castro with an exploding conch shell when he was snorkelling. Shortly after, we are told that they also tried to poison his shoes, accompanied by the chorus: "He wears pink shoelaces". Several connections are made with the song lyrics, the on-screen narrative, and the images (the sea, the shoes and Castro) , which open up the possibility of further metaphorical exchange between media. The relation here between the media is a little more complex than simply 'music as image', as I set out in the section heading, because Curtis is always shifting the perspective, and giving turn to each medium to take centre stage. However what is interesting in the context of music video and the song form, is that the open-endedness that is the hallmark of a good love song, the way that the listener can relate the song's lyrics and emotion to their own or indeed any situation, is here applied not to personal romantic sensibilities, but to an understanding of cold war geo-politics. The 'guy' in Stevens' song could be Castro, or the CIA, or even Lee Harvey Oswald, who is mentioned just after. An exact match is not important. What matters is that we constantly perceive the songs not only as a subtext for the political

⁵¹ Interestingly, the music that he cuts to, over this sequence is George Delerue's soundtrack for Jean Luc Godard's film *Le Mepris*. Curtis has never acknowledged the influence of Godard in his use of intertitles, which seem to have a very similar visual aesthetic.

narrative of the film, but as narration itself, the story of the misrepresentation of the American Dream.

At the end of the film, after a climax of very fast montage of images to the soundtrack of *River Deep Mountain High* by Ike and Tina Turner, comes one of the most shocking juxtaposition of image to song. Peggy Lee's *Is That All There Is?* was written by Jerry Leiber and is reminiscent of Brecht/Weill repertoire in its bitter-sweet atmosphere of resignation and the way the verses are spoken, and only the chorus is sung. Peggy Lee recounts an incident from her childhood, watching her family house being burned down, while the chorus plays along to an extended shot of a man being immolated (possibly during the Vietnam war). The sense of disillusionment and resignation the song conveys, along with the extreme juxtaposition of emotions, of beauty and horror, is indicative of what the film is trying to communicate. It leaves the viewer emotionally drained, disorientated and with the guilty pleasure of being seduced in full daylight by the power of compelling images and sounds, along with the sense of possibly having understood something of the complex networks of causality of our recent history.

4.7 Summary

There are many examples, which the limitations of space prevented me from discussing them in this chapter. Some of these include: digital poetry, electronic literature, opening titles in film, the use of language in contemporary art, data art, computer art, motion design and advertising. An encyclopaedic presentation of the text based work in some of these genres is covered in the catalogue of the ZKM exhibition: *Schriftfilme, Schrift als Bild in Bewegung* (in English as: *Typemotion, Type as Image in Motion*⁵²). There is also an increasing use of text projections by a younger generation of contemporary composers, including Jennifer Walshe, Joe Snape, Larry Goves, Genevieve Murphy and Andrius Arutiunian, which would be interesting to investigate at some stage.

What I tried to show in my choice of examples, is that the way text is used in relation to other media differs radically, depending on how the hierarchy of media is presented, specifically depending on what the target medium is. Some of my choices have also been motivated by personal taste, as these are some of my favourite works in any genre. Nevertheless, they form a strange history of mostly 20th century art practice, as most of the examples are to be found on either side of mainstream

⁵² Scheffer, Bernd & Stenzer, Christine & Wiebel, Peter & Zehle, Soenke (eds.). 2013. *Typemotion, Type as Image in Motion*. Hatje Cantz Verlag/ZKM.

modernism, belonging to experimental avant-garde or popular culture. This marginalisation, in my opinion, goes back to the fostering of medium-specificity in high modernist aesthetics, and the irritation many had with hybrid art forms, specifically with what Frampton sees as the "ostracization of the word" (Frampton 1983: 83).

In the following chapter I will be discussing my own music-text-films. Most of these works would fit into the category of Music as Language (Chapter 4.3), as they share much in common with the examples I presented there, and because the main concern in the use of text is, how language informs the target medium - music. Nevertheless, the media relations are never completely straightforward, and I will focus more on how the dynamics shift not only between the basic levels of melos, opsis and lexis, but between various levels of the music itself.

Part II

Music-Text-Film

Introduction

In the following chapters I focus on my own work, discussing some of the strategies used in the context of music-text-film. This discussion will be organised in four main sections, Internal Monologues, Unanswered Questions, Voiceprints, Interactive Scores, representing some of the main themes I have been exploring in my work. In the last few years, I have written about 30 compositions which would fall under the category of music-text-film. These form an exploration of projected text within concert music or sound art in the last years, and although I provide a concise description of most of these in the Appendix, I will focus on a more detailed discussion of twelve in the following text.

'Internal Monologues' explores work that highlights predominantly first-person narratives, especially those that derive from conscious or semi-conscious discourse taking place within the mind. It includes three examples of this work, namely: *Dreams of the Blind*, based on five dream accounts of blind people; *Mnemonist S*, an account by the mnemonist Solomon V. Shereshevskii (as detailed in the work of psychologist Alexander Luria⁵³) of narratives he would construct in order to recall abstract sets of phonemes; and *Memoryscape*, an ensemble work based on the recollection of fifteen early memories. These are some of my earlier music-text-film works, and I include them here because they specifically highlight some of the theoretical positions described earlier in this thesis, namely internal diegetic narrative and inner voice. Some of the other work that deals with this approach, which is to be found in the appendix is: *The Arrest*, another dream narrative; *Nerve*, an orchestral work made up of fragments of inner dialogues about stage fright; and *Walls Have Ears*, a short piece for string quartet and voice based on a poem by Mehmet Yashin, that deals with the experiences of trying to hide his native language in a hostile environment.

'Unanswered Questions' deals with a video and two sound installations that explore question and answer structures. *Machine Read*, an ode to Dick Raaijmakers, is a short video piece based on some texts from his essay *De kunst van het machinelezen* (*The art of reading machines*)⁵⁴, that encodes questions that deal with the relation between language and music. *Norms of Transposition* and *Dodona*, also deal with the aspect of the encoding of questions and answers between music and language, which in this case, take the form of installation pieces. I explore the question of what can be understood in this particular process of translation between media. These pieces represent examples of some of my work dealing with encryption of language into

⁵³ Luria, Aleksandr Romanovich. 1987. *The mind of a mnemonist: a little book about a vast memory*. Cambridge: Harvard University Press.

⁵⁴ *De kunst van het machine lezen*, 1978, Raster: no. 6, pp. 6–53.

musical structures, some of which is not represented in the thesis, because it falls outside the focus of text-film: this includes pieces such as *Politicus*, a sound installation that presents the entire book of Plato's *Statesman* as a 10-hour encoding for prepared disklavier. In the appendix there are other examples of work which does have a text-film component, *The Queen is the Supreme Power in the Realm*, and *Codex Simple* (both inspired by telegraphic code books) .

The works presented in 'Voiceprints' are examples of material constructed on the manipulation of spoken voices. Voices are put under a sonic microscope, dissected, disembodied and reconfigured. The projected text in these pieces, provides a hint regarding the spoken content, or the context in which the voices are found. *Wordless* presents twelve portraits of residents of Brussels in the form of interviews, where the words are removed to leave only the hesitations, breathing, emotional reactions and environmental sounds. *Varosha /Disco Debris* is assembled from a mass of granulated voices and a narrator, leading the listener through the remains of a ghost town. The projected text and images, shot through with holes, provide a hint of the atrophied space that is described. The orchestral work *Der Komponist*, is based on a 20 second fragment of speech by composer Helmut Lachenmann, stretched out over 20 minutes. The manipulation of voices is something that figures highly in my work, and is included in many works of the last years, which focus on the materiality of the spoken and sung voice. Examples of those pieces, that do not utilise projected text but are essentially driven by the sonic exploration of different aspects of voice are: *Paramyth*, *Lunch Music*, *The World Feels Pressure*, *The 100 Words*. Works that do use text-film and which are to be found in the appendix are: *The Musicians of Dourgouti*, an ensemble work based on the voice of a Greek-Armenian musician recounting life in an impoverished immigrant neighbourhood in Athens of the 1960's.

Finally, 'Video Scores', is about more recent work that takes the idea of projected text and turns it into notation for musicians, fixed or interactive. Departing from the notion that instead of having a separate visual code for musicians and a separate one for the audience, the two could be merged as one. This is the idea behind karaoke; texts and music notation is blurred and amalgamated into a video score that conveys musical information in an animated graphic way. The first result of this idea is therefore to be found in my interpretation of this mass medium: *Karaoke Etudes*. In the interactive tablet score *Oneiricon*, the tablet acts as notation, sound-generator, book and source for visualisation, for both audience and musicians. The act of reading is shared, the wall between musicians and audience is partially removed. In *Trench Code* the score is generated live by one group of players responding to the listening or reading of another group. Again, this is something that is evident to the audience, who share the position of the players in the performance space. These are directions and potentials that I see opening up in terms of using the current and future technologies, which could alter the way notation functions as an act of shared public reading.

Chapter 5: Internal Monologues

I'm writing you all this from another world, a world of appearances. In a way the two worlds communicate with each other.⁵⁵ (Chris Marker 1983)

Short term memory, long term memory and dream accounts are some of the topics dealt with in my early music-text-film works. I was initially drawn to these topics because they give an insight into the workings of the inner voice, or that elusive private commentary that tries to make sense of the information flow in our minds. In Chapter 1, I briefly discussed the difficulty of capturing the inner voice in flow, the chain of thought from conscious to subconscious cognition, that helps us piece together our experience of the world. I was curious to examine different domains of memory where the inner voice was useful as a narrator of these private worlds, and in turn expose that in the composition through the form of silent reading brought about in music-text-film form. The three pieces discussed in this section, *Dreams of the Blind*, *Mnemonist S* and *Memoryscape*, all deal with an aspect of internal monologue, specifically in the way that the inner voice functions as a go-between within the depths of our memory. The inner voice seems to be sending queries into the subconscious trying to access information that is perhaps only half-formed or half-remembered; fragments of a sensation, a dream, or an experience in the distant past. In the case of *Mnemonist S*, the inner voice is not only looking into a visualised memory system that is consciously created for the sake of remembering, but it is seen from the perspective of decades later.

The 'absent' voice acts as a bridge between two barely conscious private spaces: the mind of the mnemonist or dreamer and that of the reader/listener. Just as the voice of this mnemonist has been used to convey and understand the images that he or she has experienced, the inner voice of the audience undergoes a transference in the act of reading; the voice of the mnemonist is stripped of its original vocality and emotional undertones, and takes in its place a mediated instrumental one. The imagined vocal presence of the mnemonist embeds itself into the voice of the audience; the inner voice of the reader/listener becomes possessed by the narrator. The relation between the absent voice and the instrumental music, which in all three cases becomes a surrogate voice of sorts, is very clearly pronounced. The music here deliberately shadows the hypothetical rhythm and melody of the narrator, giving it some form of vocal expression, and in so doing, takes on the function of voice.

What is also explored, in a more general sense, is the question of medial transference – how the absence of one medium can place more weight of meaning onto another.

⁵⁵ Chris Marker, introductory voice-over from his film *Sans Soleil* (1983).

For instance, in the case of *Dreams of the Blind*, the absence of voice is replaced by text and music, and the absence of image is highlighted by the material of the dreams, which, because they originate in the minds of blind dreamers, rely predominantly on the senses of hearing and touch rather than sight. In the context of *Dreams of the Blind* one can describe the following sequence: Sensations of touch and hearing are transferred to narrative situations and emotions by the dreamer, which are then put into spoken words by the dreamer's account and documented as a written text by the lab researcher. I take this text, set it to music (as notation for musicians and electronic sound), and with an ensemble of musicians, perform the music and project the words on screen. The audience hears the music and reads the text (senses of hearing and sight), which leads to a recreation of a voice (an imagined voice) and perhaps a visual recreation of the imagined dreamer's narrative. In *Mnemonist S*, the multi-modal transference is highlighted even more intensely because the narrator and subject of the piece, Solomon V. Shereshevskii, was himself an extreme synaesthete. The multimodal translation of sensation is mirrored in how the text is both visualised and sonified. In *Memoryscape*, the uncertainty in the narration of people's earliest memory is reflected in how the text appears, as one block of text, but gradually fading in and out in a non-linear fashion.

In all three pieces the narration is highly intra-diegetic, the worlds that are narrated either came into existence in the mind of the narrator or have a high degree of subjective interpretation of the past. Also in all three cases one could say that words are used to describe something intangible: an emotion, a sensation, an imagined space. There is a significant effort involved in dragging these impressions into consciousness and mapping them into words. There is also presumably a significant loss of content. The unreliability or incomplete nature of narration is something that is explored in these pieces. The music opens up the space between words, and the sonic information found in this non-linguistic space gives a sense of the limits of language, by conveying a sense of the subconscious.

An interesting word, which could be used to describe the function of music here is: 'lalangue', coined by the French psychoanalyst Jacques Lacan.⁵⁶ 'Lalangue' is meant to stand for everything that grows parasitically around the fractures and fissures of language. Lacan defines 'Lalangue' as being the bedrock of the unconscious, unstructured noise of polysemy where eventual meaning emerges.

⁵⁶ This was first mentioned by Lacan in 1971 at a conference at Sainte Anne's Hospital and is connected to his concept of *Jouissance*. (Simonney 2012: 7)

5.1 *Dreams of the Blind*

Dreams of the Blind is a 30 minute suite of five pieces for an ensemble of nine instruments, electronic soundtrack and text-film.⁵⁷ It was written in the summer of 2006 for Ensemble MAE. The five movements of the piece are each based on dream accounts of blind people, collected at the 'Dream Bank' research site of the University of California, Santa Cruz.⁵⁸ Three of the five have direct references to music: *Supermarket Guy*, *Car Radio*, and *Winter Funeral*, while the other two: *Floating Table* and *Hairdresser*, are informed by a combination of touch and hearing.



Figure 38 Still images from *Dreams of the Blind*.

The metaphor of the blind dreamer is one that can be taken to generally represent how meaning is conveyed through different sense perceptions, both in the mind of the dreamer and in how these dreams are understood by a sighted reader. There exists here an inherent asymmetry in sensorial communication, because while one usually associates the idea of dream narratives in terms of images, the narratives of blind dreamers are not formed by visual perception, but, just as in their waking lives, are predominantly fed by the senses of hearing and touch.

In terms of musical narrative, *Dreams of the Blind* follows two distinct paths: there is an instrumental layer, which reinforces, interprets, and embodies the projected text; and there is a constant drone-like electronic track, made up of complex synthesis sounds, which is an attempt to metaphorically represent the physical space where the dream takes place, the body of the dreamer. There is a classic 'figure and ground'

⁵⁷ Link to recording by Ensemble MAE: <https://vimeo.com/226624063>

⁵⁸ Link to online database can be found here: <http://www.dreambank.net/>

relationship between instrumental sound and electronics, which, as I have explained in chapter 2, reinforces the perspective of narration, as it creates the sense of two separate ontological frames. This also highlights the different types of listening experiences between the two media, between a syntactic one and a more directly sensory one: words vs. body. These two strands of function are intertwined in as much as the words are intertwined with the music. The intention is that the instrumental 'voice', as suggested above, acts as the bridge between the sensorial/emotional experience of the electronic sound and the projected words. In many parts of the piece, the ensemble is, in this sense, the medium between words and body, something like a voice that exists in parallel with the listener's own reading voice.

In Part I: *Supermarket Guy*, the dreamer imagines a telephone conversation with a man who takes care of her weekly delivery from the supermarket. This is not a normal conversation: "he was playing an organ and he wanted me to play mine". The text in this movement is built on a regular pulse matching the telephonic character of the music; giving the voice of the dreamer a somewhat mechanical or remote character. Metaphorically, the movement develops along the narrative of the text: the dreamer is trying to find a particular sound on the organ and is frustrated in her attempts to do so. This translates into the music as a sequence of scale patterns, becoming increasingly dense, conveying a sense of effort that comes to no fruition.

Two clear voices are heard throughout the movement. The first is the voice of the singer, who at a certain point declaims: "wait a minute". There are similar moments in the work as a whole, when the dreamer is saying something or when something is being said to the dreamer, the singer in the ensemble voices these words. These are the only times text is heard rather than read in the projection, as if the dreamer is hearing her own voice and at the same time recounting what she has heard herself saying.

The second distinct voice in the movement is a florid violin line. This part differentiates itself from the other instruments, in that it has a soloistic character, making it identifiable with the subjectivity of the dreamer, not so much as the voice, but as something that is being expressed through her situation, an emotion or an attitude. These two distinct levels of voice (others can be said to be: the text, the electronics, the tutti ensemble) are what creates a shift in viewpoint in the narration, and which gives the listener a sense of perspective.

27

Fl.

Cl.

Tbn.

Vib.

E. Gtr.

— mine *mf* warm and fuzzy

Txt.

and I

Pn.

Mezzo.

Vln.

Cb.

8

28

Fl.

Cl.

Tbn.

Vib.

E. Gtr.

long sound *mf*

Txt.

said "wait a minute"

Pn.

Mezzo.

Vln.

Cb.

wait (airy) *mp* sul G very slow gliss. *mp* sul pont

a sul D/A

Yannis Kyriakides / *Dreams of the Blind*

9

Figure 39 Page 8-9 of *Dreams of the Blind*, showing the violin motif (page 8) and the beginning of the voice phrase "wait a minute" (page 9).

Part II: *Floating Table*, projects a very different sense of perspective and place. Already the way that the text appears and disappears with a fade, and an instability in its position on the screen, suggests a sense of floating, an unreal atmosphere. The words are always reinforced by an alternating note on the piano or vibraphone, tightly connecting the voice of the dreamer to the ensemble. The echoing of the notes from one instrument to the other is there to reinforce a sense of the crossing of a vast space, like a long reverb with an echo of about one second. These notes are coloured by the sonorities of the rest of the ensemble, they sustain the harmonies and occasionally slide between notes, with modest glissandi that mirror the subtle hovering of the text. The dream account here describes a massive room full of people, sitting around a floating table:

It felt like we were sitting, but the more I think about it, the more I think that the chairs were almost suspended, they were kind of floating. My weight didn't feel like it was pressing on the chair, but the people were just kind of talking, but there was nothing resolved or purpose to be there.

The sense of floating in the sonic image is partly achieved by a wide register gap between the very low sub-tones of the electronics, which gently oscillate in pitch and volume, and the slowly fluctuating melodic contours at least two octaves above that. The electronic part, just like in the rest of the piece, stands for the embodiment of the dreamer; the prostrate body, whose physical functions in sleep are conjuring this floating world. Aside from the image of suspension that is carried in the music, the sense of purposelessness is conveyed in the way that even though there is a harmonic development of sorts, and a gradual change of intensity, there is no real narrative resolution.

Part III: *Hairdresser* has a distinctly lighter tone. The text is presented in a more rhythmic fashion, four words at a time moving from the top to the bottom of the screen. This reinforces a sense of a longer phrase structure, as opposed to when one word is shown at a time, which leads to a greater tendency to group the musical phrases in terms of how they are visually structured. The dream describes a visit to a hair salon where robots have taken over the job of the hairdresser. Instead of the hair being manually cut, machines seem to come out of the walls to do the job:

I sat down in the chair and she pushed a button somewhere and these scissors came out of the wall, I think, and began to cut my hair. They were automatic. They would cut for so long and then they would stop. I could hear this noise, this motor, I guess they were going back into the wall or wherever they came from.

This dream seems to be very reliant on how mechanical sound events, in a place like a hairdresser's salon, can give a sense of a robotic world. Take away conversation and one's sight and one can well imagine that the job is being done by machines. In

the case of this dream, there is conversation with the hairdresser, though she seems to be directing the dreamer from machine to machine, whilst trying to put her at ease. This dialogue, like in the first dream, is sung, it is the only part of the text that is heard. The role of the ensemble is more illustrative: it builds up mechanical noises using a wide variety of extended techniques, to give a sense of an automated world. Unlike the previous movement, the voice of the dreamer is less present in the music of the ensemble, which seems to be representing the diegetic space of the dream rather than the point of view of the dreamer. There is a strong metaphoric correlation between the ensemble and the text but with less temporal coherence. What is more in sync with the text is the electronic soundtrack, though in this movement it is harder to separate between soundtrack and ensemble, because the material converges. This has the effect of flattening the perspective between the different layers of sound material, making the focusing of the voice more ambiguous.

Conversely, in Part IV: *Car Radio*, the perspectival shifts between text, musical syntax and sound are more clearly pronounced. The narrative opens with the dreamer describing the location of the dream and whom she or he was with in the dream.

I was in the hospital where I work, I was going down the hall with my friends Virginia and Penelope. We were going down the cafe for lunch.

The music is composed of a drone layer of electronic sound that acts as an unstable pedal point for the instrumental phrases that are played. The instruments seem to synchronise the beginnings and endings of the text phrases with short notes, positioning the first person voice within the realm of the music. The longer notes on violin and recorder, which accompany the text, help give an expressive, emotional or vocal quality to the words. The flow of text is occasionally interrupted by electronic sounds from the drone layer, which seem to have the effect of moving the story along, as if the sleeping body, occasionally stirring into life, directs the course of the dreamer's narrative. Early in the movement, the dreamer says:

Virginia and I ended up in her car, Penelope wasn't with us, and she said, "let's not go to the cafeteria, let's go out to lunch", and I said, "fine".

At this point, the guitar takes on a more prominent role in voicing Virginia's text, and the entry of the female voice in the ensemble somehow suggests the voice of 'the other'. Eventually the instruments overlap with each other, so that the sense of who is speaking becomes blurred. In the next section the recorder and bass clarinet set the extreme pitch range of the instrumental sound, in a repeated motif that is intended as a sonic representation of the engine of the car that dreamer is sitting in. These instruments assume the role of scene painting. Virginia switches the radio on, and the dreamer tries to describe the music she or he is listening to, and the apparent embarrassment felt at having to like what the friend is playing for her or him.

She liked it, but I didn't, but I couldn't say a whole lot, and I said "oh".

The music here switches from voicing what the dreamer's persona and friend are saying, to suggesting the possible music that is heard from the car radio, over the sound of the engine.

It was not a song I have ever heard before, it was an instrumental, it was very strange, it almost sounded like sitar music...

The narrative twist in the dream comes when she or he hears the same song twice, first in a major key, then in a minor key. This assumes the dreamer to have some kind of musical understanding. The significance of the change of key, which seems to be a crucial one for the dreamer, is reinforced in the music as a hypothetical version of the 'Indian-like' music that might be playing on the car radio.

62

Figure 40 Page 62 of *Dreams of the Blind*, partially showing 'major' key version of the hypothetical song playing on the car radio.



Figure 41 Page 68 of *Dreams of the Blind*, showing 'minor' key version of the hypothetical song playing on the car radio.

As the dream comes to an end, and the music is stripped down to the bare electronic drone, the dreamer, now reflecting on the dream, reminds us that she/he is blind:

The senses I used were hearing and touch, I could feel myself sitting in the seat and walking in the hall, earlier in the cafeteria, I could feel my shoes on the hard floor.

This reminds us that any images we might have seen with our mind's eye, as an audience reading the text, are very far removed from the world of sensations the dreamer occupies. We remain in the realm of sound, because images and seeing are never suggested, even though, as an audience, the way we gain insight into this unsighted world is, ironically, through our eyes.

This reminder of the blindness of the dreamer, also comes to the fore in Part V: *Winter Funeral*, where the dreamer leaves a funeral ceremony, where she apparently had been playing the organ, accompanied by an unknown man. Her fear and awkwardness about the situation seem to escalate as the man tells her to start running, as if they are being chased by someone. This reaches a climax when she spots the headlights of a car coming towards her. This acknowledgement of a light source, of the sense of sight acting as a signal in her dream, is puzzling. We know

she is blind, but in her dream, the visual senses seem to be partly intact; though it is also clearly connected to fear:

...and the headlights of the car kept coming closer and closer, and I woke up terrified and shaking like a leaf.

Of all the movements, this one sets up a greater divergence between the electronic soundtrack and the ensemble. The electronics are harsh, noisy and grating, without much connection to what the ensemble plays, which is in turn based on a chorale-like harmonic motif.

The only connecting point between the electronics and the text is a metaphoric one, in that the sound seems to suggest a landscape of wind and rain, which is referenced in the text (the trees in the dream, just like the dreamer, are 'shaking like a leaf'). The chorale-like motif developed by the ensemble, takes as a cue the fact that the dreamer had just been playing the organ at a funeral. Her thoughts about what she had been playing, were perhaps the driving force for the guilt feelings that permeate the dream. She says earlier on:

I was playing an organ at a funeral, and I thought it went very well, but apparently someone was very unhappy about it, because a man, and I have no idea who, left the funeral home with me.

This motif develops and intensifies together with the dream narrative. The voice of the dreamer in this case is connected directly to the organ-like material that the ensemble is playing, even though like in Part III: *Hairdresser* the text is not too closely synchronised to the music, but seems to unfold in its own pace. There is at some point a discrepancy between the fast pace of the music, which one can interpret as the increased emotion of the dreamer, and the indifferent pacing of the text. This serves to ratchet up the tension by increasing the distance between the events that are happening in the dream and the emotions of the dreamer, an increase in loss of control. The voice of the dreamer is split between what she is recounting and what she is feeling. By the end of the movement, the line 'shaking like a leaf' is accompanied by a low cluster on bass clarinet, trombone and contrabass, that seems to finally match the low drone that has been underpinning the soundtrack; as if the dream narrative is finally converging with the dreamer's emotional state.

Dreams of the Blind is representative of one of my early music-text-film works, which explores ideas of narrative voice, as exposed in Chapter 2 of this thesis. The dreamer's voice is to be found in the space between the text, the ensemble and the electronics, being mediated and shifting between different levels. The deliberate avoidance of the visual dimension that was implicit in this work, lead me to ponder about the power of withholding certain media to let others come to the fore. I began

to explore this idea in terms of what could be called an 'asymmetrical balance' between the different media – the text, the music and the visual form. The idea of the changing hierarchy of media (as outlined in Chapter 3) is one of the key aspects in maintaining the dynamic shifts in perspective, necessary to avoid one of the media becoming either redundant or too dominant, as well as highlighting the built-in inequality of metaphoric relationships.

5.2 *Mnemonist S*

Mnemonist S is 10 minute work for large ensemble, soundtrack and text-film. It was composed for the ASKO |Schoenberg Ensemble and premiered in Amsterdam in 2007.⁵⁹ A version for 5 instruments was made for California Ear Unit and performed in Los Angeles in 2009.



Figure 42 Stills from *Mnemonist S*.

Mnemonist S refers to and attempts to question some of the arguments of the primary models of music multimedia, namely the synaesthetic model, which was important in the early twentieth century as a way of imagining the symbiosis of sound and image, but insufficient as a way of drawing universal truths about multimedia perception beyond subjective experience.

My fascination with using the recollections of Solomon V. Shereshevskii as described by Aleksandr Luria in his book *The Mind of a Mnemonist*, was down to the

⁵⁹ Link to live recording by ASKO |Schoenberg Ensemble: <https://vimeo.com/13766483>

extraordinary insight into the mind of this savant and synesthete. Shereshevskii (1886-1958) was a Russian journalist, apparently of no extraordinary intelligence, who made his name as a 'mnemonist', someone who could perform, in front of an audience, extraordinary feats of memory. The reason why he came to the attention of psychologist Aleksandr Luria was because of what he described as his 'five-fold' synaesthesia, in which the 'stimulation of one sense would produced a reaction in every other' (Luria 1987:83). Sound would not only produce colour, but it would also trigger taste and feeling; a number or the sound of a word would create a very strong sensation in another mode:

I recognise a word not only by the images it evokes but by a whole complex of feelings that image arouses. It's hard to express... it's not a matter of vision or hearing but some over-all sense I get. Usually I experience a word's taste and weight, and I don't have to make an effort to remember it – the word seems to recall itself. But it's difficult to describe. What I sense is something oily slipping through my hand... or I'm aware of a slight tickling in my left hand caused by a mass of tiny, lightweight points. When that happens I simply remember, without having to make the attempt. (Shereshevskii in Luria 1987: 28)

When performing, Shereshevskii would use these sensations to construct a narrative that would function as memory strategies. Apparently he could remember all his mnemonic associations years later, which is why the texts collected by Luria are so fascinating.⁶⁰ He could recall in exact detail not only the list of meaningless words from performances decades previously, but he would also recall the narratives, the mnemonic images, that helped him remember them. From the point of view of creating a music-text-film, this is very rich material; not only does it deal with a model of synaesthesia that gripped ideologies of early multimedia music (which in my opinion was often misguided in its universalising of colour to sound experience), but it creates a fascinating example of modal transference, that is driven by an overt narrative structure; as if he is, on the fly, reinventing the very function of language and narrative.

In his recollections Shereshevskii gives us a unique glimpse into the subjective workings of his mind, frame by frame. The particular text used in this piece centres around a recollection of the memory system he had used in a performance dating from 1936, in which he was asked to recall a large series of syllable permutations, VA NA MA SA, which were read to him only once. The text suggests the characters and places he had created in his mental theatre, as a way of recalling this seemingly abstract and random series of symbols.

⁶⁰ Shereshevskii is comparable to the character in Jorge Luis Borges' short story *Funes el memorioso* (Funes the Memorious) who, after hitting his head in a riding accident, receives the gift or curse of never being able to forget.

My landlady (MAVA), whose house on Slizkaya Street I stayed at while I was in Warsaw, was leaning out of a window that opened onto a courtyard. With her left hand she was pointing inside, towards the room (NASA) [Russian: nasha, "our"]; while with her right she was making some negative gesture (NAVA) [Yiddish expression of negation] to a Jew, an old-clothes man, who was standing in the yard with a sack slung over his right shoulder. It was as though she was saying to him: "No, nothing for sale" Muvi in Polish means "to speak". As for NASA, I took the Russian nasha as its equivalent, remembering all the while that I was substituting a *sh* for the *s* sound in the original word. Further, just as my landlady was saying "Nasa", an orange ray (an image which characterises the sound 'S' for me) suddenly flashed out. As for NAVA, it means "no" in Latvian.

This is a priceless example of how the workings of an inner voice, constructed from a constant stream of sensory input, directs the imagery of a personal mental theatre. It enters the brain through various modes and is transformed by the particular architecture of a synaesthetic mind, into still frames of memory, that carry both the sensorial and emotional weight of the input medium, while at the same time acting as an encryption of the data itself.

The other source and the inspiration for the four-tone melodic character of the music is 'Simon', a cult electronic memory game from the 1980s. This is a game unit with four large buttons, in red, blue, green and yellow, each connected to a musical tone. The player follows an accumulating sequence of these tones, remembering each sequence that had been played before, until a mistake is made, and game-play ends. The particular 4-tone/4-colour aspect of this game seemed apt in this case, as it perfectly matched the material of Shereshevskii's task at hand, having to memorise the four syllable sequence of: MA VA NA SA.

The music thus unfolds in ever increasing phrases of four note sequences, first heard as an electronic sound in a basic wave form, sine-wave, square-wave, triangle-wave or saw-tooth wave, and then repeated by an instrument or spread over a group of instruments. These patterns would then be overlaid in increasing complexity, up to a maximum of 4 layers at different speeds. This rhythmic complexity is thus underpinned by the fast moving, sometimes too-fast-to-read text of Shereshevskii's mental process. What is interesting to note is that in a musical texture where rhythmic polyphony is complex and the dominant pulse is hard to pin down, the rhythm given by the visual flickering of the text, has a significant effect on how we hear rhythmic patterns in the music. Add to that, the silent vocalisation of the reader/listener and the semantic stresses within the sentence that are implied, and the seemingly neutral and mechanical musical phrases gain a depth and an emotional weight that is absent within the music alone. What is interesting to note, is that just like in *Dreams of the Blind*, (and also in *Memoryscape*, the piece described in the following section), there is a dialogue between the electronic sound and what is played live by the ensemble. In the case of *Mnemonist S*, one can say that it is almost

a dialogue of identical voices, a game of mimesis, because one is an almost direct repetition of the other.

The 10 minute piece divides into two equal parts: The first part which has a relatively straight forward visual language, where each word, centred on the screen, deals with Shereshevskii's initial confusion as to how the memory task could be accomplished. The second part [beginning at 5:05] is a retelling of the successful mnemonic technique he used, characterised by a more complex and multi-layered text animation.

The piece begins with the establishment of the common pulse. The text reinforces both the basic pulse heard in the music and initially the phrase lengths of the repeating 'Simon' pattern. Throughout the course of the first section (up to 0:40), as the phrases in the music grow steadily longer, the text begins to become more independent, as the start and end points do not always line up with those of the music, culminating in the phrase: 'the word seems to recall itself', projected at half tempo. This has an effect of stressing the words, just like the rhetorical device of deliberately slowing down speech to underline meaning. The dynamic between text animation and music is established here: a strong temporal convergence, steadily weaker metaphorical correlation (no illustration of meaning takes place) but with the variation of 'scale' (the size of the on-screen text) to visually modify expression.⁶¹

The next section reinforces the autonomy of the text, by drawing out the temporal correlation even further. 'When that happens I simply remember' is projected at a quarter of the speed of the section before, pausing on the word 'remember' for an even longer time. At the same time this is accompanied by the addition of a second layer of phrases in the music, in a triplet 8th note tempo. The bifurcation of the relative speeds between text and music sets out the basic expressive tension in the work. An oscillation between these two types of sections (indicated in the visuals by whether it is white lettering on black background or black on white) leads to the first non-semantic text interlude [1:45], where the animation itself comes more to the foreground in a visual play with the four syllables that make up Shereshevskii's list of abstract words, MA, VA, NA, SA. Although here the letters themselves don't have any rhythmic potency, the fluctuation of the background colour in time with the musical phrases, and the metaphorical coupling of the four syllables to the four notes and the four colours, gives the section a significant function in showing the relation between the media. From here on, the four colours are used more freely, along with more expressive text size modulations and placement in the projection space.

⁶¹ See Chapter 3.2 for a description of media correlation parameters.

The other aspect that begins to establish itself here, is the relation between three distinct levels of musical material. The mechanical repetitions between the instruments and electronics, the slow waves of static harmony, which help delineate the sections, and the increasing presence of the drum kit. The drum kit part, which begins as a way of reinforcing the points of the phrases, gradually takes on the role of 'voice' in the piece. The call and response phrases between electronics and instruments, aspire to represent something like the background mental activity of the mnemonist. The drum kit, which is underlining these phrases and providing synchronisation points with the projected text, comes increasingly to the foreground. This is enhanced by the fact that it is not always performing an entirely conventional role in providing a steady background beat, but seems to be piecing together phrases in a speech-like manner (see score example in the following page). This process reaches its most dense point at the half-way mark, where the function of the text shifts from describing the performance experience, to describing the world conjured up through Shereshevskii's mnemonic technique. From about 6:20 onwards, even though the musical and visual materials increase in complexity and speed, the drum kit takes a background role, giving the sense that the narration has slipped from a more vocal level to a level of deep thought, the voice withdrawing to an inner rather than outer mental state.

Figure 43 Page 15 (half way) from *Mnemonist S*, showing three distinct layers of material.

Graphically, the main shift that occurs at the half-way point is the establishment of multiple layers of text. Here, like in the music, three distinct levels of text begin to develop. Firstly, the letters of MA-VA-NA-SA are almost always present in the background in some form, implying the root of what is generated above. Secondly, the mnemonic space itself, the narrative that is generated from the association of the three-syllable words, is sometimes depicted as scrolling text, and at other times as pulsing text with specific spatial placement in the frame. Thirdly, Shereshevskii's comments on his thought process, which are not always present, but take over again towards the end of the piece. For instance, at 8:58, all three levels of text are projected.

The commentary, projected word for word in the centre of the frame, is slow enough to read while picking out the odd word from the mnemonic narrative, almost unreadable as it skims past. The perspective of these three layers is always shifting per section, just like the media relations in *Subliminal*⁶², so that one has to readjust hierarchies. Unlike in *Subliminal*, here the text is always the primary medium. Perhaps only in the more abstract sections, where one sees the four syllables floating by, does the music gain some temporary prominence.

Out of all my music-text-film works, this one has always provoked the most extreme reactions from audiences, in terms of testing the limits of their perception of text and music. I would explain this in terms of the tight hold the relation of text and music has on the audience, and the fact that the piece leads one down a path of increasing complexity, until it becomes almost impossible to keep on reading and listening with equal concentration, especially, as it happens, if English is not one's first language. This seems to be a huge disorientating factor, when it comes to the perception of words and music together. I would contend that the reward for a concentrated listening/reading experience of the piece, is the experience of mirroring the speed and peculiarity of Shereshevskii's unique mental process, together with one's own mental faculty being stretched over several parallel perceptions: a simulated synaesthetic joy-ride.

Analysing the media correlation between text and music in *Mnemonicist S* using the method outlined in the previous chapter, one notices the prominence of the synaesthetic approach, because the sensory category tends to converge more than the semantic one. There are certain ambiguities left in the semantic categories for the spectator to piece together. It is not highly divergent, though it may suggest that in order to achieve a balance between the doubling of media on the sensory level, the level of style, story and sentiment has a slightly more neutral connotative relation.

⁶² See Chapter 3.4 for the analysis of the media correlation in *Subliminal: The Lucretian Picnic*.

Media Correlation: *Mnemonist S* (Music & Text)

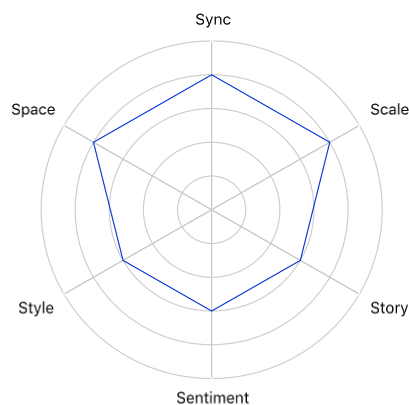


Figure 44 *Mnemonist S*: media correlation.

Sync: 4

There is a high degree of synchronisation in both the metric appearance of the text and the sectional changes, which correspond to sentence or subject shifts in the text. Only the musical phrases deliberately stay out of sync with the language syntax, and increasingly throughout the piece, the speed of text to music multiplies or divides.

Scale: 4

The music is always oscillating between the synth tones on the soundtrack and what is played live by the ensemble. This reinforces an idea of internal and external space. Other than this more metaphorical nuance, the sound which is spatialized (in quad) tends towards an immersive role around the musicians and the audience.

Style: 3

There are two worlds reflected in the piece: the mental space of the mnemonist with his personal memories from early 20th century Russia, and other hand, the early electronics game 'Simon' from the late 1970's. Both are in some way present in the fabric of the electronic soundtrack, the instrumentation of the ensemble, as they are also present in the video, in the way the text is set against the four primary colours. But because of the dissonance between these cultures, there is perhaps a lingering ambiguity about the forced nature of the metaphorical connection.

Story: 3

There are two worlds reflected in the piece: the mental space of the mnemonist with his personal memories from early 20th century Russia, and other hand, the early electronics game 'Simon' from the late 1970's. Both are in some way present in the fabric of the electronic soundtrack, the instrumentation of the ensemble, as they are also present in the video, in the way the text is set against the four primary colours. But because of the dissonance between these cultures, there is perhaps a lingering ambiguity about the forced nature of the metaphorical connection.

Story: 3

The music never illustrates the mnemonist's narrated world, it always stays on the level of the metaphorical image of memory as a game to be played. Where it does follow the general narrative of the text, is that it reinforces a sense of an increasing mental complexity.

Sentiment: 3

A general feeling of stress is the main sentiment shared by the music and text. There are many emotions expressed by the mnemonist in the text that are not reflected or illustrated in the music; the music stays in this respect relatively neutral.

5.3 Memoryscape

Memoryscape was written for the German ensemble musikFabrik and premiered at their WDR concert series in Köln on the 27th February 2010⁶³. A 32 minute work scored for 15 instruments, soundtrack and text-film, based on the concept of early memory.



Figure 45 Screenshots from *Memoryscape*, showing four memories. The difference in the brightness/contrast of some words shows various stages of the projection, as the words randomly fade in or out.

As opposed to the often disorientating speed at which text appears in *Mnemonist S*, *Memoryscape*, another work which deals with the idea of memory, utilises video in which the appearance of the text is rather slow and non-linear.

Around the time that I began working on the piece I had started collecting the early memories of close friends. I would ask them to recall their earliest memory and immediately make an audio recording of them recounting it. Early memories are one of the most personal windows onto a person's psyche, and it is a question most people respond to with a great deal of uncertainty. Not only because of the intimacy that this question exposes, or because the memories often describe a situation of significant emotional impact, but because these memories are elusive and unclear;

⁶³ Link to Premiere performance recording by WDR: <https://vimeo.com/226623184>

and it is sometimes impossible to ascertain if things indeed happened as one actually remembers them.

It is, at times, very difficult to pinpoint what is exactly the earliest memory we have. Often the date has to be inferred by some facts that we know about our life. Nevertheless, there are usually a handful of memories from the ages between 3 and 7, which remain accessible to us into adulthood. I was fascinated by the fragility of these memories, and also the reason why we have access to some memories and not others. It is not so much the fact that we do not have the capacity for memory at an early age, but that much of what we retain in our minds as children, is lost as we move into our teens. Psychologists call this 'childhood amnesia'. (Bauer 2008: 1). For memories to survive into adulthood, sensory information encoded by synapses between our brain cells must be regularly accessed for consolidation purposes. Thus it is no surprise, that most of our earliest memories do involve an event, which has left a significant emotional response, positive or negative.

The underlying backbone of *Memoryscape* is formed by 15 solos, one for each of the instruments of musikFabrik, which are directly based on 15 of these early memories. These solos are influenced by either the narrative content, the speech rhythms or speech melody, or the timbre of the voice recounting them. Thus, an imprint of the rememberer's voice ends up in the contour of the solo material. This ranges from the more direct melodic transcriptions of the voices of, for example, the various string solos, to the more abstract ones, such as the trombone, piano, or clarinet, where the speech has been filtered or reduced in some way. Below is an extract from one of the violin soli, which is more closely derived from the melodic contour of the voice:



Figure 46 Violin II solo from *Memoryscape* [at 12:30].

My first memory is when I was two years old almost three - I was lying on the top of my bunk bed - I was sick - I was usually sick because it was just after the Chernobyl disaster - and we were always sick then - and I was lying on my bed - and asking my mum every morning 'is it my birthday? Am I already three years old?' - because I wanted to grow up very fast. (R.V's memory from *Memoryscape*)

The idea behind the characterisation of the solo material was to create some personality or identity connected to the instrument, that would relate, in some metaphorical way, with the memory material. So for instance, the clarinet solo, which is based on a memory of a bee, takes on the character of a circling, buzzing pattern, that shifts to the contours of the speech.



Figure 47 Clarinet solo from *Memoryscape* [at 15:27].

I remember the sun is shining - it felt summery because I could hear my dad outside mowing the lawn - and then this bee started buzzing around my head - and I remember just listening to it and wondering what was it - and what was it going to do - a very distinct memory of this bee buzzing round my ears. (B.E's memory from *Memoryscape*)

In another more traumatic memory, the piano solo, takes the contour of the voice quantized into 32nd notes, where the melodic contour expands into different registers, hardly leaving an audible trace of voice, which is only apparent in its structure. The solo tries to depict in the way the pianist has to constantly reach out to the extreme registers, mirroring the sense of panic depicted in the memory.



Figure 48 Piano solo from *Memoryscape* [at 9:25].

My first memory is when I almost drowned - I remember being on a lake on a campsite at my aunt's camping ground - and I was floating on an air mattress - the air mattress tipped over - and I was sinking towards the bottom of the lake -looking up and seeing the sunlight playing through the water - when all of a sudden I felt somebody grab my arm - and yank me out of the water - it turned out to be my dad- who saved me (P.S's memory from *Memoryscape*)

A different type of virtuosity is required for the tuba solo, where the score, based on a close transcription of speech, does not necessarily sound as written, because the Tuba player is required to play using a baritone sax mouthpiece, which entirely alters the logic of the valve system. The grotesque nature of the sound in this solo relates to the subject of the memory:



Figure 49 Tuba solo from *Memoryscape* [at 18:35].

I think I would have been around three - it was a birthday party either mine or my sister's - I went to the living room - I might have just gone to have a pee - and I then watched TV and there was this weird programme on - never forget it - this is where this black weird character comes from - in these recurring dreams I had when I was a kid - there was a programme on TV - and it had Big Daddy the wrestler on it - and I remember watching Big Daddy and he was against this guy - who was totally black and skinny like Spiderman - and there was this horrible scene and it stuck with me for ages - where he's back stage - and he lifts this black mask of his face - and underneath is just what can only be described as crunchy peanut butter - and I had this recurrent nasty dream based on this character for about four years. (A.C's memory from *Memoryscape*)

The electronics in *Memoryscape* play an ever-present role in framing the ensemble, and also taking some of the sonic identity of the solo material and transforming it. Sound particles that emerge at some point in the work appear then in others, as if an imprint remains on the sonic fabric, which is sustained for the duration of the work. These temporally and spectrally transformed imprints of acoustic sound form the basis of the electroacoustic landscape. Memories of the ensemble heard at different points, are stretched into slowly evolving layers of sonic matter, creating an immersive field of sound, that is somehow meant to represent the polyphony of reminiscing voices, an omnipresent landscape, that steadily grows throughout the piece. Similar to *Mnemonist S*, there is a relation between live acoustic sound and soundtrack in an ever-changing dynamic. There is a three-fold figure and ground perspective created between solo instruments, ensemble and soundtracks, which share similar material, that is passed between the levels in different scales of magnification. The soundtrack stays mainly in the background, though in the opening and closing passages, and at various times throughout, it comes to the fore and blends with or alternatively swamps what is being heard live.

At the time I wrote this piece, I was experimenting with a compositional process that had some resonances with the way memory is shaped and transformed into sensory information. I would begin by trying to recall a sound pattern that would come to mind. I would then encode it as musical information, either in notation for instruments, or by compiling layers of electronic sound. Later I would constantly revisit this pattern by placing it in a temporal frame and repeating it, reassessing and updating this memory and transforming it, until it either disappeared, consolidated into something static, or evolved into another state entirely. This would happen in an intuitive rather than a formal process, where focus would rest on a particular detail, which might become magnified, or where a particular subjective way of seeing a pattern would be exaggerated. At times this transformation would be very minor, at others it would turn into something entirely different from where it began. Sometimes it might even split into contrapuntal branches, by exploring various possible variations of an idea simultaneously. This latter multi-layered approach,

suited my musical preferences at the time, for a type of polyphony, which encompassed radically different time scales, though originating in one musical idea. For example, a material developing very slowly in one direction, while another, could be moving rapidly and almost imperceptibly in an anterior direction. Another example could be the way of combining electronic and acoustic sounds, which at first merge, and then diverge. Below is an example of the piano, percussion and string parts near the beginning of the piece, with phrases slowly evolving over the first solo (viola), articulating the memory: 'coming down the stairs...' (see figure above)

Coming down the stairs crying - because I'd just been playing with this neighbour - we were very good friends - a boy my age - and I just told him that his father is shit - and he'd gone to his father immediately - and told him what I'd said - and I think - the father walked into the room in which we were playing and told me to go home which I did of course (A.H.'s memory from *Memoryscape*).

The musical score is divided into two systems. The first system covers measures 29 to 36, with a time signature of 1:20. The second system covers measures 37 to 44, with a time signature of 2:00. The instruments are listed on the left: sd (soprano), pm (piano), perc (percussion), v1 (violin 1), v2 (violin 2), via (viola), vcl (violin), and cb (cello). The piano part features a solo in the viola, marked 'SOLO' and 'on the string - full bow pressure - speech like'. The percussion part includes a high woodblock and a shinghol. The strings provide a harmonic background.

Figure 50 Excerpt from the piano, percussion and string parts [1:20-2:16] of *Memoryscape*.

Figure 51 Continued excerpt from the piano, percussion and string parts [2:16-3:18] of *Memoryscape*.

There is an added theatrical element in *Memoryscape*, in that the musicians only enter the stage at the moment just before they play their solo. They take one of two positions at the front of the stage, and then proceed to take their proper place in the ensemble, which is placed in a diagonal formation. They remain there until almost the end of the piece, when gradually, one by one, they exit the stage. The piece, therefore, begins and ends with an empty stage, amid the backdrop of the sonic landscape. It was my intention that this action would reinforce the transience of the moments described in the memories. A collective remembering, which emerges for a short time and then collectively vanishes.

The text-film in *Memoryscape*, always appears as one slide (as shown in the figure at the top of the section). The words fade in gradually (the speed depending on the length of the solo), in a random order, and then fade away again, in a random order. Because the text is a transcription of something that was spoken spontaneously, it does not tend to fall into formal sentence structure. I chose therefore to highlight the sequence of phrases by putting them each into a single line, and changing the font size and spacing to accommodate the different sizes on each slide. This makes it easier to read, as the eye picks up single phrases as they appear on-screen, though

because of the random order of build-up, the whole memory is probably not read at first in a linear fashion, the mind, perhaps only later, piecing these parts together.

The gradual materialisation of text, as read by the inner voice of the audience, could be said to exist in a space between ensemble and electronics; between the solo instruments, which represent the voice of the rememberer, and the soundscape which, in my mind, represents their embodied thoughts and feelings. Because the first person narratives in this piece are understood as being voiced by the soloists, sonified in instrumental form, the audience one can easily project the voice that is being read onto the instrument that is soloing, especially as in most of the cases they are at the front of the stage. The focus is relatively clear: one narrative, one instrument. What is less clear, and hopefully adds to the poetry of the piece, is what happens to the traces of these voices, and their reflection in the virtual space of the electronics, once the individual soloists enter the collective of the ensemble. This is the 'memoryscape' that is referred in the title of the piece. A hypothetical collective space, made up of traces of people's earliest memories - transitory, disordered and raw.

The three pieces cited above, dealing with narratives of dream and memory, highlight an aspect of inner voice discussed in Chapter 1, and a form of narration described in Chapter 2. These are essentially first-person narratives that toy with a sense of the 'focus' potential of the voice. This is because there is always some ambiguity in how the words on the text-film shift between the ensemble, the electronics and the audience's own inner reading. Each of the three works utilises different correlations between electronics and instrumental music, in which the relation is never static, resulting in an ever-shifting focus. The 'asymmetrical balance' between the media helps generate this dynamic shift in perspective.

From an analytical point of view, what becomes apparent in describing the relation of text to music in these works, is that there seems to be an important three-way correlation. Not necessarily between text, music and image, but between text, soundtrack and ensemble. The narrative voice that is summoned by the projected text, seems sometimes to align itself more to the electronics and at other times more to the ensemble. This has a huge effect on how the voice is eventually perceived by the audience. The voice moves and attaches itself to different aspects of the sonic environment, thus defining its character and becoming alternatively distant or immersive (depending on how much it connects with our own voices). Whether the text correlates to either sound medium on a metaphorical or a perceptual level, influences the 'what' and the 'how' that is being communicated.

Chapter 6: Unanswered Questions

In the following section I focus on three of my works which explore question and answer structures found in language. They also incidentally address the question whether music does indeed function in the same way as language. The works in question are *Machine Read*, a short video piece made for a presentation of Terras magazine as a homage to Dick Raaijmakers' essay "De kunst van het machine lezen", and two sound installation works, *Dodona* and *Norms of Transposition (Citizenship)*. What is common in all three works is that the question and answer structure spans different media, where a question is either translated from words to music and vice versa, going through other media in the process. These pieces take on a playful perspective on the idea of music as a surrogate of language, by using encoding methods that are both absurd and opaque, and which, certainly in the case of the two installation pieces, try to use musical or sonic aspects to break down the coherence of language.

The intriguing relation of music to language has been the subject of much speculation in studies spanning anthropology to neuroscience, and the question: whether music is indeed a language has been asked in many different ways. While a detailed answer to this question has first to deal with a clarification of each domain's specific semiotic structure, a simple answer might point to the many shared similarities. In evolutionary theory, some have argued that music and language stem from the same evolutionary stage, what psychologist Steven Brown has coined 'musilanguage'. (Brown 2005: 271)

Many recent studies in neuropsychology have shown the brain's shared subcortical processing of music and language, whether spoken or written (Besson & Schön 2003: 269). Whether we accept that music is primarily communicating as a language or simply resembles language is something that will continue to be argued for years to come, and will largely depend on how we define music and how we define language. As a composer and sound artist, this discussion has been a source of fascination for me, precisely because of the understanding that music can be made to behave as a language depending on the context in which it is used. In the last years, I have focused on a body of work that deals with the issue of how music communicates, how musical structure can function as a language, and the translation of semiotic information from one medium to another. This interest grew out of a piece, *a conSPIracy cantata*, based on so-called 'number stations', which I wrote in the late 1990s.⁶⁴ These number stations are shortwave radio signals transmitted by government agencies, such as CIA, M16, and Mossad, to message

⁶⁴ CD release: Yannis Kyriakides, *a conSPIracy cantata*, Unsounds 01U.

agents in the field. Some of these number stations have a seductive musicality, using tones, pitches, and even tunes to carry the message. Background research reveals that the use of music in cryptology and the transmission of secret messages has an illustrious history, stretching back from the careful deployment of church bells by the Vatican (Sams 1980:80), to the invention of the vocoder⁶⁵ and its use in messaging across the Atlantic during the Second World War.

The possibility of using musical structures as a secret language, utilising music as a surrogate for language, also drew in many classical composers throughout the centuries: Giuseppe Tartini, Georg Philipp Telemann, Robert Schumann, Michael Haydn and Edward Elgar (amongst others) were all in some way fascinated by the encoding of messages into music (Sams 1980: 80). From an ideological perspective, there have been several attempts to create a 'universal language' out of musical tones, that would be "equally speakable by all nations and people" (Wilkins 1970).

The philosopher Gottfried Wilhelm Leibniz proposed something similar using a binary system that encodes language into pitch intervals (Woolhouse 1994: 439). A successful example of an artificial language based on music was Francois Sudre's 'Solr  sol', developed and used at a school for the blind in France in the early 19th century, which consisted of musical formulas of up to six notes with specific stress placed on different parts of the phrase, resulting in a vocabulary of about 2,500 words.⁶⁶ These ideas eventually lead to the development of Morse code, probably the most widely used encoding form of language, resembling music in its utilization of binary difference to create seemingly complex rhythmic structure.

Can these examples of text-to-music be considered as 'translations' in terms of the common practice, that seeks an equivalent of source-language in target-language? Here, the problem lies not so much in the idea of translation, conversion, or encryption, but in whether we can accept that the target-language is, in fact, a language. In music, the rhetorical device of the question and answer structure operates as the basis for many traditional phrase constructions: two consecutive phrases, the first ending on a weak cadence, the antecedent, the second on a strong cadence, the consequent, give a sense of a question and its answer. There are countless musical examples of how this structure is used, from call and response forms in popular music, to larger scale structural antiphony found in the symphonic form. One interesting example is Charles Ives' *The Unanswered Question* (1908), alluded to in the title of this essay. In this programmatic work, a solo trumpet poses 'The Perennial Question of Existence' seven times over a muted string chorale, which represent the 'The Silence of the Druids - Who Know, See, and Hear Nothing.' A

⁶⁵ For a fascinating history of the vocoder, see: Dave Tompkins, *How to Wreck a Nice Beach*, Chicago: Stop Smiling Media, 2010.

⁶⁶ Jean-Fran  ois Sudre, *Langue Musicale Universelle*, 1866.

woodwind quartet representing the 'Fighting Answerers,' answer each time, but unsatisfactorily:

'The Fighting Answerers', as the time goes on, and after a 'secret conference', seem to realize a futility, and begin to mock 'The Question' - the strife is over for the moment. After they disappear, 'The Question' is asked for the last time, and 'The Silences' are heard beyond in 'Undisturbed Solitude'.⁶⁷

What interests me about this work is the way in which Ives uses the technique of musical collage to highlight different levels of discourse, or, perhaps, different language systems that render the attempt of answering the question impossible. This idea of the problematisation, of 'the question' is the concern at the heart of the two works discussed below. A situation set up in one medium is given an answer in another as a way of highlighting the problematic aspect of the question itself, and music's incapacity to function in exactly the same way as language in this regard. In both cases, questions are posed in words, input into a system that communicates them in a musical structure, and are then translated back into words. In all three cases, the idea of power structures, either in the posing of questions (as in *Norms of Transposition*) or the answering of questions (as in *Dodona and Machine Read*) are critiqued by the process of musical translation, where the ambiguity of meaning is used to highlight the specific nature of the initial questions.

6.1 *Machine Read*

Machine Read is a 5 minute fixed media music-text-film, that was made for an event in 2012 organised by Terras Magazine⁶⁸ as a celebration of Dick Raaijmakers' 1978 essay "De kunst van het machine lezen" (The art of reading machines).

The piece is based on texts from section 100-102 of Raaijmaker's essay, in which he discusses the tape recorder, typewriter and piano in the context of language, where he tries to find the common factors in their function. Raaijmakers' essay analyses the concept of the machine from the point of view of how it is used, thus arriving at an understanding of its status and historical function. Typically for Raaijmakers, he takes a detached point of view, and tries to observe how each machine affects our ways of perceiving the world. What he tries to articulate in the context of technology, is the relationship between *form*, which is the result of an action, and *idea*, inherently expressed in the technology itself. He draws parallels between disparate media, as in the case quoted here, where a piano, a typewriter and a tape recorder are shown to

⁶⁷ From Charles Ives' explanatory note in the score of *The Unanswered Question*, Miami, FL.: Southern Music Publishing Co. Inc, 1908.

⁶⁸ Link to online version: <http://tijdschriftterras.nl/a-reflection-ideas-dick-raaijmakers/>

share similar attributes. He binds these machines through the metaphor of language and suggest, that just as the typewriter is a writing machine, and the tape recorder is a reading machine (even when it records, he says, it "reads whatever comes into the input"), the piano falls somewhere in a category in between.

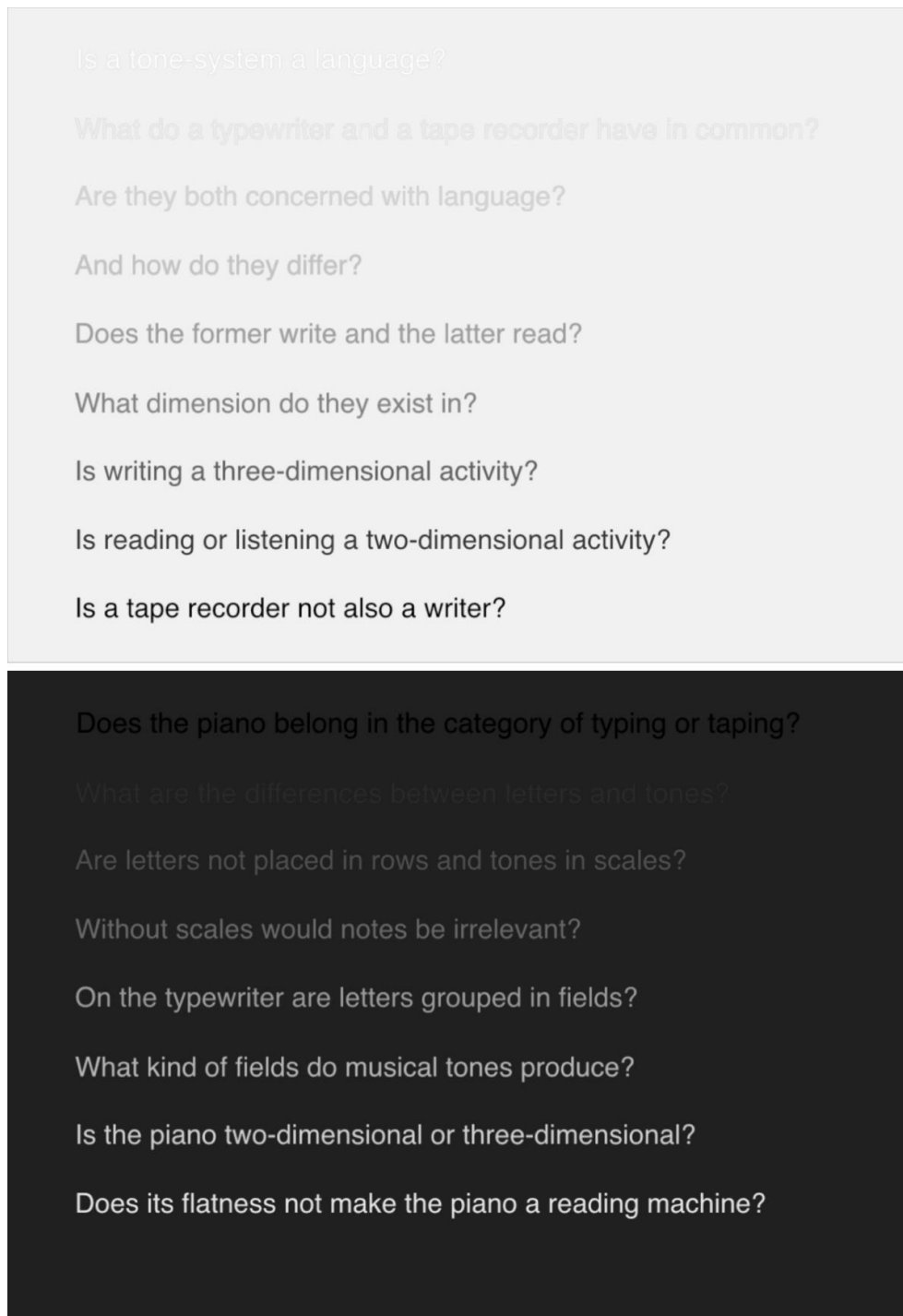


Figure 52 Screenshots from *Machine Read*, showing the list of questions from both parts of the piece.

over klok opzij, van links naar rechts:
NB "De kunst van het klok-lezen"!

hij heeft de klok
opzij en omkeert.
Maakt de klok
brek.

99 Ruimte-beeld

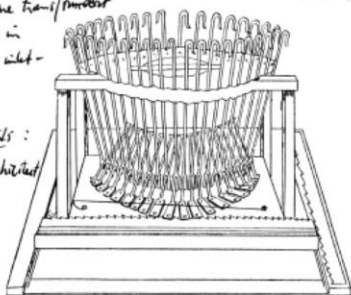
Wanneer machines samengesteld zijn uit meer delen, of seriematige handelingen verrichten, of min of meer gekompliceerde operationele handelingen, dan verandert hun voorkomen in de ruimte. Het BEELD hiervan zal dan op zijn minst iets moeten zeggen over de taakverdeling van de componenten binnen de machine en op welke wijze ze dit realiseren.

Met voorbijgaan aan allerlei obskuraantistische methoden om dergelijke machines perspectiefisch, fotografisch of holografisch in beeld te brengen, moet het BEELD - wil het op een juiste wijze gelezen kunnen worden - aan bepaalde voorwaarden voldoen. Om hieromtrent enkele criteria te kunnen opstellen vergelijken we een schrijfmachine, een bandrecorder en een piano, en vervolgens een piano en een orgel met elkaar en wegen hun (verschillende) verschijningsvormen tegen elkaar af.

100 Taal (1) *moese! telegram linst pas op!*

Een schrijfmachine en een bandrecorder komen in zoverre met elkaar overeen dat ze allebei met 'taal' van doen hebben. Zij verschillen echter hierin, dat de eerste *schrijft* en de tweede *leest*. Schrijven bij een schrijfmachine betekent, het realiseren van een schrijfplan via een aantal 'armslagen' in gedrukt schrift: machineschrijven is dan ook een activiteit die zich - net als sport - in de ruimte afspeelt. Lezen zoals een bandrecorder dat doet is daarentegen een eendimensionale activiteit, zoals uit de eenparig voortbewegende tape valt op te maken. Het argument dat een recorder toch ook schrijft (immers hij heet niet voor niets 'recorder') berust op een waarnemingsfout: hij *leest* wat aan zijn ingang verschijnt, en *leest* die - ogenblikkelijk of later - hardop voor. De konsekwentie van dit alles is dat de bouw van een schrijfmachine *ruimtelijk* is (- móet zijn) en van een bandrecorder *plat* (- kán zijn).

Een schrijfmachine transporteert
langzaam bewegende
hamslagen in richt-
schijf -
Een bandrecorder
transporteert hams:
die ingekant dekkings-
te reproduceert
dekkingsbest!



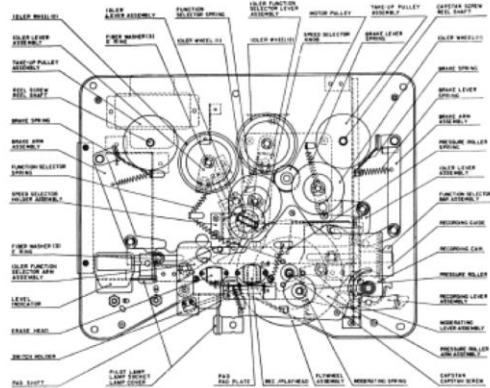
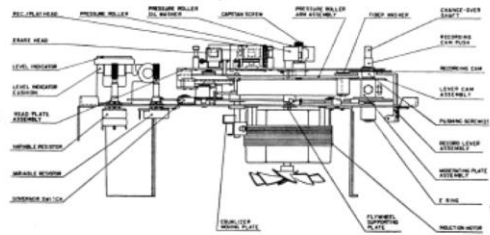
'Plume Typographique', Xavier Progin, 1833. Het willen schrijven met behulp van een machine betekent de derde dimensie in huis halen. Immers een hamslag vraagt nu eenmaal meer 'ruimte' dan een penhastreek.

hij bevestigt de signalen om van zijn gang te verschuiven, te spreken en - ogenblikkelijk of later - hardop na.
Een schrijfmachine transporteert (hamslag in hamslag in schijf);
Een bandrecorder transporteert (hamslag in hamslag in schijf);

101 Taal (2)

De platheid van een bandrecorder blijkt pas goed door zijn tape-transport-systeem van boven te bekijken en dat beeld te vergelijken met dat van opzij. Immers door zijn aard komt het bewegings-systeem - dat voornamelijk uit een stel van wielen, rollen, snaren en assen is opgebouwd - overeen met het wijzer-transport-systeem van een klok, waarvan de raderen en tandwielen van opzij bekeken lijken op 'smalle balken' die dicht op elkaar en tegen elkaar gestapeld zijn, en waarin geen beweging te ontdekken is.

Een klok 'leest' men dan ook niet van opzij. Een van-boven. En zo staat het ook met de bandrecorder: van-opzij heeft het aan ons geen boodschap.



Hitachi bandrecorder, 1965, van-boven bekeken

Hitachi bandrecorder, 1965, van-opzij bekeken

Net zoals bij het transport van een pen over papier vindt het transport van tape in een plat vlak plaats. Het tape-transport-systeem bezit dan ook van-opzij bekeken geen feitelijke dikte, temeer daar de breedte van tape, evenals de dikte van een lijn, slechts aan kwalitatieve maatstaven hoeft te voldoen. In dit verband ziet het tape-transport-systeem van-boven bekeken er uit als een route gemarkeerd door druk- en glijpunten langs schijven en assen.

NB. In typetijl is de hamslag van boven is een bandrecorder in hams, niet symmetrisch!

De hamslag is een schrijfmachine van-voor niet symmetrisch
- niet van boven! (goed te herkennen.)
Schrijfmachine is een individueel op de manier stand
opgesteld.

opnieuw
pas op!
in hamslag
in hamslag
(Maakt hamslag
p.100)
De hamslag is
niet te herkennen
van! De
hamslag is
niet.
Dus: een hamslag
opzij is niet
te herkennen
Fels is de hamslag
5 opzij bekeken
niet te herkennen
in hamslag
hamslag is
hamslag is
hamslag is

Van opzij en bekeken

Van opzij en bekeken

33

Figure 53 Page from the original publication of Dick Raaijmakers' "De kunst van het machine lezen", with his own annotations.

My piece takes some statements that are made in this text, either directly or implicitly, and converts them into questions:

Is a tone-system a language?

What do a typewriter and a tape recorder have in common?

Are they both concerned with language?

And how do they differ?

(First four questions from *Machine Read*).

These questions are then posed to 'Siri' (the personal voice assistant of Apple's iOS technology), and the answers are either played back using Siri's voice (in the first part of the piece) or encoded into virtual piano notes (in the second part of the piece). In the case of the questions, the exact opposite happens. In the first part, they are encoded into piano notes, as the letters appear, typewriter style, while in the second part, Siri takes on the role of questioner⁶⁹. Using the technology of voice recognition, and all the mistakes that occur with it - ('reading' becomes 'Reading, England'), a connection is made between the typewriter and the piano. One could even say that the tape recorder, whose sound also lurks in the background of the piece, is a forerunner, in certain aspects, of the personal voice recognition system, in that the sound is 'read' into the machine and converted into code. Language is recognised by its sonic patterns and converted into another modality (I can only wonder what Raaijmakers would have made of smart phone technology?).

The main underlying question, that is posed both in the piece and in this part of Raaijmakers' essay, is whether music or a tone-system can be considered a language. This question is posed twice in the piece, at the beginning, and at the end,⁷⁰ and both times remains unanswered; because there can only be a metaphorical answer to this metaphorical proposition. Music is language, in so far as attributes of language can be aligned with attributes of music. This metaphor is re-proposed by Raaijmakers, as a typewriter is a tape-recorder is a piano. In the piece, the voice, like in many other pieces cited in the thesis, acts as a go-between across these two domains. But its role here is a complicated one, as *Machine Read* plays on the idea of unequal superimposition of voices. In the first part, piano + type-written text (the inner voice of reader) is answered by computer voice. In the second part, computer voice + type-written text is answered by piano. There is always a doubling of voice in the question, which gives the answer a less emphatic tone, more ambiguity; certainly in the second part where the piano is clearly insufficient in conveying any semantic information, other than the length of the answer phrases, or the emotional effect communicated by the widening of registers. This, in effect, stands for an answer to

⁶⁹ In practice, this second part, involved me typing out the questions on a computer, and having them read by the same computer voice (Daniel) that was used in Siri's first generation voice. The answers on the other hand were the genuine answers given by Siri, when posed through an iPhone.

⁷⁰ In the installation form, the loop reinforces this question as it is posed consecutively.

the question of whether a tone-system is a language, but the question remains, does this tell us anything more useful than Siri's answers in the first part?

The encoding system used in *Machine Read* is based on the Polybios cipher system that I have used in many of my piece, from the viola and contrabass piece *As They Step* (2007), to the book encoded in *Politicus* (2011). It is a 5x5 grid, where each letter is given two values, a coordinate. In this piece, it is mapped onto two sets of unequally tempered pentachords, which are asymmetrically aligned across the whole keyboard. They are heard in the piece, one in each side of the stereo field. The discrepancy between music and language is also reflected in the differences between what is visualised and what is heard. There are small differences highlighted in how the question is visualised or communicated by voice/piano. For instance, in the first part, towards the middle, a question is posed which because of the changing gradient of greyscale, is impossible to immediately read. Siri responds: "I cannot read your incoming question". When a similar situation happens in the second part, with inverted greyscale, the question is read as normal, even though we experience a temporary loss of inner reading voice (which hopefully highlights the background activity of the inner voice due to its absence).

There are two connected aspects, which I have mentioned though not addressed fully. The correlation made between the background tape hiss, which towards the middle of the piece starts resonating to the frequencies heard in the virtual piano, and the gradual change of background colour, from black to white and back again. The whiter the background, the more resonant the tape hiss. Similarly, each new line that is typed has a change of colour tone from white to black, which is then reversed in the second part. This cyclical form is meant to highlight the sense of a gradual movement from one part of the metaphor to the other, as if the movement from music to language, from question to answer, from piano to typewriter is a smooth one (which it can never be). The tape hiss also stands for an additional thing, which is hinted at in the beginning. After the second question Siri says: 'Let me think', after which there is a long pause. During this time our attention shifts to the sound of the tape, as if the noise represents, in some way, either the machine's thinking process, or our own, as we try and make sense of the strange metaphor that is proposed. The gradual modulation of the hiss, from background noise to a resonant harmony texture at the mid-point, feels like a significant emotional moment, like an illumination almost, especially as it is accompanied by a brightening of the background colour. What I wanted to convey in this moment is the idea of background mental activity gradually becoming voice, entering the harmonic space of the piano, connecting with what has been up until now, the surrogate voice; in Raaijmakers' terms, the tape-recorder becoming reader.

6.2 Dodona

Dodona is an audiovisual installation first shown at the SPOR Festival in Aarhus, Denmark in May 2013. It consists of a large suspended metal sheet mounted with two transducers, a microphone, one or more fans, speakers, a beamer, lights, a computer and controlling software.

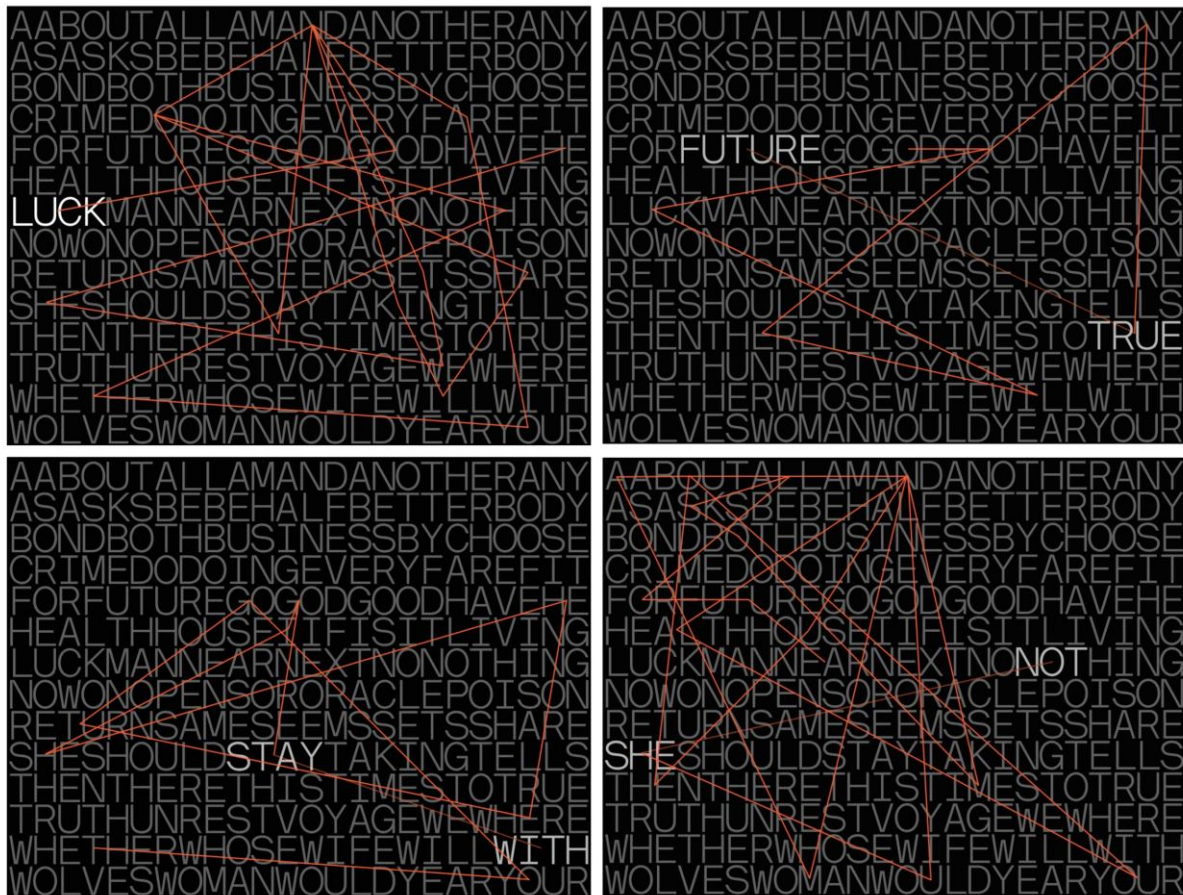


Figure 54 Screenshots from *Dodona*, showing the questions encoded as geometric shapes.

Questions appear, projected on one side of a metal sheet, hung in the middle of a room. They appear as lines joining words set out on a letter grid. These lines are drawn accompanied by white noise, which is filtered according to the position of the word on the grid. This noise is played back through transducers on the metal sheet. Once the question is complete, the words fade, leaving the image of a geometrical shape, the graphic representation of the question's path.

At this moment, on the other side of the metal sheet, lights turn on to reveal a microphone on a stand and a fan being switched on and off, projecting air onto the microphone at various rates. This is amplified through the speakers in the corner of the room. On the metal sheet, words begin to appear at the corners of the geometric

shape, rapidly flickering in correspondence with the undulations of the air in the microphone, and their changing size directly correlating to the loudness of the air's sound. So the question, which has been encoded as voltage, changed and sent to the fans, receives an answer from the translation of the sound of the air hitting the microphone - a sound generated flickering of words.

This self-answering oracle machine is a reference, as the title suggests, to the ancient oracle of Dodona in northwest Greece. Many different accounts exist detailing what the divination process consisted of; talking birds, the rustling branches of an oak tree, wind, flowing water and the sound of copper vessels vibrating, have all been mentioned by classical writers such as Homer, Xenophon, and Herodotus (Eidinow 2007:71). What is certain is that questions were written on lead tablets, hundreds of which have been excavated around the site of Dodona in the last century. These tablets, dating from the 6th century to the 2nd century BCE, contain questions raised by all casts of society, such as personal consultations regarding emotional insecurities, resolving crimes or questions about citizenship:

Would I do better . . . if I took a wife?

Did he (or she) introduce a poison (or potion) to my children, or to my wife or me..?

Shall I request citizenship this year or next? (Eidinow 2007: 71)

Most of the answers have been lost, while some remain written on the back of the lead tablet. Though if we should follow the example of some oracular pronouncements from the more famous oracle at Delphi, it is unclear whether the answers were concrete or ambiguous. Nevertheless, part of the skill required in receiving a useful answer from the oracle relied upon the formulation of the 'right' questions. In the third book of the *Anabasis* Xenophon recounts how Socrates gave him advice on the matter. Instructing him to frame the question as an either/or so that he can get a clearer response:

Should I do x or should I do y? (Eidinow 2007: 71)

Dodona plays on the idea of the finite, or quantifiable aspect of a question. A grid of about 80 words written without spaces to separate them, forms the total vocabulary available for what can be asked. This imposition of limits upon the enquiry is reinforced by the reduction of the question into a geometric form. The form is constituted of a sequence of lines, moving from one centre of a word to another. This is reduced to a set of Cartesian coordinates that represent the basis for the question's translation both into filtered noise and, subsequently, into the voltage changes that drive the fans. The fact that the question is translated through several different media, mapped onto coordinates, visualized, sonified and converted into voltages, is almost akin to an energy conversion model in physics—where, rather than Joules being lost in the conversion, it is meaning that falls away. The metaphor of the wind

as the all-knowing voice of the oracle is heard as air blowing on a microphone. Yet the complexity and power of the resulting noise still remain captivating. The chaos is filtered and translated through the algorithm into a semblance of meaning, as a recognisable pattern emerges from the spinning words.

Finally, the translation of the sound back into text, happens through a computer algorithm, based in part on the vocabulary that appears on the grid. Initially I had wanted to use a commercial speech-to-text algorithm, that would convert the clouds of noise into intelligible text, but being so far from recognisable speech, only errors ensued. Thus an algorithm was created, that would analyse the pitch and volume of the incoming sound and map it, using a similar grid type system as the initial conversion, onto a dictionary of words. The louder the sound the larger the size of the printed word, illustrating a direct correlation of text production to incoming sound. The words would spin very fast on the nodes of the geometric shape, and eventually come to rest on a selected number of words, the oracle answer, as it were. Because of the slight difference in the air currents and the small differences in the pitch and volume detection, which would lead to totally different set of resulting words, each repetition of the cycle would, in practice, give totally different answers to the same questions. This perceived arbitrariness of the answers is not a totally unwanted element of the piece, it is in fact of the essence of the piece; that noise translated into a recognisable pattern will always entail a very subjective or arbitrary filtering of information.

A question addressed to an oracle is essentially a question addressed to one's self, projected onto an external system. The first step is its formulation into language—the translation of feelings into words. Subsequent steps involve mostly arbitrary processes, be it the pattern in tea leaves, the sorting of yarrow stalks, or the position of stars. What I wanted to highlight in this self-enclosed oracle machine, regarding the translation of text to sound and back again, is how the process of finding meaning in sound always depends on the mediation—the way in which the hierarchy of media is set up in the construction of a metaphorical exchange. Applying the metaphor model described in Chapter 3 to the analysis of *Dodona*: in the question phase, the text can be said to be the target domain, and its sonification and visualisation become source material in the construction of its meaning. In the answer phase, the sound becomes the focus, the target domain, and the resulting text, because it lacks clear coherence, remains only as a source of possible meanings.

6.3 Norms of Transposition (Citizenship)

Norms of Transposition (Citizenship) is an interactive audiovisual installation created for the exhibition *In Crisis*, curated by Yiannis Toumazis shown at the Municipal Art

Gallery, Nicosia (July 2012 until August 2013). The installation consists of a cubicle, a chair, a table, a screen, and a small piano-keyboard.

A civil servant at the immigration department has been of great help to you. How do you thank them?	You enter the room for a job interview. There are two people behind a desk. What do you do?
fight small soul understand game sing breathe breathe breathe	do far last fill water last always to song guy sorry stare room
You are stopped by a policeman, who request to see your identity card. You do not have it with you. What do you do?	You have lived here for five years. You would like full citizenship. Who can give you information?
notice point down open found much gone fly such everyone	morning more sleep we sun that could could even word child whisper

Figure 44 Four screenshots from *Norms of Transposition (Citizenship)*, showing the questions and possible answers.

Unlike the self-addressed oracle questions of *Dodona*, the type of questions raised in this piece, are addressed from an external position of power. Questions are posed on the computer screen. Each word of the question, as it appears, is underlined by a piano tone, so that the questions form coherent musical phrases. Once the question is completed, the lower half of the screen turns white and the spectator is invited to input his or her answer. There is only a piano-keyboard to type the answers with, and when a key is pressed, a word appears, soundlessly.

The way the words are mapped onto the piano-keyboard is never made clear to the spectator, and it is almost impossible to give a deliberate answer to the question, though there is a hidden structure to how the words are arranged, which is given in the relation between words and sound in the questions. When there is no activity for a few seconds, the next question is posed. A total of 25 questions repeat as an endless loop.

The questions are taken from European citizenship tests, predominantly tests from the Netherlands (circa 2012), which were circulated on newsgroups around that time. Under normal conditions, these questions would be posed as a multiple choice. The candidate would be given a choice of four answers. These would include a correct answer, perhaps two near correct answers, and one answer that presumably

would be totally unacceptable. But giving the correct answer to these questions is never clear-cut:

Your boss has a problem with you. He does not want to pay you, what are you going to do about it?

- A. Go to the police
- B. Talk to the governor.
- C. Go to the union.
- D. Take matters into your own hands.

Other questions range from moral conundrums in the workplace, social behaviour in the community, understanding of the welfare or the political system, to the acceptance of the supposed liberal values embodied by the society. These are some more sample questions, used in the piece:

You go to the doctor, but you are not making yourself understood. What can you do about it?

A civil servant at the immigration department has been of great help to you. How do you thank them?

You are stopped by a policeman, who requests to see your identity card. You do not have it with you. What do you do?

You hear from the immigration police that you will be removed from the country.

Who can help you?

You enter the room for a job interview. There are two people behind a desk. What do you do?

The job centre offers you some work that you find unappealing. Do you think it is within your right to decline the offer?

The overwhelming feeling of social conformity that this form of integration testing promotes, raises questions about the validity or even effectiveness of the procedure. This has been highlighted by the academic Ricky van Oers, who questions the validity of the testing of "core liberal values" in the way the questions are framed. Rather than testing the legality of certain situations, the 'what is right' and 'what is good' aspects of these questions test the candidates' knowledge of social norms, and mistakenly suggest that there is a moral consensus in society (Oers 2014: 126). What the subtext of the questions reveals, is the relation between the power structures of society and the aspiring citizen. The so-called 'inburgering' (the name given to the Dutch integration procedure) is somehow expressed as a process of translation of one's mode of behaviour, or moral system, from one culture to another. The language here serves as a thinly veiled mask of the ideological agenda of the power structures that frame it.

The encoding of questions in *Norms of Transposition (Citizenship)* occurs on the level of word parsing, the analysis of a text into its syntactic components. The sentences of the questions are parsed into their separate functions: noun, adjective, verb, pronoun, adverb, preposition, conjunction, interjection; each then assigned to a specific register on the keyboard. Other than an ascending melodic structure, giving the phrase an interrogative quality, there is no other expressive mode used in the voicing of the text. This neutrality, or unemotional musical expression, is an important feature of the construction of the voice of the questioner. It is a bureaucratic, impersonal voice, perhaps even a passive-aggressive one, hiding a contempt existing in the subtext of the question.

The encoding of the answers from piano key to text uses a similar though inverted structural approach. A vocabulary of 500 words, the most used 500 words of the English language, were divided on a 49-key keyboard, so that every note housed an average of about 10 words, selected by keyboard velocity (the weight that the note is struck). Instead of the parsing of text being divided in octaves as in the question part, the parts of speech (as much as it can be understood from a single word out of context) were placed on consecutive notes of the keyboard. This was repeated 6 times across the keyboard. The system of encoding was not entirely transparent to the participants who interacted with it, though repeated use and awareness of the causality of repeated words, could lead one to some kind of control of the text output. The piano-keyboard represents the idea of a hierarchical musical system—a metaphor of the authorial system it is encoding. It reorders and translates from one musical order to another scrambled linguistic one. Ultimately, the idea is to underline the feeling of frustration experienced in citizenship tests, or communicate a sense of the absurdity posed by these hypothetical questions, by forcing the spectator to answer these through a medium that cannot give sensible responses.

This system of music symbolically replaces everyday language. It asks the spectator to accept it as a system of communication, though the inability to use it like a language highlights it as inaccessible and out of reach. Here, the question and answer structure is used to highlight the divide between those in positions of power and those outside the system. This is reflected through the specific use of music as a surrogate for language. The musical code is built upon an existing language, but the key to the encryption is not made accessible to the participant. This renders the music unusable as a working surrogate. In spite of this the answers that the system produces are meaningful enough to engage with the questions in a poetic sense, and the carefree feeling one can develop with the installation, derived from the push-button response, is probably a fitting way of answering these absurd and problematic questions.



Figure 45 Installation view of *Norms of Transposition (Citizenship)*, as shown at the Municipal Art.

In the three works cited above, the translation of words into music and their subsequent problematisation, underline the ambiguity inherent in these specific question and answer structures. In all three cases, a request for information and the

ensuing answer in another medium reinforce the sense of a communication structure, even though the questions are never answered coherently. This places a focus on the act of translation between the media, and the possibility of meanings that are thus generated.

What can be inferred from this particular process of translation between media? The poet and classicist Anne Carson once said in an interview, that translation between languages always involves looking between the cracks of meaning generated between one language and the other.⁷¹ Translation between media involves an even greater metaphorical fracture, where meaning is always generated as one moves from one medium to the other. The direction of this movement is made very explicit in a question and answer form because the request for meaning is reinforced. Moreover, in the examples given in this essay of the two audiovisual installations, the metaphorical hierarchy between music/sound and language, between question and answer, are reversed. Analysing the relation in terms of conceptual metaphor, one could say that in both cases, language transforms from target to source domain and music vice versa, in the process of moving from question to answer. In the question phase, the sound simply reinforces the meaning of the text, while in the answer phase, the resulting incoherent text functions as a possible source of meaning to understand the musical/sonic act.

In a purely musical context, a question and answer phrase will underline the sense of voice, because it is mimicking a rhetorical device. It also creates the idea of completion because the answer phrase resolves any melodic or harmonic uncertainty that is set up in the question phrase. In contrast to this, the answers in *Dodona* and *Norms of Transposition*, like in Ives' *The Unanswered Question*, do not sufficiently resolve the question. They problematise the question—the act of questioning itself comes to the fore. In *Norms of Transposition*, a straight-forward answer is sought for a complex social conundrum expressed both in words and music. However, the answer, which has to be given through a musical system, is frustrated by the withholding of the encryption key, the translating dictionary, so to speak. The answer follows the logic of the music system rather than the language system. Similarly, in *Dodona*, the question that is addressed to one's self receives an answer through the chaos of wind noise and the algorithm, by which it is translated back into words. This change of frame creates a critical distance that generates many other layers of meaning, putting more focus on the voice of the questioner and the process of translation, rather than the question itself. In both cases, because of the lack of tangible meaning, what is left is the voice itself, taking shape in both musical and visual form, as well as in the subvocalised inner reading of the spectator.

⁷¹ Interview with Eleanor Wachtel in the programme *Writers and Company* from CBC Radio, 08.05.2016.

Chapter 7: Voiceprints

The three works under discussion here, *Wordless*, *Disco Debris/ Varosha*, and *Der Komponist*, all deal with the material of spoken voice, that has been analysed, dissected, disembodied and reconfigured. These are works where the semantic content of speech has been removed, either by editing, granulation or re-synthesis, to focus on both the musical potential of the material, or to underline something about the identity of the speaker.

'Voiceprints' is a term that has come to be associated with the biometrics industry, covering identification and security. In practice, these are spectrograms, or frequency graphs of a voice that are analysed by a computer in order to detect entrenched patterns of phonemes, that are unique for each individual voice. These can be used for identification as accurately as a fingerprint or an iris scan. The acoustic patterns of a voice betray both the physical characteristics of a person, their sex, their body build (chest and abdomen), the size and shape of the vocal cavities (mouth, throat and vocal cords), their age, as well as cultural indicators, emotional state, and behavioural patterns such as dialect, speed of delivery, pitch deviation and use of paralanguage.⁷² In practice, voiceprint authentication involves not only comparing a sample voice to a database of voices to confirm identity, but involves comparing the traces of the technology that they are captured on. Jonathan Sterne, discussing voiceprint technology used in the authentication of Bin Laden Tapes by the US military, writes about the complexity of dealing with indirect samples of the voice:

At every turn in the voiceprint authentication of a bin Laden tape, uncertainty wells up through the gaps. Voiceprint analysis compares (1) copies-of-copies to (2) copies, in order to determine if the voice on (1) is the same as on (2), which will then, in turn, let us know that, if (2) has already been "validated," (1) is also an authentic copy. Is your head spinning yet? (Sterne 2008: 91)

Voiceprint as a metaphor comes close to what Roland Barthes describes in his famous essay '*The Grain of Voice*' as the 'geno-song':

the space where significations germinate from within the language and in its very materiality. (Barthes 1977: 182)

⁷² The meta-communication around speech that comprises of sounds like gasps, sighs and other conscious or subconscious modifiers of speech. This is discussed in more depth in the section dealing with *Wordless*.

Barthes is here discussing, vocal performance, rather than speech, and is stating his preference for performance with 'grain', where the expression of a singer's individuality trumps the vocal codes set out by cultural conventions (which he defines as the pheno-song). Barthes choice of the word 'grain' (also 'grain' in French) is evocative in a digital music context because the use of 'grain' is understood in a very material sense within granular synthesis techniques. The 'grain' is here the miniscule piece of sonic data, a fragment of a soundwave or sample, often between 10 and 100 ms, used to construct a cloud or train of sound. Can the identity of a speaker be recognised from a grain of sound? There is certainly enough spectral information in a grain to deduct general qualities of a speaker, and with enough grains and time variant information, the voiceprint can be assembled. What I wish to highlight here, is that a forensic approach to the sonic dimensions of voice, separating speech from its primary semantic agency: words, syntax, language, does not drain it of meaning. On the contrary, another meaning comes into focus. This is not only the difference between what is known in semantics as 'speaker meaning' as opposed to 'sentence meaning'⁷³, the intention of the speaker, but also the identity of the speaker and the physical context in which the speech was made.

7.1 *Wordless*

Wordless is a suite of 12 'sound portraits' lasting 50 minutes, based on interviews with and by residents of Brussels, created as a commission from the Argos festival and the BNA-BBOT archives, and premiered at the Beursschouwburg, Brussels in October 2004.

The concept of *Wordless* revolves around the removal of semantic language from speech to leave only the so-called 'paralinguistic' aspects: hesitations, gasps, sighs, as well as emotional reactions and environmental sounds. Which kind of semiosis remains after this lexical dissection? What remains of the personality of the speaker when words are taken away? Usually it is very difficult to notice even the vocalisations that envelope language, because our attention is so accustomed to only focusing on the discursive aspect of speech. But there is much more occurring around speech than merely words. Linguist Fernando Poyatos defines this space around words as 'paralanguage':

the nonverbal voice qualities, voice modifiers and independent utterances produced or conditioned in the areas covered by the supra-glottal cavities (from the lips and the nares to the pharynx), the laryngeal cavity and the infra-glottal cavities (lungs and oesophagus), down to the abdominal muscles, as well as the intervening

⁷³ Philosopher Paul Grice suggests that the intention of the speaker is not the same as what is understood by the words (Grice 1991).

momentary silences, which we use consciously or unconsciously supporting, or contradicting the verbal, kinesic, chemical, dermal, and thermal or proxemic messages, either simultaneously to or alternating with them, in both interaction and non-interaction. (Poyatos 1993: 6)

The primary qualities of speech, discussed in paralinguistics, the loudness or pitch of the voice, have some relevance in the discussion of *Wordless* but are not the main focus of the work. We get some sense of the physiology of the speaker from the quality of voice we hear at the beginnings and endings of words, but because the main core of worded speech is removed, there is more focus on the 'differentiators' and the 'alternants'. Differentiators are parts of speech that give clear emotional information about the speaker, they include: laughter, crying, sighing, gasping, panting, coughing and other more involuntary sounds that interrupt speech. (Poyatos 2002: 59). Alternants, more relevant in the case of *Wordless*, are almost a complete subsystem of speech, that encompass expressive interjections around words, which consciously or subconsciously communicate some emotion, or qualify in some way what is being expressed in words. An example of this are the many expressions denoted by the sound "h'm", which can denote:

approval, disapproval, hesitation, unbelief, admiration, acknowledgement, interest, disinterest, curiosity, anger, contempt, surprise, pleasure, displeasure, concern, suspicion, pondering, superiority. (Poyatos 2002: 143)

schoolboy 0032 The city seen through the eyes of one of its youngest citizens. The interviewee is a 7 year old boy from St.Gilles. He is certainly not shy or lacking in things to say. He talks about his school, hobbies, friends and future dreams. When he grows up he wants to become a pianist in the opera house. He's also not a big fan of the city, he'd actually rather go and live in a tree-hut in the forest.	pensioner 0496 This pensioner from St.Gilles is of Italian origin. She arrived in Belgium in 1951 with her children. Her husband worked in the coal mines. In 1962 they moved to Brussels. Their financial situation was difficult and they encountered many problems with racism. They found whatever work they could in order to support their children who according to her are now settled and living well.
rai singer/ electrician 0522 A singer of Algerian rai music, who also works as an electrician. He explains to a pupil from a school in Passerelle, his taste for music, his studies and the thoughts of his parents on him wanting to sing. During the interview he demonstrates his talent by singing a famous rai song "Didi" by Khaled.	cleaning lady 0500 At the age of 10 she came from Morocco with her mother and sister to join her father in Belgium. She stayed 7 years. In 1982 she went back again on what she thought was a holiday, but her father announced that she was to remain permanently in Morocco. It was a nightmare, and furthermore, she was forced into an arranged marriage. At the death of her father, and going against the wishes of her brother, mother and future husband she decided to return to Belgium. She thought she was going back to her childhood paradise, but it turned out to be hell.

Figure 46 Four slides from *Wordless*, showing the information given for the persons portrayed.

The fabric of the sound world of *Wordless* is constructed from elements of the paralanguage of twelve interviews from the BNA-BBOT archives.⁷⁴ The particular interviews were chosen for differing reasons. For the fascinating vocal quality of the interviewee, for the background sounds, for the surrounding narrative or simply for the remarkable character of the person that it was portraying.

In both the performance and installation versions of the piece, texts describing the interviewees and their narratives are projected in the space, slowly fading in and out as a single text block for each movement, remaining visible for between about a quarter to half the duration of the music. Why project the words? Or at least a summary of the words, when the point was to remove them in the first place? Paradoxically, the use of projected text in *Wordless* came out of a desire to restore some context and background narrative to the disembodied voices, precisely because they had a large part of their semantic capacity removed. Having achieved the shift of focus away from the spoken word to the paralanguage and the 'grain of the voice' (Barthes 1977: 179), I felt there was no reason not to reinstall some of the missing personal information, particularly expressed as it was in another medium. This, I feel, does not greatly interfere with the sonic fore-grounding of the wordless aspects of voice, and has the advantage of adding another layer of meaning, through which the listener can contextualise the sounds. The texts are kept as objective as possible, specific to the context of the interview.⁷⁵ These remain anonymous, just as they are presented in the archives; and the movements are titled either by the occupation or according to an overriding attribute of the interviewee. Next to the title, is the classification number of the interview as it appears in the BNA-BBOT archives, in case anybody would wish to hear the full interview. One could describe the use of projected text here as 'paratextual', as it functions in a similar way to programme notes, providing a context or a description of the source material. Even though the piece takes as its central precept the idea of voice, the projected texts retain an aspect of objectivity, they are not 'voiced' in the music. Below are two further examples of some of the texts projected:

Blind 0425

Born with minimal eyesight, he slowly lost all his vision as he was growing up. In the interview he explains how blind people go about their lives in the world of the seeing and what kind of appliances and devices are used to help. He doesn't rule out anything as impossible, he even names photography as one of his obsessions. He explains how sighted people experience unsighted people and vice versa.

⁷⁴ BNA-BBOT (Brussel Nous Appartient -Brussel Behoort Ons Toe is an organisation that focuses on the oral history of Brussels. Their archives contain thousands of recordings of the diversity of people and cultures of the city.

⁷⁵ Some of the texts are direct quotations from the way the interviewer described the recording in the database of the archives.

Drummer 0404

A Gambian drummer is interviewed by two children about his life as a musician, the different instruments that he plays, and the cultural adjustments that he had to make coming to Europe from Africa. The interview is held in Dutch, which for all the participants is their third language. In the course of the interview the children burst intermittently into fits of giggles, which seems to be ignored by the adults. It turns out that some amorous noises coming from an upstairs room are attracting their attention. At the end of the interview the drummer gives an example of his djembe technique.

The sound aspect of the composition is conceived for a 4-channel system, where two channels are played back through headphones, using the silent-disco system distributed to members of the audience as they enter the space, and the other two channels through a PA. The material sent to the headphones and the speakers differ in kind. To the headphones were sent all the intimate sounds, the edits of the interviews, which left the paralinguistic and the environmental sounds. These were lightly processed, compressed and filtered mostly, binaurally enhanced, looped and sometimes layered. The idea was to have the voice as intimate as possible, close to the ear. On the external channels were electronic sounds: samples, resonances, pulses and noise. These sounds are based in some way on the voice or the sound world of the interviews, they are electronic imprints of these voices. The musical structures created around the remainders of voice are a way of restoring some sense of a musical narrative, a portrait in sound of an aspect of the personality of the interviewee.

The splitting of the sound between headphones and PA creates a sense of inner and outer sound space, the voices transmitted binaurally on the headphones while the electronic sounds create a sense of spatiality, physicality and awareness of the listening environment, that has an almost holographic sonic effect. Depending on the degree of openness of the headphones, certain frequencies will be heard naturally through the headphones, others with more difficulty. High-frequency sounds have the least chance of passing through, so that in some places, these have been faintly mixed into the headphones. This has the effect of making the sound space more complex, as high frequencies tend to be a more pronounced indicator of spatiality than low frequencies. Turning direction while hearing sound from the headphones and the fixed sounds in the space, alters one's sense of location in a psychoacoustic way, accentuating the sense of an inner voice, though it is one that does not speak any words. The words are, however, read from the text projection, resulting in a strange symbiosis between the wordless voice-in-the-head and the inner reading voice.

This is the main axis by which the different media of the piece operate. The space created between the voice of the interviewee, heard as an intimate wordless murmur close to the ear, and as sense of their personality, or the situation of the interview, which comes through the projected text, that is read and imagined by the audience. The music of the piece drifts between these two poles of imagined voice, in turn highlighting aspects of one or the other. The fact that we never get a complete speech image, the absence of spoken words or even images of the people, stimulates the imagination to fill the narrative in with what is supplied by the music and the soundscape. Analysing this from the point of view of the metaphor model: one can state that the target is always a sense of the missing person, whether it is the wordless voice or the anonymity of the text, and that this is always placed along side the information given by the music. Furthermore, there is a metaphoric relation between sound and projected text, which is in many ways is more dynamic than the one described above, because the degree of presence of the interviewee is always shifting between the media. Whichever media seems to be providing a clearer picture at any given moment, will assume the target role in the hierarchy. The fact that text always fades mid-way through each piece, changes the dynamic, so that what we are left with is wordless voice and personification in music, a dynamic between sound heard within and sound without.



Figure 47 Installation version of *Wordless* shown at M HKA, Museum van Hedendaagse Kunst Antwerp, 2012.

7.2 Varosha / Disco Debris

I will have spent my life trying to understand the function of remembering, which is not the opposite of forgetting, but rather its lining. We do not remember, we rewrite memory much as history is rewritten.⁷⁶ (Chris Marker)

Varosha/ Disco Debris are two related pieces that share the same material. *Disco Debris*, the first manifestation of the piece, is an interactive installation created for the inaugural exhibition of the art collective Suspended Spaces⁷⁷ held at Maison de La Culture, Amiens. This version was later shown at M HKA, Antwerp and Centro Centro, Madrid. It is based around the idea of an invisible architecture of sounds and voices that are triggered by the spectator using a video tracking device. The performance version, *Varosha*⁷⁸ was initially created for a presentation at the Centre Pompidou, Paris and has been performed in many places subsequently. It uses similar sound materials to *Disco Debris*, but in a fixed timeline rather than depending on the interaction with the audience. It is constructed around a clear narrative form, and uses a video to convey the context of the material.



Figure 48 Photograph of the abandoned hotels of Varosha, as seen from the water, by Marcel Dinahet.

⁷⁶ From the film *Sans Soleil* (1983) by Chris Marker.

⁷⁷ The collective 'Suspended Spaces', is an art group based in France, whose work focuses on places and spaces, whose development is obstructed by political or economic conflicts.

⁷⁸ An online version of *Varosha* can be found here: <https://vimeo.com/192369559>

Disco Debris initially came out of an invitation to create a work based on the abandoned tourist suburb of Varosha, Famagusta in the military zone of Eastern Cyprus. It had been left uninhabited since the summer of 1974, evacuated during the time of the Turkish invasion; a ghost town inaccessible to anybody except military personnel. In its heyday, Varosha, as this suburb was known, was one of the most popular holiday destinations in Cyprus, having undergone extensive hotel development to cater for the expanding tourist industry. Since then the area has been fenced off and kept as a possible bargaining-chip for peace negotiations, which finally never transpired. Today the buildings are beyond repair, uninhabited and unmaintained for what is now over forty years. The resort is crumbling, the once lavish hotel structures yielding to nature's persistent pressure.

I had previously created work based on the recent history of the Island of Cyprus, *The Buffer Zone* (2004),⁷⁹ because of which I had been asked to contribute to a group exhibition on this subject. I also have a very personal connection to Varosha, as it happens to be the location of my own earliest memory. My family, like many residents of Nicosia, used to spend the summer holidays there. At the time of the invasion, in July 1974, at almost 5 years of age, I remember hearing sirens and running into the basement of what was possibly the 'Hotel Loiziana'⁸⁰, where we eventually spent the day drawing pictures on the concrete floor, using chalk that was being dislocated from the limestone walls by the bombing outside. The next day we were evacuated south, thus possibly becoming some of the very last civilians to have ever been in Varosha.

In 2008 while researching the project, I visited Famagusta, now known as Gazimağusa (as it is in the Turkish part of the island), in order to view the stretch of coastline going south into the buffer zone. No photographs are permitted anywhere near the vicinity of the ruined resort, and some of the artists from the Suspended Spaces collective that I was with, were in fact taken into custody and forced to erase their camera's internal memory. Returning there and seeing the derelict façades of these hotels from a distance, was an unnerving experience. Knowing that one's earliest memory was formed on the very day that time there stood still, was like being confronted by the concept of memory itself. I had the idea of using the metaphor of a discotheque to convey the sense of the hedonism of Varosha in its heyday. This was related to the fact that before the invasion, my father used to run a live music venue in Nicosia⁸¹, where many singers and musicians from the Middle

⁷⁹ Recorded on *Unsounds* (11U), written about in (Kyriakides 2007: 129) *Voices in Limbo*. In *A Fearsome Heritage: Diverse Legacies of the Cold War*.

⁸⁰ The hotel's name is based on my mother's recollection. I could not find any evidence to back this up.

⁸¹ The venue went by the name of Montparnasse, and was also temporarily occupied during the invasion, and later used by UN forces.

East used to perform. The idea of using fragments of Greek, Turkish and Arabic pop songs from the early 1970's was reinforced by the fact, that while I was in Famagusta in 2008, I had an interesting encounter and conversation in a music store, whose Turkish Cypriot owner, made me six compilation CD's of Turkish golden oldies from the 1970's, containing most of the musical gems that I used in the piece.

The way the piece was initially conceived was that voices, disembodied and granulated like the state of the buildings left standing in Varosha, would be mapped into different paths and planes in a neutral space, so that the visitor to the installation, would traverse and collide with these fragments, as if uncovering an invisible architecture. The public would enter one at a time into a dark room, in which they would experience the sensation of metaphorically walking through sonic debris. One would stumble onto a landscape of frozen voices, barely recognisable shards of 1970's pop music, static bird song, broken pulses of disco music reduced to an almost Geiger-like clicking and ghostly resonances. These imaginary spaces were mapped onto a topography of intersecting voices and sounds, slowly transforming over time. Technically this was achieved by using a video tracking system that mapped the movements of the audience onto a granulated moment of sound. The position of the person in the space would determine which moment in time would be heard, as if they were a play-head of a tape machine, or a stylus of a record player traversing the remains of a debris of vinyl. Technically, the software was handled by a patch developed using STEIM's Junxion⁸², which handled the mapping of the video camera and the granulation algorithms created in Kyma. There would be 4 or 5 paths created at any given moment, which would be mapped onto different sounds and voices, so that if one would walk in a particular direction at a given speed, a normal playback of the file would occur. If one would walk in the other direction, the sound-file would sound reversed; if one would stop, the resulting sound would be a granular freeze. Crossing the path would result in a fragment of audio being momentarily audible. These paths and sounds would gradually morph through the process of time, changing their identity and function, so that one never had a complete sense of the full audio landscape.

As well as original recordings of songs from the 1970's, the voices used were made up of different recordings of people re-voicing these songs. I played audio of different songs, in Turkish and Greek, to non-native speakers and asked them to sing along as best as they can. I wanted the tentative, hesitant quality, which would emerge when tracing the contours of these songs, as if remembering an image from the distant past. As well as voices, there were various other layers of sound. Pulses made out of typical 1970's drum beat of rock and pop patterns, which would also granulate and change tempo as the visitor moved in the space. A siren-type layer

⁸² Junxion is an application that routes data from different 'sensors' including HID, Video, and Arduino.

which would come on intermittently, comprising of two triangle wave drones that would change pitch depending on the position of the visitor, and different audio field recordings made in the area, capturing the sounds of birds and insects, the principle inhabitants of the area these days.

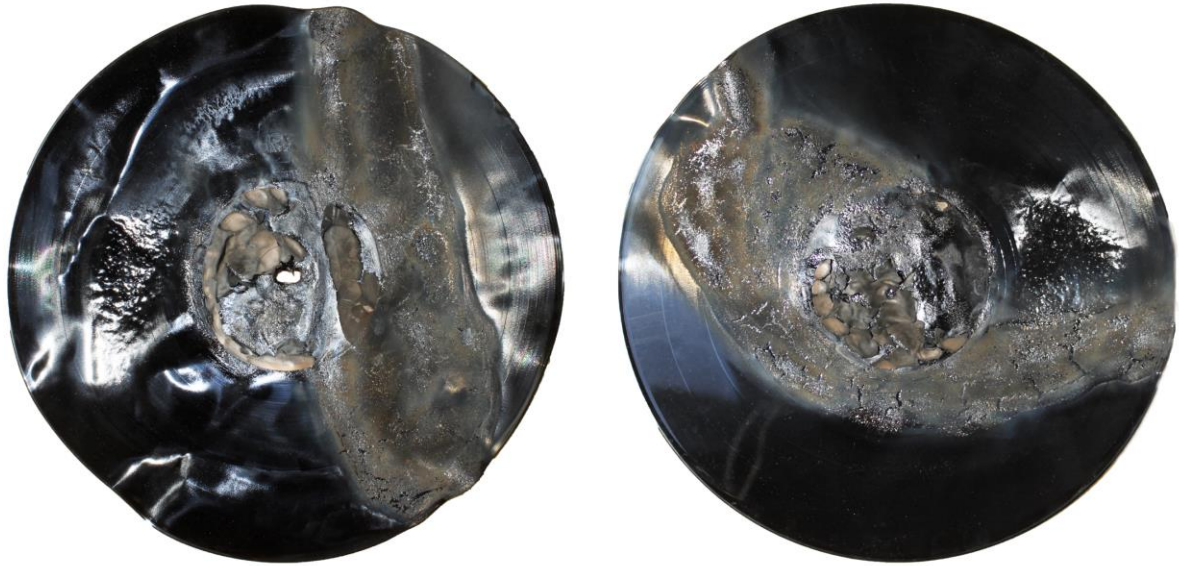


Figure 60 *Sides*. Blow-torched vinyl of Demis Roussos' *On the Greek Side of My Mind*.

Another source of sound, which eventually became a side-project in itself, was a Demis Roussos⁸³ record on which I blow-torched outlines of the map of the buffer zone and military bases around Famagusta. I used the aptly named record: *On the Greek Side of Mind*, released in 1974, the year of the invasion. The audio recordings of these warped records were also used as material for the piece.

The installation was created in a sealed-off space, made neutral by covering it with black cloth. Four theatre lights delineate the active space within. Inactive, the space just goes through a playlist of Turkish pop songs. As soon as a person enters the active space, the music would cease and the sound of the installation would be triggered, a violent rupture of the soundscape. Inside, there is a disorientating effect caused by the brightness of the lights. There is nothing else to look at, only shadows cast by the movements of the visitor, orientation is done by sound alone. There is an invitation to explore the space of sonic fragments, and an awareness of how one's movements effect the sound.

⁸³ The reason I used a Demis Roussos record has to do with the fact that he was a then friend of my father's, and in my mind the music represented by Roussos was exactly the music that I wanted to permeate the piece with.



Figure 61 Installation view of *Disco Debris* at Maison de la Culture, Amiens, (January 2011).

Seeing how audience reacted to this installation revealed several things to me about how interactive sound work functions in a visual art context. The amount of time people spent in the space was far shorter than I had imagined and intended. Sometimes people would walk in, wave their hands about, recognise the interaction and immediately walk out again. I was disappointed with the lack of patience in exploring the sound paths and piecing together the narrative of voices within the work. The sound manipulations were perhaps too abstract at times for a non-music audience to perceive what was going on structurally, and perhaps because of the complexity of layers, not transparent enough. If I would remake this installation, I would simplify this aspect of the structure, in order not to have any structural changes in time; keeping everything much more static, and letting it be animated by the visitor alone. I over-estimated the time a visitor to an art installation would spend exploring the piece, as opposed to the time duration that is usually spent in a concert situation.

I was excited about the possibilities that this installation opened up, in terms of stretching out composition time through a space, a letting a listener move through it, yet also became aware of the difficulties in keeping an audience motivated to engage in it. This touches on the problem of immersion vs. critical distance that I describe in Chapter 1, how it is difficult to maintain the balance between immersive activity, where an audience navigates its own path through the music, and have enough

critical distance to the work, in a passive role, to really listen. The unresolved issues around this led me to create a fixed-media version of the piece, that resulted in the 30 minute audio-visual performance *Varosha*.

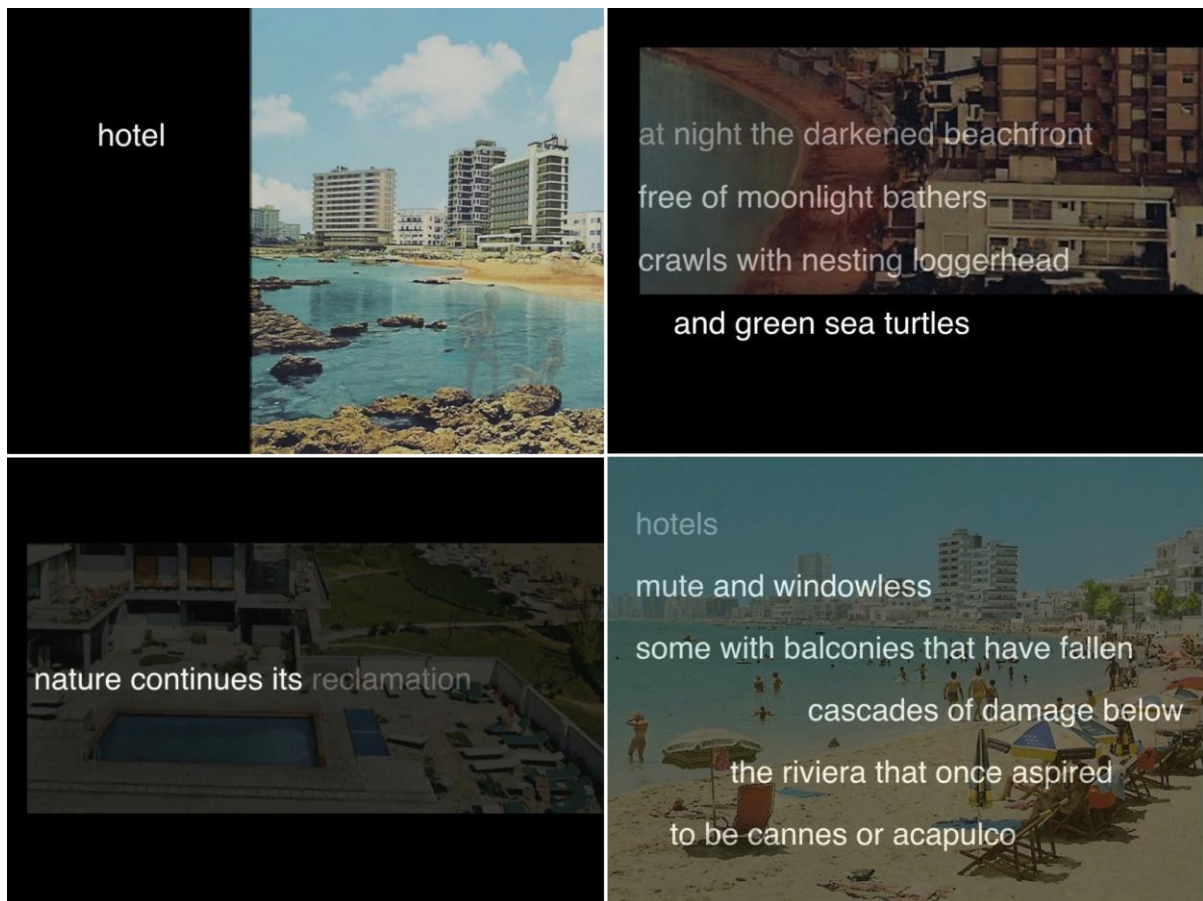


Figure 62 Stills from *Varosha*.

In this version of the work the same layers of materials are used, structured around a half-spoken, half-projected narrative that weaves together texts by journalists who have visited the ruins inside Varosha at different times throughout the last 40 years.⁸⁴ A quasi rondo-like structure is created, that periodically returns to warped versions of the Turkish pop tunes, while in different episodes, the piece explores the imaginary spaces and situations described in the journalist accounts of being inside a collapsing Varosha. This text is spoken by a narrator⁸⁵, and processed in various ways through the piece. There is a gradual shift between what is spoken and what is projected as text. At first almost everything is spoken, but gradually the intervention of the spoken voice becomes diminished, ending with only text. This shift from spoken voice to reading voice parallels the narrative's description of a gradual movement from a man-made environment to one where nature has completely taken

⁸⁴ These texts are related in Alan Weisman's book *The World Without Us*.

⁸⁵ On the recording Ayelet Harpaz narrates. When it is performed live, I do the narration myself, while processing my own voice.

over. There is a constant refrain in the piece, which is projected several times, 'Nature continues its reclamation project', leading to the final section, which describes the variegated botanical culture and turtle colonies currently thriving in the area.

This particular form of narration, that interweaves an external spoken voice and an internal reading voice, creates an interesting dialogue between inner and outer language, especially when the rate of word flow is unhurried. Friends who have experienced the piece, relate how there is a strange echoing of words, read or heard, that feeds into the anticipation of what is to come, resulting in sometimes not knowing whether the word was imagined or spoken by the narrator. The holes left by the missing words, both in the text and in the speech are sometimes answered by the sound or voices in the music, which shifts from illustrating or evoking the landscape, to recounting meta-narratives in song. The idea behind this was to create a disintegrating narrative space full of gaps, like the environment it is describing. The layers of voices in the work: disembodied, live-processed, sung, granulated and silent, are always in the process of shifting hierarchy within the field of our attention. The voice is ever present, articulating different perspectives between subjective experience and the objectivity of the journalists' account. Gradually by the final song, we hear the voice of the singer, disintegrating as it slowly morphs into a granular texture of sped-up field recordings.

The narrative here is also supported by images. In the first version of the piece I avoided any images, as I wanted the place to be imaginary, to be constructed solely by the sound and the words. At the time, the installation also did not need images, as it was shown within the framework of an exhibition about Varosha, with many artistic responses and historical perspectives on the subject. When presenting the fixed-media version of the piece, I found that I needed to preface the performance with contextual information, so I gradually began to show images, though not the shocking images of the abandonment of Varosha as it is now. The images that ended up being used in the video, are based on the 'doctoring' of existing postcards from Varosha in its heyday. Artist Isabelle Vigier, who created the design for the CD release of *Resorts and Ruins*⁸⁶, came up with the concept of taking the postcards and 'photoshopping' away all traces of people from them. These are used in the video projection as slow fades from the original postcards to the manipulated versions, as a way of reinforcing the sense of a gradual removal of human presence.

What are the essential differences between the installation and the performance version of the piece? These two entirely different forms of presentation are based on the same material. They aspire to the same idea, that of presenting an atrophied

⁸⁶ *Resorts and Ruins* is the Unsounds CD release of Varosha, where Isabelle Vigier's photo series *The Golden Seaside* appears as a set of postcards.

space where the voice is the primary metaphor for expressing a sense of ruin. Both forms refer to the absence of human presence, through traces of their past activity in this representation of a ghost town. They flirt with the possibility of immersion, where the immersion is incomplete: in the installation the movements of the visitor are answered by corresponding sound and voices, though always in the state of flowing away, like an apparition that cannot be grasped; in the fixed media piece, the immersion is cognitive, the reading and listening audience placing their voices into the gaps in the narrative, trying to piece together the ghostly images suggested.

Beyond the materiality of the abandoned architecture and the sense of the alluring ruin that it evokes, a triad of related concepts underlines both pieces: time, memory and history. The installation articulates in its form a relation between the self and time, the way a memory exists as a frozen moment, which can be revisited and rewritten in many ways. This applies to history as well, and indeed the recent political history of Cyprus is something that is still under negotiation; being constantly rewritten. Charlène Dinhut, who wrote an essay for the catalogue of the initial exhibition of *Suspended Spaces*, articulates the particular perception of time that is drawn out through *Disco Debris*, though it can apply equally to its sister piece, *Varosha*:

Disco Debris also manhandles the chronology of historical time, the piece updates a past that has been frozen for decades. In this present time, which also does not put a stop to being frozen in the past, ruins are forever being produced, too. The time of Varosha thus seems to be a 'hole' of History, a time without qualities where the forward direction like the backward direction, to borrow the image of the sound machine, do not make the situation either go forward or backward. This is a time which is not evolving and which might itself fall into ruin, while the illusion of the arrow of History breaks into smithereens. (Dinhut 2011: 248)



Figure 63 Photo looking into the crumbling hotels in the forbidden zone of *Varosha*.

As a way of understanding the correspondence between text/image and sound in the performative version of *Varosha*, below is a 'media correlation' analysis. What transpires is that there seems to be more convergence on the semantic level than on the sensory. The fact that context is such an important element in this piece, rather than an abstract synaesthetic relationship between sound and image is underlined in the chart.

Media Correlation: *Varosha* (Music & Text/Image)

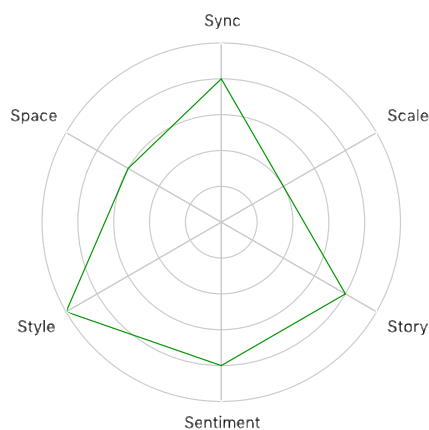


Figure 64 *Varosha*: media correlation.

Sync: 4

There is a tight alternating and synchronised relation between voiced text and written text that is also reflected in how the soundtrack responds to the text. Although there is more information in the sound, which is not represented on-screen, the structural changes are clearly demarcated visually.

Space: 3

This piece is about negative space, the re-imagining of a space no longer accessible. In this sense the music and visuals are constantly trying to conjure it and remind us of its absence. How this comes across as a sensory experience in performance has very much to do with the narration, both live and on the film, which reinforces the connection between the performative space and the narrated space. Music and text are likewise bridging these spaces, but in different ways.

Scale: 2

The scale of the re-imagined *Varosha* that is evoked by the soundtrack is immersive (surround sound) and plays on the idea of an invisible sonic architecture, which is only hinted at in the text and images. The visuals provide only a restricted window into the narrated space, whereas the scale of sound is more overwhelming.

Style: 5

The materials of the work all relate in some way to the space and time depicted - 1970's Greek and Turkish pop music, field recordings from the area, voices and samples related to the subject. Because this work originated in a visual art context, there was more focus on making the sound world more representational than abstract.

Story: 4

There is inherent ambiguity in the narrative of this ghost town, between the effects of geo-political conflicts, and the transience of human civilization and its reclamation by nature. Both soundtrack and film try to convey these ideas in their own way, in the subtle erasure of the found material.

Sentiment: 4

There is a clear feeling of nostalgia throughout the piece, but it is constantly undermined, by the edits and the objectivity of the narrative - sound and text both contribute to this.

7.3 *Der Komponist*

Der Komponist was commissioned to celebrate Helmut Lachenmann's 80th birthday in 2015. It was performed in Helmut Lachenmann's presence by the Philharmonie Zuidnederland, conducted by Bas Wiegers at the November Music Festival in 's-Hertogenbosch.⁸⁷

Whilst *Varosha* reanimates disembodied voices, using granulation as the predominant technique, *Der Komponist* takes a snippet of speech, a quotation from the composer Helmut Lachenmann, and time-stretches it using both granular and re-synthesis algorithms. This material of time-magnified voice, stretched from about 20 seconds to 20 minutes, is used to build the harmonic and melodic structure, which is performed by the orchestra. The piece is essentially a meditation on a single spoken phrase, and while the first performance of the piece did not permit the use of video for practical reasons, subsequent performances of the work were shown with a text-film that served the purpose of prolonging the focus on the meaning of this phrase.

In an interview given at the Ruhr Triennale in 2013, Helmut Lachenmann declares:⁸⁸

"Der Komponist, finde ich, hat nicht was zu sagen. Komponist hat etwas zu machen und das was er macht wird viel mehr sagen als was er sagen könnte... sollte auch ihm selbst was sagen...."

("The composer, in my opinion, has nothing to say, the composer must make something, and whatever he makes, will say more than he himself can... it should also say something to him")

The recording of Lachenmann uttering this sentence forms the sonic and conceptual spine of the work. The piece slowly scans through the contour of his voice, which becomes a landscape through which the orchestra navigates. The grain of voice is the space of the composition. The micro-fluctuations of his paralinguistic have a significant structural effect on what is sonified. The two main techniques used to process the voice: granulation and re-synthesis, where mixed and filtered on many levels to produce a complex texture, that gradually morphs throughout the course of the piece. This deconstruction of the voice fluctuates between granular windows of barely milliseconds up to several seconds, picking out transients that result in cycles of instrumental timbre looping, contracting, expanding and dissolving⁸⁹. Granulation often produces a noise-rich texture, especially when 'sibilants' and

⁸⁷ Live recording: <https://soundcloud.com/yannisky/der-komponist-for-orchestra-an-electronics>.

⁸⁸ Interview with Helmut Lachenmann about *Das Mädchen mit den Schwefelhölzern* <https://www.youtube.com/watch?v=yTIN807E2w> Quoted at [7:40].

⁸⁹ Three main granular algorithms were used for this: Paulstretch, Ableton Live and Kyma.

'plosives' make up the phonemes of speech. When time-stretched by a factor of 60, a second of this phoneme might create a noise texture that hangs for about a minute. Because of the structure of the German language, which uses many consonants at the ends of words, and because of the audible breath before and after words, phrases in this piece seem to come and go into a cloud of noise. Another interesting occurrence, which happens about half-way through the piece, is that we hear background pop-music, presumably from the cafe or foyer next to where this interview was filmed, coming through during a short pause in Lachenmann's speech. When this is time-stretched, the momentary major 6th chord gains an unexpected, though not unwanted, prominence in the work.

Figure 65 is a musical score for a section of a work, spanning measures 237 to 252. The score is divided into two systems. The first system covers measures 237 to 244, and the second system covers measures 245 to 252. The score includes parts for Comp., Perc. 1, Perc. 2, Perc. 3, Pno., Vln. 1, Vln. 2, Vla., Vc., and Cb. The score features various musical notations, including notes, rests, and dynamic markings. A specific section of the score is highlighted, showing a major 6th chord. The score is labeled with 'WHIRLY TUBE' and 'crossfade air'.

Figure 65 Section in the score showing the transcription of the 'major 6th chord' from the background of the interview.

On the other hand, the re-synthesis time-stretching technique used (based on an algorithm in Kyma), does not result in the smudging of unvoiced parts of speech. In this algorithm, the voiced parts of speech are mapped onto a single frequency. This results in a tonally pure, melodic line, with a strong vocal quality, as if it is being sung.⁹⁰ Mixed together with the granulated material, this provides a clear focus for the melodic contour of the voice. Throughout the piece the perspective of these materials is dynamically altered. Moreover, there are momentary punctuations with frozen 'transients', parts of speech that accentuate that a change of phoneme is occurring. These are used sporadically in the electronics, to highlight important moments of transition.

The image displays a musical score for a section of *Der Komponist*, spanning measures 327 to 333. The score is arranged in a multi-staff format. The top section includes staves for Horns 1 and 2 (Hn.1, Hn.2), Trumpets 1 and 2 (Tpt.1, Tpt.2), Trombones 1 and 2 (Tbn.1, Tbn.2), and Tuba (Tba). The middle section includes a Computer (Comp.) part with a note 'pulse fade - voice melody in' at measure 328, and three Percussion parts (Perc.1, Perc.2, Perc.3). The bottom section includes a Piano (Pno.), Violins 1 and 2 (Vln.1, Vln.2), Viola (Vla.), Violoncello (Vc.), and Contrabass (Cb.). The score features various musical notations, including notes, rests, and dynamic markings such as *pp* (pianissimo) and *f* (forte). Specific performance instructions are noted, such as 'mute-on' for the horns and trombones, and 'light fast bow' for the strings. A 'harmonic gliss (partial 2 to 28)' is indicated for the contrabass at measure 328. The score is marked with measure numbers 327, 328, 329, 330, 331, 332, and 333.

Figure 49 Example from *Der Komponist*, showing the polyphonic tracking of the voice.

⁹⁰ A similar algorithm was first used in my piece *Paramyth* (2010), where the voice of a blind storyteller was time-stretched to the extent that it seemed she was singing.

The digital manipulation of the audio of Lachenmann's voice forms the basis of the orchestral material. The granulated time-stretch is spectrally analysed and translated into sequences of chords⁹¹, which vary over time. These chords are orchestrated in various ways throughout the course of the composition, sometimes cross-fading into each other, and sometimes reinforcing a pulse, which is heard in the loop-scanning of the voice. This harmonic grid forms one of the ever-present layers of the orchestra which comes in and out of focus at various times in the piece. Within that, the re-synthesised, melodic aspect of the time-stretched voice is translated to the orchestra as single voices, soli and sectional writing. These melodies trace the contour of the voice, gliding between notes in one continuous curve. At times this stays close to the material it is tracking, at times it deviates into its own polyphonic trajectory:

The text-film in *Der Komponist* serves the function of sustaining the meditation on Lachenmann's words throughout the course of the piece. For this I used an algorithm created in the java based Processing platform, based on the 'Geomerative' library by Richard Marxer. This is based on a generative geometry of letters. It maps a true type font into vectors and allows the script to morph the geometry, so that the text can move slowly between legibility and abstraction. I thought this was an interesting correlation to what happens to the speech in a time-stretch, words move in and out of semantic cognition, whilst always maintaining the sense that it is based on an act of communication. Below are screen-shots of the first text in various stages of evolution. Within about a minute this evolves into unreadable script.

There is a clear correlation between text-film and the music in this piece. We read what is being said (albeit in English instead of German), and the manner it is being presented is akin to the slow disintegrations of voice being heard. Correlation between orchestra and electronics is also extremely close, more than in any other piece I have written. The electronic and the orchestral sounds at times merge into one entity. On a narrative level, what is being said has a somewhat ironic or contradictory aspect. The fact that Lachenmann is saying that the composer has nothing to say other than what is heard in the composition, is at the heart of the paradox, because it is being expressed all the way through the piece. But who or what is saying this? And what exactly does he mean by this? Is he suggesting that meaning can only come from the music itself, rather than from how the composer frames the work? Or that the artist is powerless to say anything about their work, because the intentions in making something are never the same, as what is communicated by the piece itself? What struck me as very poignant in this phrase, is the understanding of the creative process of the artist. At some point in the creative act one surrenders to the piece itself. Instead of forcing meaning onto it, the artist listens to what the piece is trying to communicate, what form it needs to take, where it is evolving towards; and this does indeed change the perception of the maker.

⁹¹ Using a combination of the software Spear and AC Toolbox

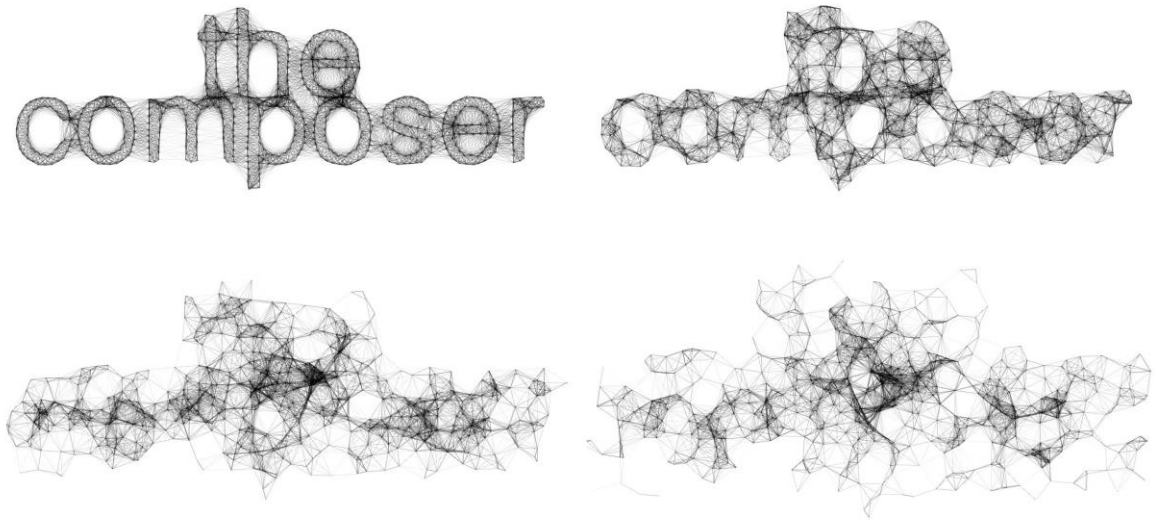


Figure 50 Stills from the first minute of the text-film *Der Komponist*.

The sound world of *Der Komponist* takes place within the microscopic grains of Lachenmann's voice. It follows the oscillations and micro-fluctuations of his vocal utterance, as if emptied of language signification, the physical mechanics of voice and language become the fabric of the compositional discourse. Each phoneme and transient used in the phrase uttered by Lachenmann, is analysed and spectrally transformed into the orchestral material. These spectral voice prints are used as building blocks to construct a musical commentary, that has to do with both identity and communication; perhaps alluding to the question of who or what is "speaking" in a musical work? The composer has nothing further to say.

Chapter 8: Interactive Scores

To call computer media "interactive" is meaningless – it simply means stating the most basic fact about computers. (Manovich 2001: 55)

The above quote may be applied to many other media. How interactive is traditional notation? Can one quantify the interaction of musicians or even an audience, if only on a sensory or cognitive level? What are the differences between a musician's and an audience's engagement with the music? Is it simply a case of one being the transmitter and the other the receiver? I think in practice the relation, or bond, created between the various roles in a music environment is more complex and more interactive. Christopher Small's coining of the word 'musicking' suggests that there is a shared ownership in the act of making music:

To music is to take part, in any capacity, in a musical performance, whether by performing, by listening, by rehearsing or practicing, by providing material for performance (what is called composing), or by dancing. (Small 1998: 9)

Much of what I have been suggesting in the earlier chapters of the thesis, of the very active role that the listener must take, especially when dealing with music-text-film, places the audience's role on par level with the musicians. The way a listener filters and fluctuates between different listening perspectives, drawing out or focusing on aspects of the various information flows, is a form of perceptual or cognitive interaction, that makes the notion of a static relationship insufficient. In practice, how does the idea of fixed media operate in a situation which suggests interactivity? Does a more fluid approach to notation, concert protocol, instrument technology or even architecture, stimulate a more profound merging of musician and audience experience?

In this chapter, I would like to focus on an aspect of my work born out of the early music-text-film, where the video component is intended for the audience only, to pieces where the text-film is used as score (with interactivity) for the musicians as well as the audience. There are two aspects of structural interactivity that exists in contemporary sound art practice, that lead on from the ideas of hierarchies of media and the point of view of the audience that I have discussed earlier. One deals with the interactivity of audience (described in the previous chapter on *Disco Debris/Varosha*) and another deals with the interactivity of musicians, specifically in the latter case, interaction with the score. There are two basic questions pertinent in the pieces discussed here, *Karaoke Etudes*, *Trench Code*, and *Oneiricon*. What happens when the audience and musicians share the same information? And what is the function of a score reaching beyond print technology?

One of the unsatisfactory aspects of staging performances of music-text-film in the last years, has been the realisation that audience and musicians have a very different experience of the piece. Most of the time, the musicians are unaware of the text-film, since they have their backs to it and are looking at the sheet music on their stand. The conductor may be only half-aware of it, as he or she is very much fixated on the score. This reinforces an unequal relationship already existing in a traditional concert environment, delineating and separating the role of musician and audience even more than is already the case. In the classical music tradition, this division is clearly demarcated by the concert ritual, dress code and architecture dividing the space between active and passive roles. The division is also manifest through the score. There is a production line going from composer to audience, and notwithstanding the cases where the music has been internalised by the musician, the 'score' manifests itself on a music stand between the player and the audience. Is the score the medium? Or is the role of the performer a 'medium' between composer and audience.

The production line also extends to the process of transforming the composer's ideas and encoding them into musical notation. An interactive process which is set into a fixed medium. Like so, there is interaction in the performers' handling of the musical instruments, to convert what is essentially 'text' into audible sound for the audience's ears to detect.

And what of music-text-film? Here the musicians are looking at fixed dots and lines on a page, while the audience is looking at words flashing on a screen. There is even no eye-contact between the two groups. One could point to a beautiful asymmetry in that the performers create sound in response to one set of graphic symbols, and the audience listens to sound through another set. But what if they shared the same set of symbols? With these thoughts in mind, I wanted to explore a situation where the musicians and the audience were on a slightly more equal level with the 'video-score/text', although their roles in interpreting it are very different. The musicians will make music and the audience will absorb what they see on screen and add this to what they hear the musicians produce, and what they see happening on stage (including witnessing the musicians following the video score) to form the total of their experience.

As increasingly more musicians I know started to use tablets to screen their score, I wondered why the tablet was not being used as more than just a PDF display. I thought of exploring a type of score creation that would only be possible with computers and tablets. Music scores are generally considered to be fixed media. The technology behind their proliferation, the Gutenberg printing-press, grew out of a need to have a uniform way of conveying and encoding information, less uncertain and prone to error than music conveyed by a scribe or aural memory (McLuhan

1962: 61). Composition developed together with the commercial benefits of this technology, but it also made use of the complexity that can be achieved by encoding in paper a multiplicity of synchronous sounds and musical gesture, to be performed by large groups in different places with no need for the composer's presence. Thus the medium of musical notation, just like the medium it is encoded in, developed hand in hand. The question could indeed be asked, whether at the twilight of print technology, are we seeing the end of traditional 'paper' composition?

An interesting example showing the limits of one technology, pushing the boundary into another, is John Cage's Variations series. In *Variations I* (1958), dedicated to David Tudor, the musician(s) has to construct the score from a series of transparent sheets of squares, lines and dots by 'dropping perpendiculars', measuring them and translating the data into musical parameters. Obviously, the construction of the score in *Variations I*, has to be performed before the performance (although one could incorporate this as part of the performance itself). This piece, like the rest of the Variations series, does not qualify as one possessing a graphic score, in that there are specific guidelines about how the elements are to be measured and interpreted. What is fascinating about the series, specifically *Variations I-IV*, and also a related piece, *Cartridge Music*, is that Cage not only utilises abstract graphic elements printed on transparent sheets, but that the musical "output" is largely dependant on the user physically manipulating the materiality of the score. Cage has managed here to make a fixed medium, namely print, behave in a fluid way. Examining the piece from the vantage point of the 21st century, one could observe that Cage was perhaps trying to dematerialise the medium of paper years before digital media managed it (the transparencies could be taken as a sign of the dissolution of fixity).

In recent compositional practice, a number of composers have been experimenting with animated notation, generative and interactive scores,⁹² this includes members of Icelandic composer's collective S.L.A.T.U.R, Jesper Pedersen, Guðmundur Steinn Gunnarsson, Ingi Garðar Erlendsson, and also Þráinn Hjálmarsson, Ryan Ross-Smith, Nick Collins, Lindsay Vickery, Cat Hope, to name a few. Approaches vary from the openly visual to the more recognisably traditional musical notation, albeit generated by the computer in real time. In my view, digital renditions and generative strategies to music notation will proliferate in the future, and with it the idea of fixed composition.

The three works presented in this chapter, progress from a fixed video score, to a score which can be generated live and controlled centrally, to finally a score that is generated live by each musicians for him or herself, or even by the audience on their smartphones. One of the great musical democratisers of our time is 'karaoke', a form

from which my first attempt at experimenting with alternative musical notation together with text, and visualising it for the audience, *Karaoke Etudes*, took inspiration. It toys with the conventions of the medium rather than duplicates its social interaction. Nevertheless, the musical code is revealed to the listener, which in my mind creates the possibility of different ways of listening and relating to the musicians. *Karaoke Etudes* eventually led to *Oneiricon*, a piece specifically written with the idea of using the tablet as an interactive score, instrument and screen for the audience. This process led to changes in the way I perceived the musical performance space. Sharing the screen with the audience, that is, making it accessible for anyone to read, meant the line dividing the audience from musicians might also have to be redrawn.

8.1 *Karaoke Etudes*

Karaoke Etudes was commissioned and written for Champ d' Action (Antwerp) on the occasion of their LAbO workshop in March 2011. The solo parts were performed by the ensemble itself, and the tutti part by the 30 or so participants of the workshop. The make-up and number of musicians in *Karaoke Etudes* is open, and can vary from piece to piece. Even though originally the pieces were written with a fixed idea of particular solo instruments, these can vary and the work has indeed been performed with many variable instrumentations. Recently a version was toured by two Canadian ensembles, Thin Edge New Music Collective and Ensemble Paramirabo , and they chose to perform each solo as a musical duel, a battle.⁹³

Twenty-first-century life is karaoke—a never-ending attempt to maintain dignity while a jumble of data uncontrollably blips across a screen. (Copland 2010: 9)

Karaoke is Japanese for 'empty orchestra'. More than simply a performance practice, Karaoke is a cultural phenomenon that in the words of Dubravka Ugresic supports:

less the Democratic idea that everyone can have a shot if they want one and more the democratic practice that everyone wants a shot if there's one on offer. (Ugresic 2011: loc 102).

In conventional Karaoke form, the music is a well-known pop song, that the participant chooses from an enormous catalogue of pop hits from all eras and tastes. The song is played minus the lead vocals, in what is often a basic and mechanically sequenced arrangement. At the same time a low budget video is shown, which could comprise, for example, of a boy-girl narrative in a picturesque setting, along with the

⁹³ A recording of this performance can be heard: <https://www.youtube.com/watch?v=HEPaAtrnZag>

lyrics to the songs displayed with a moving symbol, a coloured bar or ball, letting the participants know which lyric has to be sung at a given moment. As an idiosyncratic form of participatory music making, there is a curious form of role-play involving the singer and the overseeing public. Firstly, there is the aspect of a possession ritual at play, where the participant takes on the guise of a pop personality; the songs are not merely well known pieces of music, but are unmistakably tied to a performance artist or band, so that when a song is performed, there is both an aspect of identification and anonymity at play. The fact that audience and singer share the same crib sheet, and are complicit to the mechanics of the performance, gives both a parallel point of view of the performance. There is the memory of the original song, the imagined performance taking place in the audience's mind whilst reading the lyrics, and then the actual performance and interpretation by the perhaps drunken participant. According to Ugresic the discrepancy between celebrity culture and anonymity that is at play in Karaoke, is at the core of its attraction.

Karaoke is entertainment of ordinary people, who, within given codes (shaped by technology and genre), and protected by the mask of anonymity, fulfil their suppressed desires within their own communities, or fandoms. (Ugresic 2011: 132)

It is the epitome of model for a form of music-minus-one that involves the participant in combination with technology, resulting in a transformative act that blurs the border between performer and audience.

In the collection of 5 pieces named *Karaoke Etudes*, I have attempted to explore the themes and concepts of Karaoke. The conventions of Karaoke are taken as a starting point for the creation of a video, which functions both as visual information for the audience and as a score for the performers. The ontological levels of Karaoke: song - video - instrumental - performer - audience, are expanded in these 'Etudes' to song - video - instrumental - performer - ensemble - audience. In this sense, and on a certain level, the audience becomes an active participant in the Karaoke performance. Another reason for using this extra layer is rooted in the fact that in Karaoke the vocal melody is always a given. It does not need to be represented graphically, it is simply incorporated in the instrumental layer. In *Karaoke Etudes* the solo part becomes a written out improvisation based on the melodic material which is projected in a simplified form on the video. The soloist can choose to perform more or less exactly what is written, improvise around it, or make up something entirely new. The fact that the audience (or at least those who are musically literate) can be a witness to any deviations by the performer, gives the interpretation an added layer of representation. The main substance of the video is an animated graphic score based around the lyrics of a given song, which acts in the piece as a surrogate of a particular audience. The 'friends' of the soloists sending messages of encouragement or banter during the performance (as opposed to the audience, who

in this metaphor could be the rest of the public in the karaoke bar). The form of the graphics and how they should be interpreted varies from piece to piece; the code is not difficult to understand, it has an immediacy that a non-musical audience can understand, but there are certain rules, which guide the musicians in interpretation.

The lyrical content in the pieces is used playfully. To a certain extent I expected audiences to guess the given song (there is always a sample used in the electronic layer), but later I titled each movement, making the original source more evident, so that the audience could have a relation with the lyrics, triggering a memory of the original song (songs often carry many personal memories). There is also a sense of misdirection and re-contextualisation in the use of the lyrics of the songs, which often functions as a part of the material of the score.

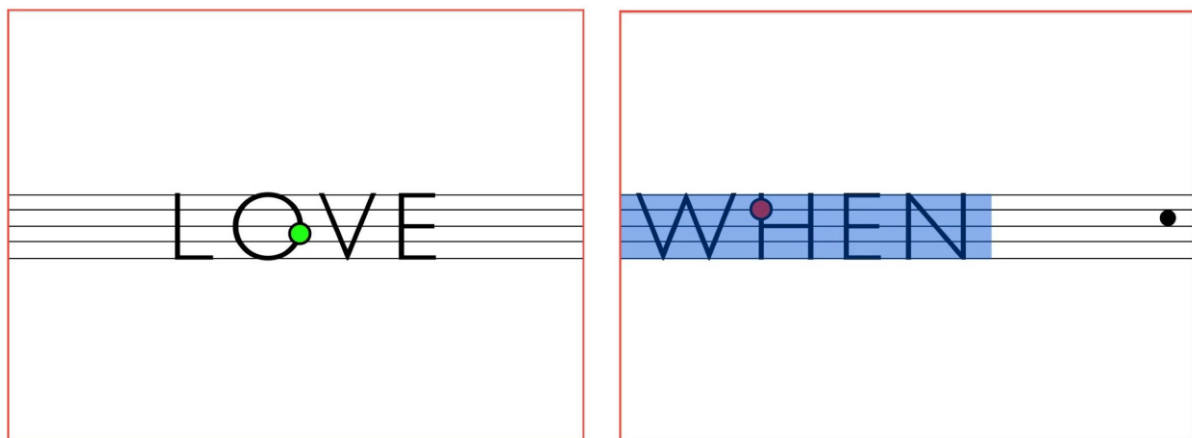


Figure 51 Stills from *Karaoke Etudes 1*.

In the first piece, based on Marvin Gaye's *Heard it on the Grapevine*, one sees an empty staff with a cropped text (a lyric from the song) that changes every four beats to a new image. Attached to the text is either a blue, green, red or yellow dot. This refers to the beat of the bar in which a note should be played by the ensemble. If a note begins on the first beat of the bar, a guide note is given in black during the preceding bar, to leave time for preparation. In this movement there is a certain precision as to what note should be played when, whereas instrumentation and manner of interpretation are left open.

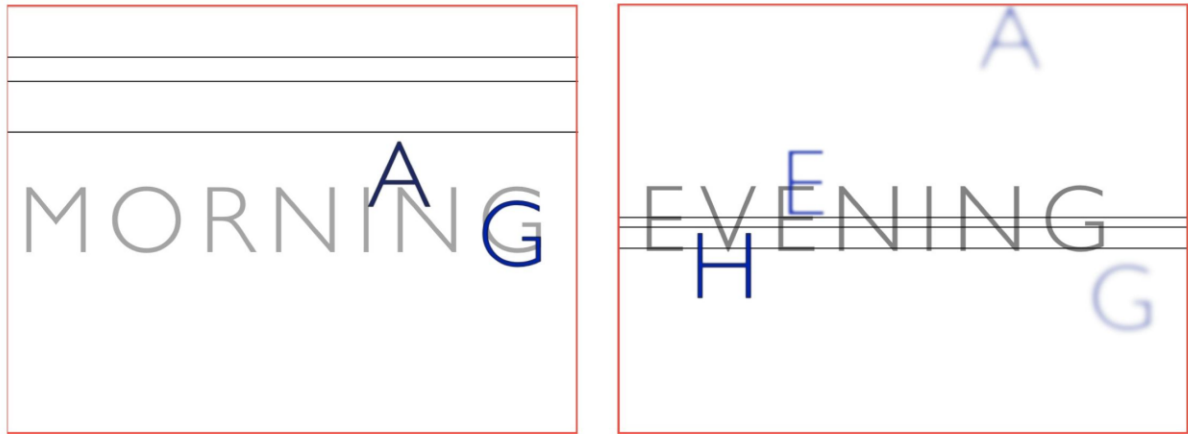


Figure 52 Stills from *Karaoke Etudes 2*.

In the second movement, based on Bob Marley's *Sun is Shining*, the graphics are comprised of freely floating lines and letters, that come in and out of formation to form meaning and trigger note events. These note events are described by the letters A to H floating in and out of the lines. The logic behind the disintegration of the stave, has to do with the fact that the letters of the words take on the function of musical code, leaving the lines of the stave to suggest a more expressive idea. This suggests a rather unsettled, fluctuating field of pitches and textures to accompany the solo flute part. As opposed to the on/off character of the first movement, this movement conveys a sense of gradual evolution and fluid intonation. At the climax of the song, a note appears for the first time, momentarily inflates and fades, like a giant sun, a reminder of the redundancy (in this movement) of the dot on the stave.

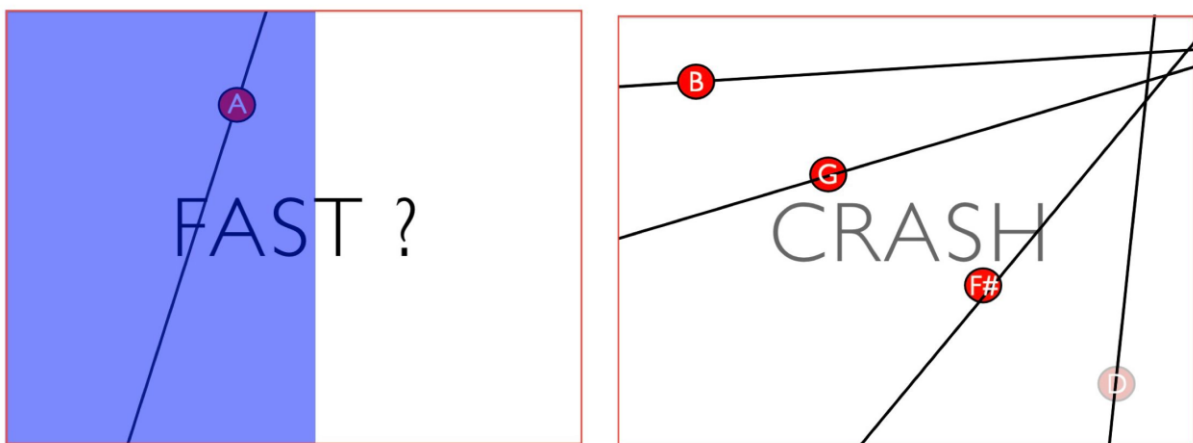


Figure 53 Stills from *Karaoke Etudes 3*.

In the third movement, based on Gil Scott Heron's *Fast Lane*, the notes return, but without the staves defining their pitch. Instead, a letter inside the circle suggests the pitch, while the lines behave in a much more erratic way, mostly moving in circular paths around the notes, conveying a sense of timbre change. The soundtrack in this

movement is quite hectic, and suggest a strong pulse. In performances of this movement I have always encouraged the musicians to interpret this video with a strong feeling of pulse (though I'm happy if they interpret it otherwise).

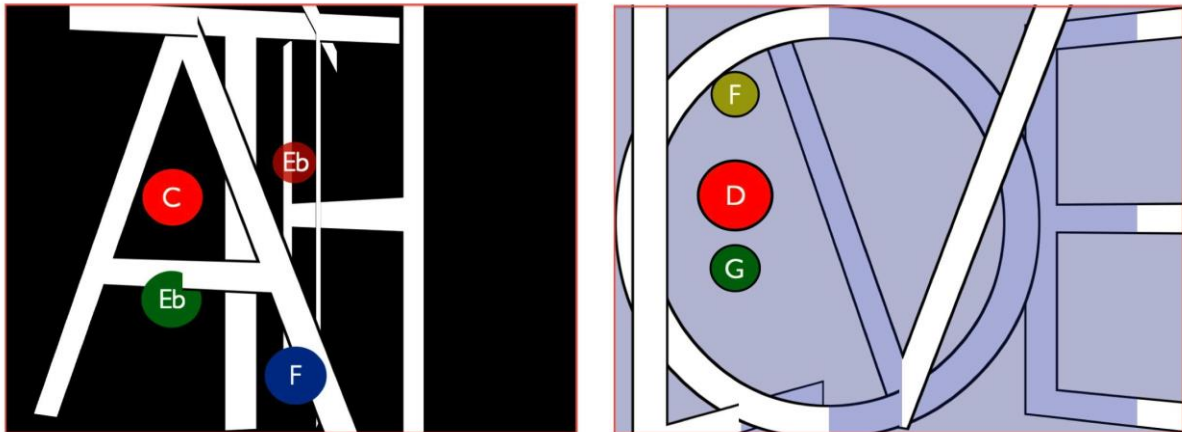


Figure 54 Stills from *Karaoke Etudes 4*.

In the fourth movement, based on Leonard Cohen's *Everybody Knows*, the lyrics of the song take the form of 3D objects, where coloured notes expanded and contract from them. In this movement the staves have disappeared entirely. The words have grown to architectural proportions, and act as a distortion element as they bury and reveal the notes in various stages. The blue-line indicator, which has adopted various functions in the previous videos, is now used to wipe the slate periodically. The soundtrack is built around a time-stretch of Cohen's voice, comprising of noise and pitched elements, reinforced in the cello/contrabass solo. The pitches shown on-screen do correspond to coherent harmonies, but convey a slightly amorphous quality in the way they appear and fluctuate in size, which is meant to have an effect on how the musician articulates their dynamics.

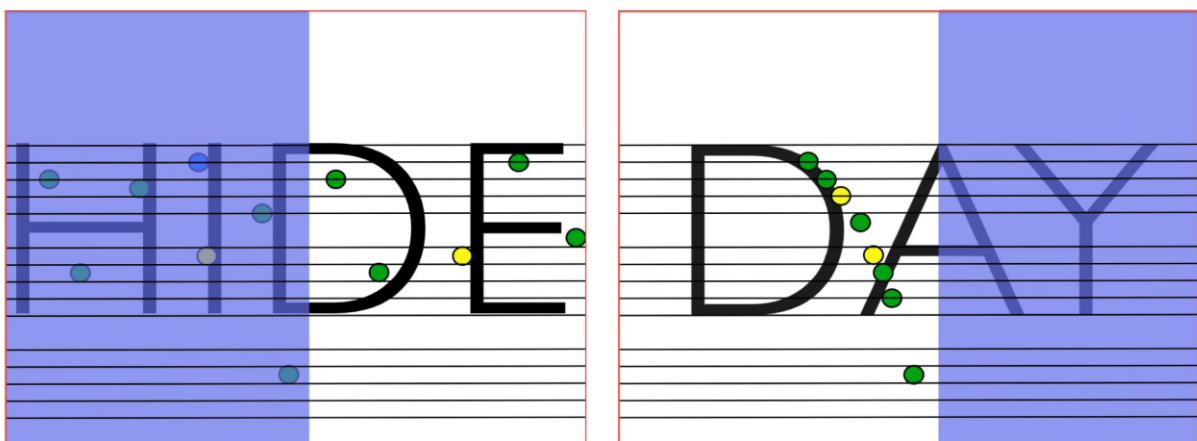


Figure 72 Stills from *Karaoke Etudes 5*.

In the last part based on Nina Simone's *Sinnerman* there is a return to a more conventional notation. The staves are back, the notes have a more fixed function and the blue 'now' indicator plays an important role in synchronising the attacks. The notes either move towards the blue border, or the blue border moves towards the notes. The colour of the notes indicate whether they are short or long. The return to a more conventional function of the symbols was not a conscious narrative decision, but, interestingly, because the relation between video and sound is very clear in this last part, the audience immediately understands the rules. This understanding also arises out of primitive video-game aesthetics that the last movement evokes. But there is a danger in this: the close synchronisation of animated notation with the resulting sound, and the expectation it creates in the audience, can have the effect of narrowing the window of time experience in the music to the ever-present 'now'. Without making a judgement on whether it is a good or bad thing, this is one of the side-effects of using animated notation, especially one where the tempo is controlled externally. This has two repercussions, that the musicians do not control the tempo, and that the focus of the audience cannot so easily drift back and forward in time.

8.2 Trench Code

Trench Code consists of 3 interactive videos, which act as scores and instruments for an open ensemble. Each video is based on a particular code book used in the trenches during the First World War, presented in a graphic form and encoded into musical material. A communication system is set up amongst the ensemble, utilising the vocabulary and codes found in the three trench code books. The scores are partly computer generated, though always relying on the encoding of the text as appears in the books, and on the operation of the score-player⁹⁴.

The three code books⁹⁵ used in the piece are:

- The *Mohawk Code* used by the American Military (1918)
- '*Schlüsselheft*' used by the German Military (1917)
- The '*BAB*' Code Book used by the British Army (1917)

These code books were used to encrypt and decrypt messages into numbers, that were encoded and intended to be sent through the radio. Radio communication was a relatively new phenomenon at the time. Both sides of the conflict were beginning to use it, while also eavesdropping on each other's communications. Radio

⁹⁴ A recording by Ensemble MAZE can be found: <https://vimeo.com/226869061>

⁹⁵ I am very grateful to John McVey and Fred Brandes for sending me scans of these rare code books.

would have made the need to lay wires through dangerous or inaccessible terrain in the trenches redundant, though in practice telephony and telegraph would have still been used to communicate from the front lines to the units further back, these code books providing the necessary encryption (Kahn 1996). Encryption methods used during World War I were quite primitive by today's standards. Initially, a basic substitution cipher would have been used, which later in the war evolved in to using code books. Although ciphers were in general easy to crack (an example is the German ADFGVX cipher, whose decipherment by French codebreakers led to an important gain for the Allies), code books ran the danger of being captured by the enemy and made redundant. This meant that books would be updated on a regular basis. The Germans changed their code books twice a month, and used different books along different sections of the front (Goebel 2014: iv).

My interest in code books and the transmission of secret messages arose initially out of a piece which I briefly discussed earlier: *a conSPIracy cantata*. This led to a general interest in codes as a communications system and the relation to music. Another work which is close in culture to *Trench Code*, described briefly in the appendix, is *The Queen is the Supreme Power in the Realm*, an environment for improvising musicians, written for the Köln based ensemble musikFabrik. In this piece, systems of communication are set up between different parts of the ensemble, using the language of telegraphic code books. In *Trench Code*, a system of communication is also set up, but this time mediated by the score software and the score-performers. Each of the three distinct trench code books are rendered into a computer application and projected as a score. Three distinct groups of musicians of any size playing from one score, communicate between each other within certain rules. There is an interaction between the player who is controlling the score and the rest of the musicians in their group, but not necessarily between each group. They are not necessarily in a metaphorical 'state of war', but they are aware of each other's reading. Each group has a particular musical characteristic that is highlighted, and the idea is that the three layers can co-exist well in the same sonic space. In order to underline the non-combative metaphor, and to subvert the destructive conditions that these books emerged from - it is encouraged that the musicians move around freely in the space between the different scores, thus suggesting the idea of warring factions is not implied in the way the musicians are organised.

Initially the work was commissioned by the Brugge Concertgebouw as part of First World War reflections/commemorations that were taking place in 2015. It was performed there as three separate movements with the Veenfabriek Sirene Orkest and musicians from the group Okapi. In this version the work functioned as a more traditional video score, where the musicians would interpret the graphics and rhythm of the score in the particular ways possible with the instrumentation of

sirens, intonarumoris⁹⁶, synthesizers, paper instruments, turntables and radios. The score operator, at the time myself, would respond to the musical material being generated in terms of tempo and rate of change of material, and there would be a natural feedback between musicians and score. The score operator functions almost like a conductor, though more in the sense of an improvising 'conduction'⁹⁷ than in a classical orchestral setting. In the revision of the score in 2016, I thought it might be more interesting to let the scores be played simultaneously and provide instructions to how the work can be thus set up.

In the new version of *Trench Code*, there are three laptop score players, positioned behind or near three screens set up around the audience. The musicians move between the zones of the three screens, or stay fixed in one position. There is no fixed structure as to how the piece should begin or end, so that different situations can arise in unexpected moments and in different performances. This is to highlight the idea that the musicians are creating the situation by their actions, which are not being too tightly controlled by the video score. This sense of de-centerilization of control is specifically what is being problematised in this version of the piece. Because the material of the *Trench Code* books inherently deals with conveying and obeying very specific and clear orders, I wanted to set the piece up in such a way that what is being communicated is never clear cut, is ambiguous and always relative to what other sound or action is happening in the space.

The score consists of three 'apps' created on the 'Processing'⁹⁸ platform, a java based language. These programmes need to be performed by a player, as they do not just run automatically. In all three scores, the player controls the rate of change of information, and has an influence on the choice of visual presets at any given moment. The constant interaction and engagement with the players is an important feature of the piece, because it is by no means a fixed video score relationship. In all three scores there is also sound that accompanies the visuals being generated, which reinforces the active aspect of the score, and allows the musicians the opportunity to respond from aural rather than visual stimuli. Below is a description of the construction and function of each of the three scores:

⁹⁶ Intonarumori was a type of instrument invented by Italian Futurist Luigi Russolo.

⁹⁷ Conduction is a term coined by Butch Morris to describe a lexicon of signs used by a conductor to shape an improvisation.

⁹⁸ <http://www.processing.org/>

Mohawk

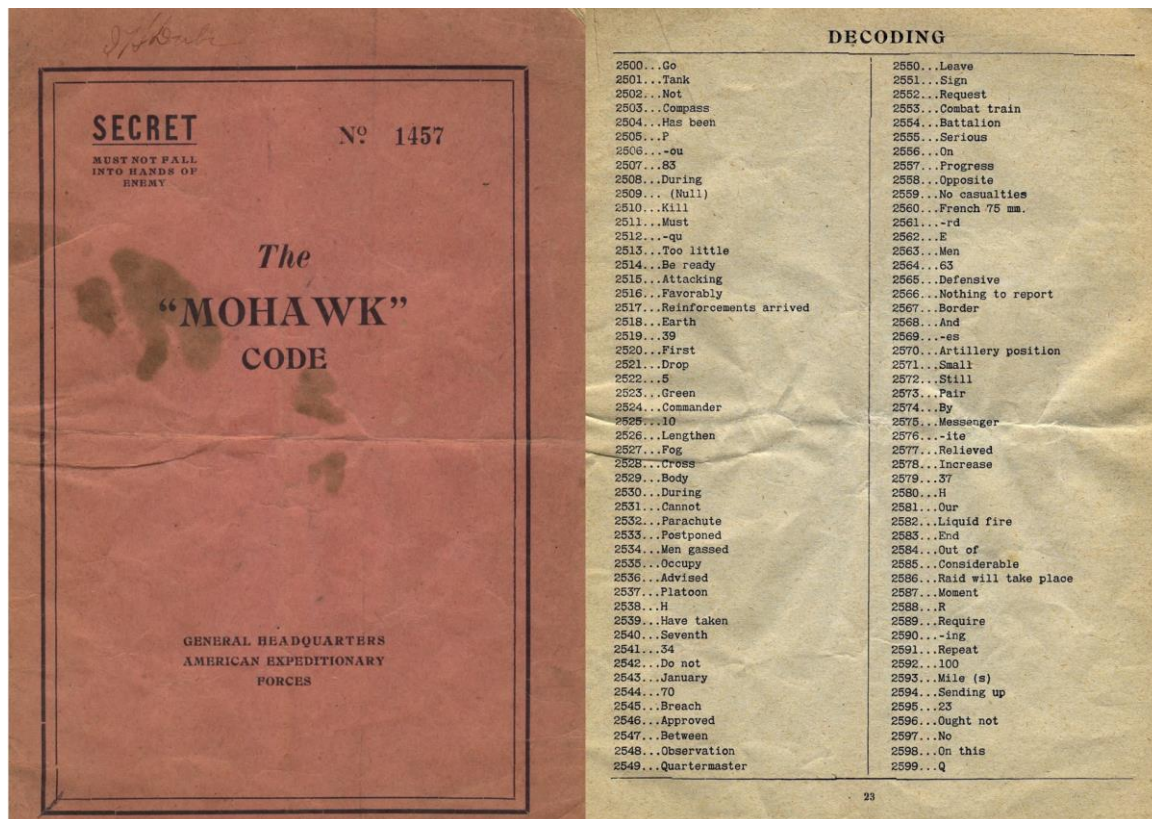


Figure 73 Cover and sample page from *The "Mohawk" Code* (1918).

Mohawk is based on a trench code book from 1918 issued by the American Expeditionary Forces. It is split into two main parts, one part for Encoding messages where the words are listed alphabetically, and one for decoding, where the list is numerical. The eventual message would be sent as a series of numbers, with very clear instructions given as to what would be the safe protocol; avoiding repetition, using random 'null' messages, sometimes spelling words and sometimes using phrases.

In this score, the basic idea is that words are mapped onto lines and dots on the screen. The complete 'decoding' section of the book is used as a data set in the score. The words appear preceded by a number and follow numerically at a speed controlled by the player. If the mouse is clicked, then the words will quickly and randomly scroll until it is released and a new sequence of words starts to appear. The mapping of the words to graphics has a few simple rules: if the first letter of the word is a vowel a line will appear, if it is a consonant an ellipse of a certain size will appear at a certain height on that line. If the coded word starts with a number, a shaded black bar will appear on the screen, if a dash precedes a word then random small dots will be splashed across the screen. These graphics build up over time, until the mouse is pressed and the screen is refreshed. Each new build up of the

graphics has a quasi-random distribution of the lines in the space. This is important as it determines the rhythmic identity of each sequence.

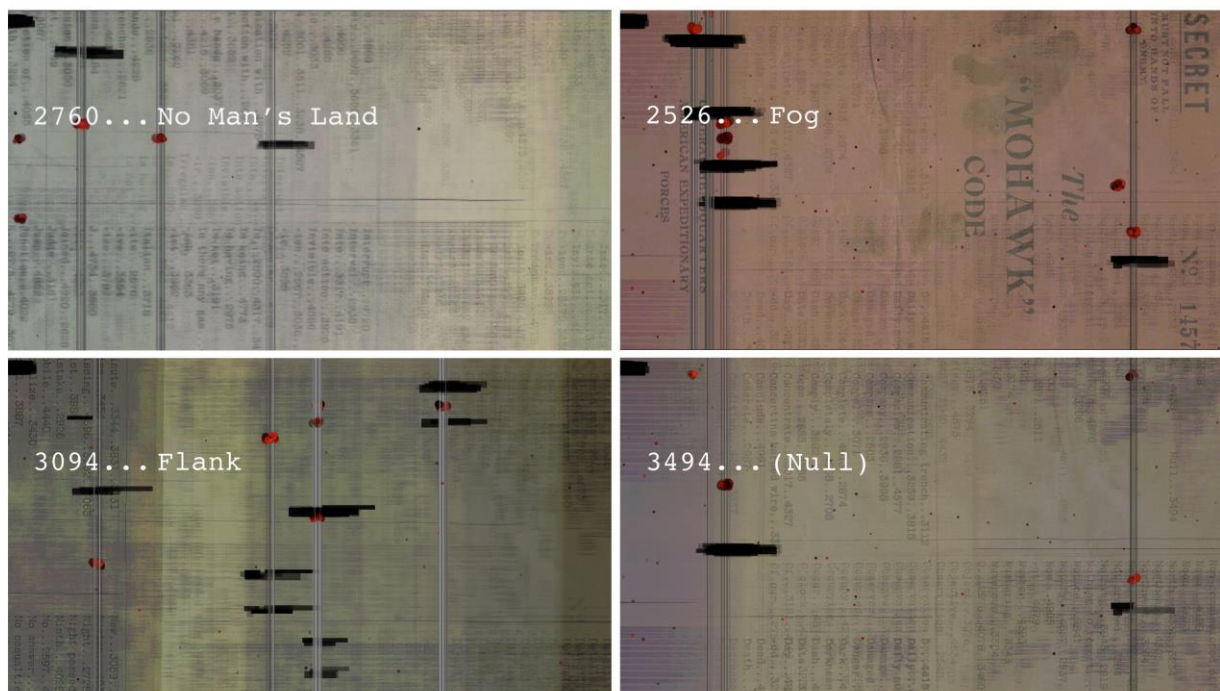


Figure 55 Randomly generated screenshots from *Mohawk*.

The lines are interpreted as rhythmic divisions of a given duration or tempo. A 'bar' is the X-axis of the projection. The circles are events on those rhythmic divisions. In the example above (top left) the bar might be subdivided into something like (2)+2+3+7 - with some events happening on those points indicated either with red ellipses or black bars. In time, other vertical lines might appear. This is all reinforced by the sound, which is automatically produced by the score. Five randomly selected percussive sounds are assigned to each of the vertical line regions and repeated over the period of anything between 1-10 seconds (chosen by the player). So again, the exact subdivisions of the bar can be heard and felt by the players, without needing to visually calculate the exact rhythm. However, having the visualisation gives the musicians an idea of the musical material that can be used: red ellipses could correspond to more pitched material of a given range, and black bars could correspond to something more noise based. The resulting musical character of Mohawk is that of precisely articulated rhythmic patterns unfolding over time, with sometimes sudden changes of speed, alternating stasis, chaos and movement.

On top of the graphic and text layer of the score, there is a background layer of scans of the original code book, randomly called up and manipulated visually with movement and blurring effects. These are controlled by the player, and although it was not the original intention, because of the varying processing drain each

algorithm has on the CPU (processing power of the computer), it does have an effect on the tempo of the graphics, and therefore the musical result. I tried to correct this, but in the end accepted the occurrence of this glitch as being part of the nature of the score.

Schlüsselheft

Schlüsselheft was a tri-numeral code book introduced by the German army in 1918, before major offensives began that Spring. The book itself was not updated, but the key used to encipher the numbers, the '*Geheimklappe*,' was changed twice a month, and towards the last days of the war, it was changed daily (Dooley 2016: 66). Even though the Germans went to great length to ensure the security of the code book, even if it was captured, this so-called super-encipherment was regularly cracked by the allied intelligence services. The book itself has several sections including various keys and maps, that create the feeling of an all-purpose code book, compared with the *Mohawk* and *BAB* code books. The material I used in the score for this app is not the tri-numerals and their associated words, but rather the '*Buchstabenzeichen*' (letter character) encodings of the '*Allgemeine Verkehrszeichen*' (general traffic signs) and the '*Zeichen für Fliegerdienst*' part of the book. These consist of two letter encodings of general phrases, where the two letters used are also marked within the phrase in bold letters (see below right).

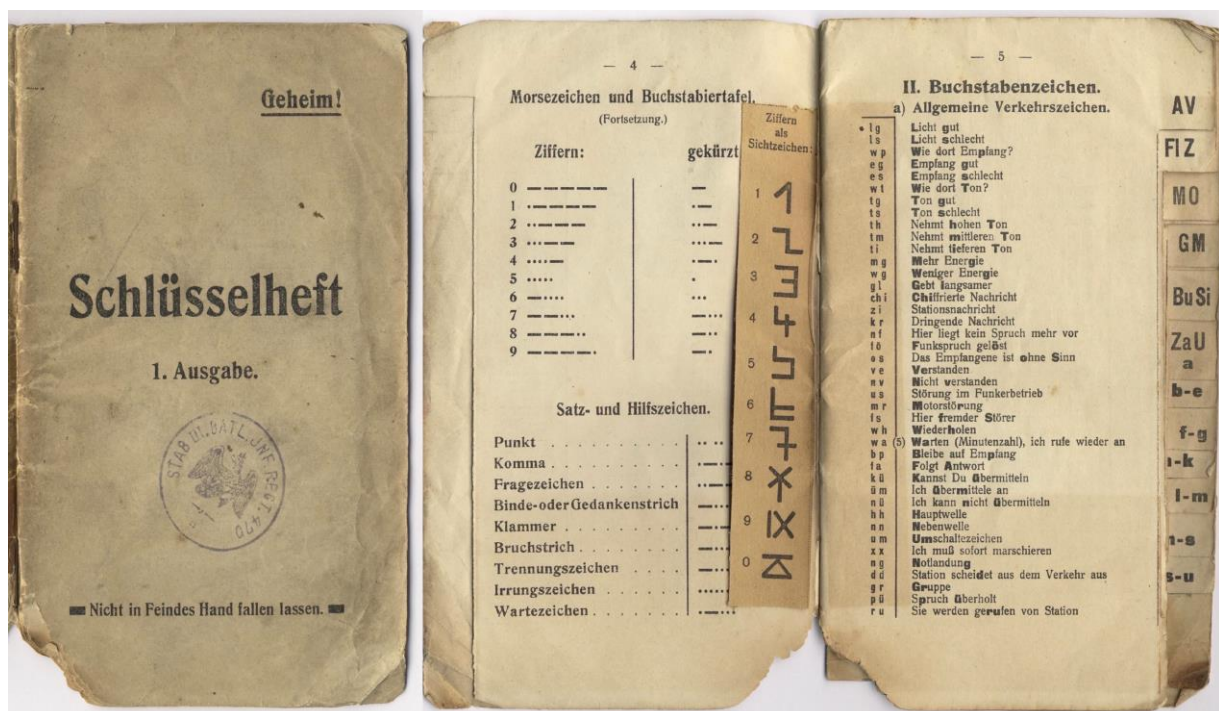


Figure 56 Cover and sample page from *Schlüsselheft* (1918).

I took the idea of how maps are segmented into quadrants and labelled with letters as the main visual concept of the graphic encoding. The screen is separated into 24 quadrants of 6x4, and the 24 letters used are assigned to these in alphabetical order. When a phrase is called up, the two letter code is rendered as a line drawn between two segments of the screen. The line is not always straight but might be segmented into two parts. As phrases are called up by the app operator, the screen is gradually filled with lines of a certain geometric order. Beneath this, maps of trenches (around Arras) are called up and similarly to *Mohawk*, are rendered with some random blurring and other image effects.

The clue as to how to musically interpret these lines is found in the sonification heard in the score, which is made up of slow glissandi, taking the Y-axis as the pitch axis and the X-axis as the time axis. In the app these glissandi are in fact filter resonances of noise, where the bandwidth of the resonance filter is randomly called up each time the mouse is clicked. There are 10 different pitch ranges of the glissandi called up by the score operator, which correspond to 10 different scales of the maps shown. It is not entirely necessary for the musicians to follow the exact pitches heard in the score, but they are encouraged to relate to it in some way, as well as to simultaneously relate to the visualisation.

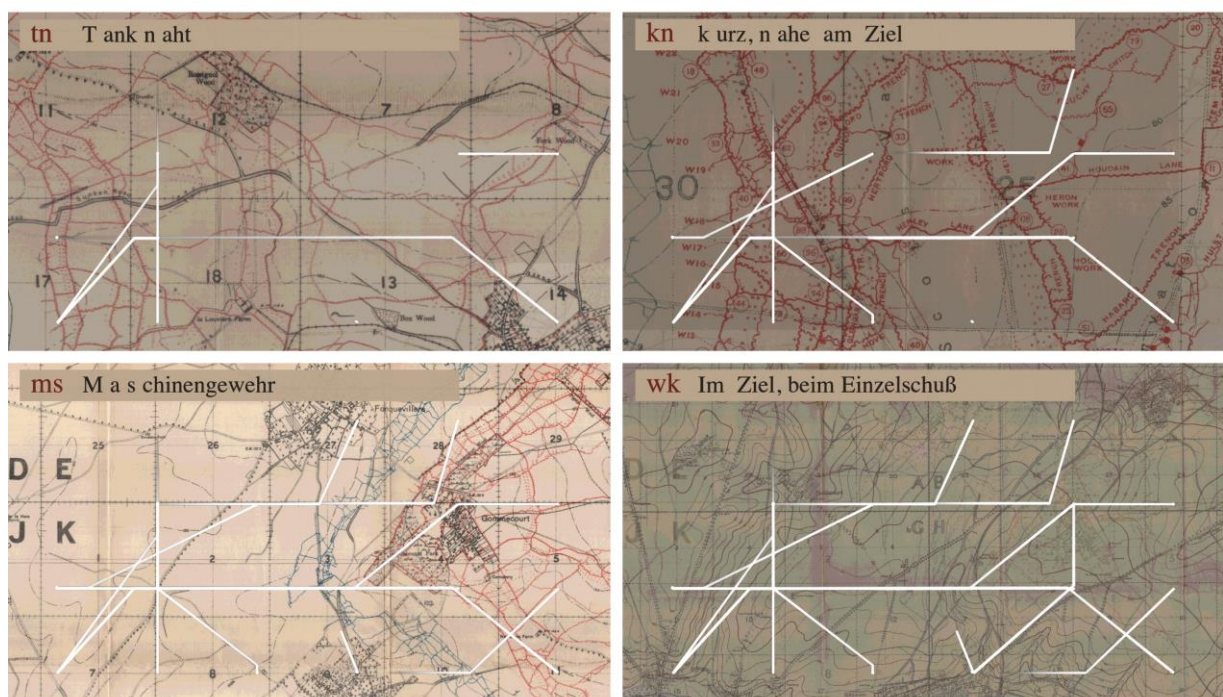


Figure 76 Randomly generated screenshots from one sequence of *Schlüsselheft*.

BAB

The third code book used is the "BAB" *Trench Code No.4*,⁹⁹ used by the British Army from 1917. This is similar to *Schlüsselheft*, in that it is largely a tri-numerical code book, grouped in terms of subject matter, and written in the convention of many commercial telegraphic code books. Phrases that share the same words are grouped together and proceed through various syntactic variations (see below):

175 We / (...) is / are / being shelled
176 " / (...) " / " / " rifle grenaded
177 " / (...) " / " / " trench mortared

Another interesting feature of this code book, which gives it a somewhat poetic quality when read through, is the frequent occurrence of the phrase: 'no meaning'. Similar to the phrase 'null' in the Mohawk code, the corresponding code was to be used frequently in the message, in order to compound the complexity of the signal in case of interception by the enemy. From the perspective of 100 years on, 'no meaning' becomes a refrain of the pointlessness of this inhumane trench war and the lives lost.

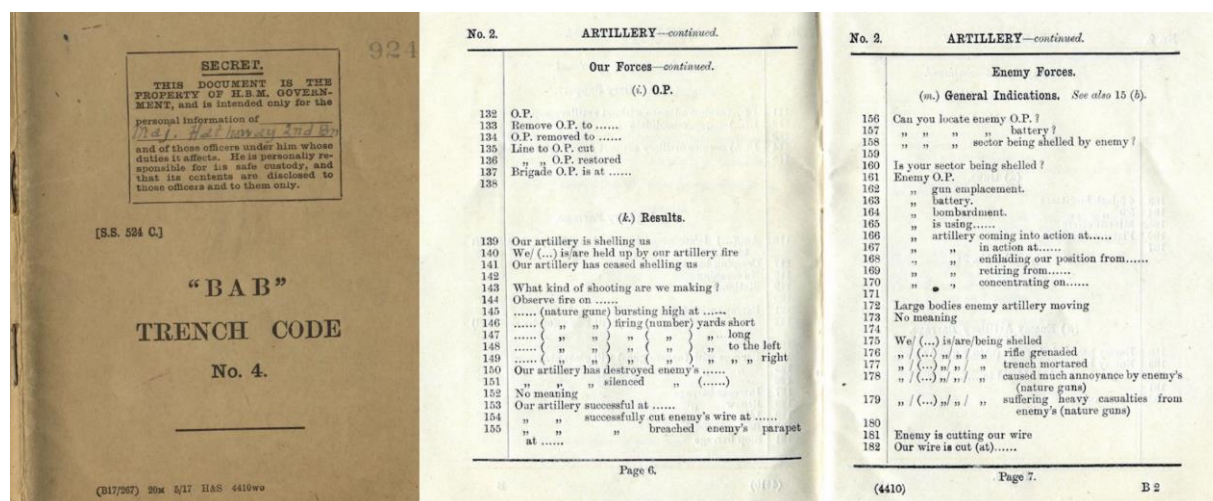


Figure 77 Cover and sample page from "BAB" *Trench Code No.4* (1917).

248 Are we to use gas?
249 You/(...) will make a gas attack
254 Gas has began to be released
255 " " ceased " " "

⁹⁹ According to The History of the 33rd Divisional Artillery in the War: 1914-1918, the code book was named after the initials of Lt Col B A B Butler: "The Division suffered great loss.... in the death through shell fire of Lt Col B A B Butler, of the Divisional Artillery. Lt Col Butler will always live in history as the author of the B A B Code, universally used throughout our army during the war as the official secret trench code, and thus named from his own initials" (Macartney-Filgate 2012: 143).

256 " " blown back
 257 Our gas is preventing our advance
 258 No meaning

The concept behind the encoding in this third score is based on the translation of the letters of the words are called up, in the order they appear in the books, into a geometric pattern of connected lines. At the points of these lines a number is called up which refers to either an instrument or a particular sound decided by the player. Each new word calls up another number, and the musicians are free to interpret the position of the event in the space in their own way. In the sonification of this pattern, the numbers called up, trigger specific noise samples connected to that number, and are modulated depending on their X-Y position in the space - X determining the length of the sample, Y the speed of playback.

Similar to the other scores, the laptop player controls the speed of the text, the refreshing of each of the line 'trees', where the numbers hang off, and the background collage of scans from the book, which in this case are split into 3 areas with different blending and tint algorithms. The character of this 3rd part of *Trench Code* is more soloistic, noisy, explosive, and speedy. The text and graphics appear in bursts of energy, contrasting with the slower build up of texture in the other two parts. There is also a clearer narrative component, as the succession of words is built up from phrases as they appear in the code books, and is more suggestive of meaning. This speed combined with semantic coherence, gives it a more voice-like character, in contrast to the more distant nature of the other two parts.

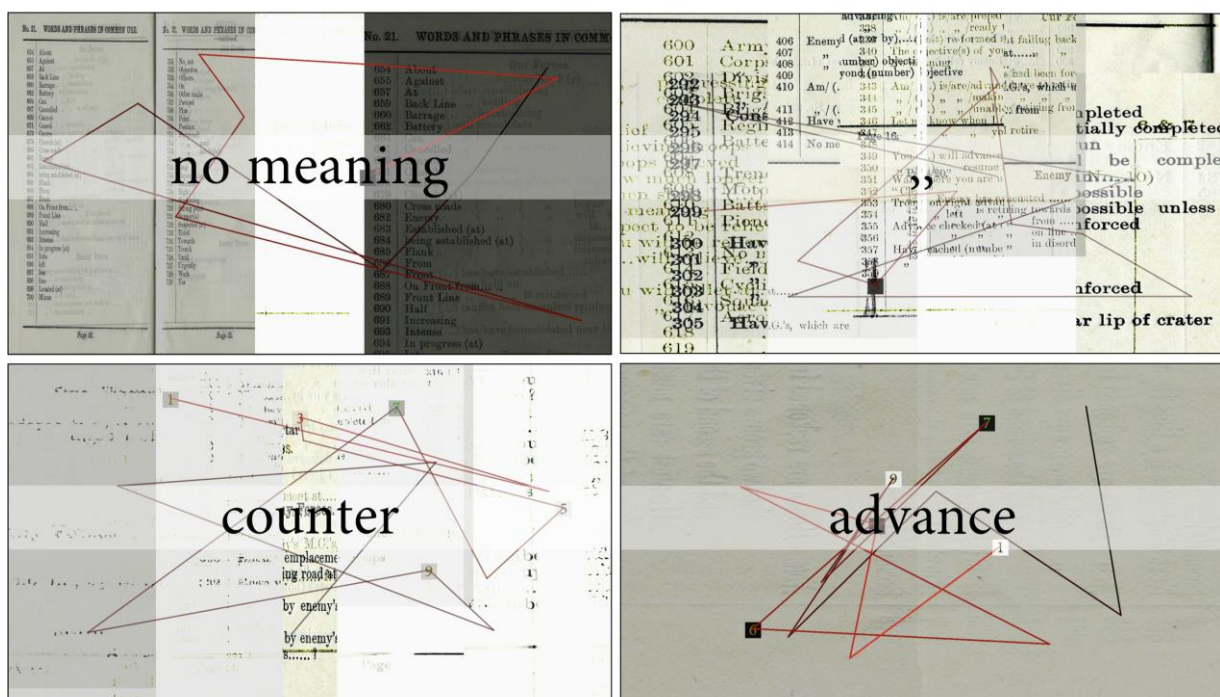


Figure 78 Randomly generated screenshots from BAB.

The way in which the musicians have to engage with the score, having to balance both sound and visual information at the same time, is characteristic of the three scores as a whole. The information is encoded in a parallel way, giving the musician space to interpret it from either one or both sources. The information is in one sense amplified, but the result is always slightly enigmatic because it is the encoding rather than the message that is being reinforced. This puts weight on the abstract nature of the mapping, which is constantly being undermined by the dramatic and weighty war references that crop up in the texts. Meaning is always being imposed onto the score, though what is being generated by the algorithm and the musicians, concerns the structure of the letters supporting the language alone, not the language itself. Seeing the words, both musicians and audience cannot help but reflect on said meaning, or the meaning that is being generated by their superimposition.

8.3 *Oneiricon*

Oneiricon is an interactive score devised for computer and tablets (iOS¹⁰⁰ and Android) based on an early Byzantine dream interpretation book (the dream-book of Daniel). Written for MAZE, it was first performed at the Angelica Festival Bologna, in May 2014¹⁰¹.

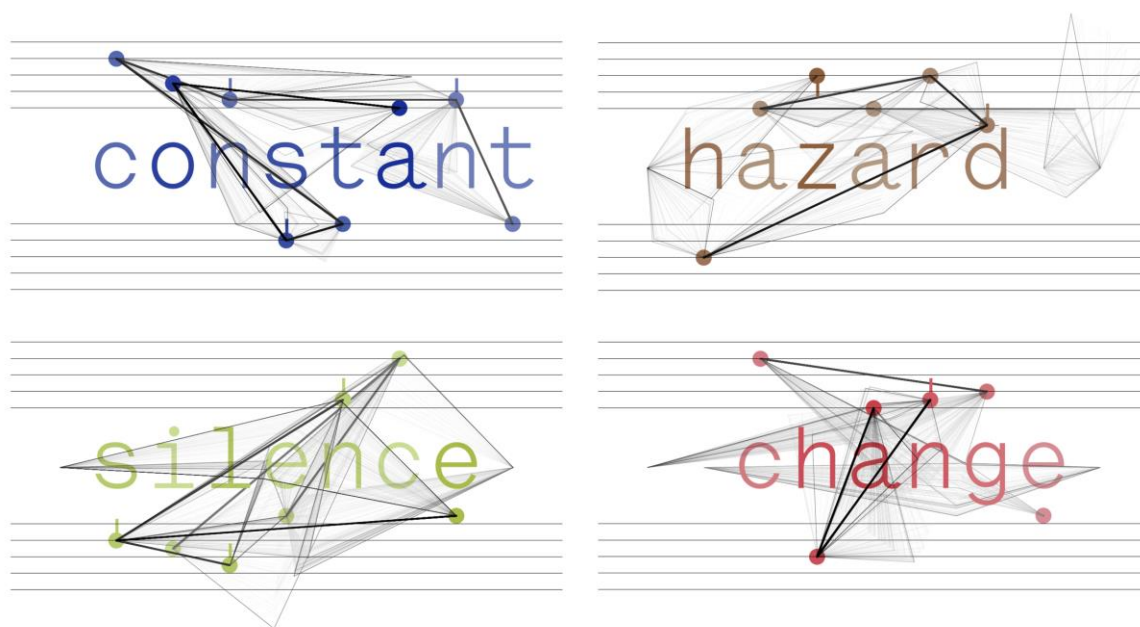


Figure 79 Randomly generated screenshot from *Oneiricon*.

Oneiricon is an open form composition, initially written in java script, on the Processing platform, making use of a self-generating, interactive score, that each

¹⁰⁰ Link to App Store to download app: <https://itunes.apple.com/us/app/oneiricon/id1293741939?mt=8>

¹⁰¹ Link to performance by Ensemble MAE: <https://youtu.be/GroHLc9QXTk>

musicians can control and navigate in. There is no universal score, but each musician controls and navigates through their own self-generating score-screen. This 'score' serves three functions: it generates a score, it serves as a basic sound generating instrument, consisting of sine waves and pulses, and it creates a visual text environment for both audience and players. Since the medium of the iPad or computer is multi-sensory and is capable of sensory input as well as output (unlike fixed media such as paper, film, or fixed audio or video playback), it is ideally suited for the creation of such a 'score'. (I use the word 'score' in inverted brackets, to highlight the fact it is not a score in the traditional sense).

The score-app itself has two modes. The first is the 'reading' mode, in which the player scrolls through a book, a dictionary of sorts, an 'oneirocriticon', which is a dream manual from the eleventh century. The players scroll through this word for word, controlling the speed by moving their finger on the screen from left to right. In this mode, each word appears one by one, centred on the screen, in the order it appears in the Oneirocriticon of Daniel¹⁰² (from which all the texts originate). The text obeys syntactical and narrative order. Phrases such as the following appear:

To dream of oneself beheaded signifies getting rid of great oppression.
Taking off one's clothes is good for sick people, but bad for all others.
A thistle plant sprouting up points to a rising-up of one's enemies.
A bear noiselessly approaching you signifies grief and dishonour.
Receiving a kiss from a dead person signifies life.

In this way the score is also an e-book of sorts that can simply be read, the intention being that a performance of the piece would give a strong sense of the musicians' act of reading, and the sharing of this activity with the audience.

The second mode of the 'score', triggered when the finger is removed from contact with the screen or trackpad, is one in which words slowly appear in random order, and are gradually morphed with each successive word. Letters appear one by one, in random order, completing each word, to be overplayed with the next word, letter by letter. The letters of the words are assigned to specific notes, based on a fixed mapping, created using a letter frequency analysis, so that the more common letters such as E, A or T, are mapped to more consonant pitch relations, and least common letters such as Z, Q, or X are mapped to more dissonant or distant pitch relations. As the letters appear, a note appears with it, placed on a traditional double five line staff. A simple indicator, pointing up or down, specifies sharps or flats. Further, no other traditional notation is used. A thick black line is generated between each successive note in order to suggest to the musicians a way of phrasing. Thinner grey lines are also generated by the movement of the finger and the position of the letters

¹⁰² Translation used: Oberhelman, Steven. 2008. *Dreambooks in Byzantium: Six Oneirocritica in Translation*, with Commentary and Introduction. Ashgate Publishing, Ltd.

on screen, in order to give feedback to the musicians. Once running, the score can play indefinitely, generating near infinite combinations of letter superimpositions and permutations, though because of the fixed letter to note mapping, tonality and melodic formulas tend to cohere.

In the first performances of this piece (by my own ensemble MAZE), I mapped a set of rules by which the score should be interpreted:

1. Navigate or play at will.
2. Sing notes in inner voice before playing.
3. Play notes as they appear - repeat notes if no change is seen on screen.
4. Change register freely but only one note at a time.
5. Think of stress and phrasing as if you are articulating the words - this can shift with repetitions.
6. Order of notes can shift but only one note at a time - new note can come in between.
7. Expressive elements such as glissandi (strings), timbral changes, multi-phonics (wind) and articulation can be used, but without expressive intent.
8. When new colour/word appear change to the next phrase.
9. Try to coordinate/interlock with each other (use pulse as reference if audible).
10. When full word appears (in play mode or after scrolling), speak the word quietly to yourself or to those near you.



Figure 80 MAZE performing *Oneiricon* at the Kontraklang Festival, Berlin, January 2017.

At the time, I had a specific interpretation in mind, one in which phrases would evolve in a similar way as they do visually, so that the music would come across as an act of reading, moving in and out of synchronisation with the ensemble, as if the reading was at times a private activity, at times a collective one. In subsequent performances of the piece by other ensembles, I felt that in the true spirit of an open score, I should leave the interpretation more in the hands of the performers; so apart from giving some technical information about the navigation of the score, I left the musicians to collectively decide how to approach the piece, how long it should be, how the musicians should be positioned in the space, how to relate to each other and how to create phrasing.

Instrumentation is naturally open (there exist versions of the score for different transposing instruments), and the work can be played with any level of musicianship. Moreover, because the score contains a sound component intended as a way of giving feedback to the musicians and acting as 'shadow' instrument, the work can even be played without other instruments or voices. The sound component is made up of two layers of sound, one is a sine tone playing back the pitch of the notes projected, with the possibility of microtonally bending the notes as they appear, by moving fingers along the screen. This is done in order to create complex tuning relationships and beating effects with the rest of the sine tones in the room and with the same notes played by the musician on their instruments. The other layer is a slowly modulating pulse, which can be manipulated in speed and in resonance, also by the position of the finger on the screen. The idea behind this sound was to create a subtle and complex rhythmic grid, which the musicians could freely use to synchronise the phrasing. The sound image I had in mind, was inspired by the effect of crickets chirping at night, a hypnotic tapestry of discrete pulses, spread around the space, coming from individual laptops or iPads, which would somehow resemble a nocturnal soundscape, to accompany the activity of dreaming.

The work investigates similar territory to that of the video-scores explored in earlier works, such as *Karaoke Etudes*. The sharing of the score with the audience, the connection between a text, which can be read by all, and notation that seems to have a link with what the musicians are playing, either in a direct or indirect sense. One could argue there is an open cryptographic approach to how the text and pitch information come across to the musicians and audience, and that there are semantic relationships formed between what we are reading and hearing, because they are slow enough to follow note per letter; though the kind of meaning they engender is, of course, ambiguous and subjective.

This ambiguity of meaning is a concept that resonates with the subject matter of the piece. The culture of dream manuals arose out of a desire to give fixed, unambiguous meaning to what has been one of the mysteries of human consciousness, to give signification to something, which seems to be communicating

with us, in the language of strongly emotive images. These 'Oneirocritica' books or records, offering the reader an interpretation of their dreams through the methodical description of key images, have a highly unambiguous ascription of meaning. There seems to be no doubt in the pronouncements that:

Dreaming of falling into a pond signifies danger, or
Drinking olive oil signifies illness.

What exactly the danger or the illness this pronouncement refers to is unclear, and clearly belongs to an impending future. This pattern underlies the idea behind dream interpretation in the pre-Freudian era, where dreams were seen to present the dreamer with an oracle, a future forecast from God(s). Dream books were written throughout antiquity, and were found throughout Assyrian, Mesopotamian, Ancient Greek to Roman culture, though the proliferation of 'oneirocritica' occurred during the Byzantine era (330-1453). Dream books from the Hellenic period were popular in Byzantine Constantinople, but being pagan, they presented a threat to the ruling Christian orthodoxy. In response to this popularity the patriarchy issued its own spiritually authorised readings, which seemed to strike the correct moral tone, as well as borrowing on many Arabic and Hellenic (Artemidorous) sources. The Byzantine Greek public could rest assured, that these interpretations were sanctioned by good authority, as they carried the names of patriarchs, emperors, prophets and saints (Oberhelman 2008:2). The *Daniel* dreambook from which the text for this piece is taken, is an early example of this, and served as basis for the many dream books to come in the medieval age. The motivation behind *Oneiricon* lies in taking this sense of authoritative and ultimate interpretation and opening it up to the ambiguity of meaning, that an encoding into music brings. Both text and musical narrative are deconstructed into half formed words and phrases, which in collective readings/performances would interlink into something, that seems to be loaded with meaning, though this meaning remains random, evasive and constantly modulating.

The exploration of the possibilities inherent in using new media for the creation of notation and scores, has been very much in my focus in the last years. The possibility of making scores self-generating, fluid rather than fixed, receptive to incoming data, and adaptable to different situations, is something that seems to be both a natural development, brought on by the shift from print technology to digital media, and a reflection on the changing relationship between composition, musicians and audience. The experience of working with a score like *Oneiricon* with ensemble MAZE, and discovering the many different approaches that one can develop through it, confirms my feeling that there is still much to explore. Specifically, the way in which the basic function of the score is redefined, as it becomes a score with ears, eyes, and mouth. The idea of reading and listening, that permeates the music-text-film pieces is relevant here as well, but with the veil between audience and musicians lifted: the musicians can in this case read the text, while the audience are able now to see the score. My current interest in developing these ideas further, is in

developing computer audio and visual input to effect parameters of how the score can self-generate, which might eventually involve using words in the form of speech recognition, but also listening to audio which is being generated by the musicians.

Rather than focus on a few large-scale works, I decided to present a broader output of work (still only a fraction of my creative work from the last years), as the specific area of research in each piece allows a significant perspective on the areas of research I have set out in the first part of the thesis, namely the ideas of inner voice, narration and dynamic perspective between media. The four distinct areas of focus that these pieces are grouped in: the inner voice, the idea of translation between language or other media and music, the materiality of voice, and the exploration of interactive scores, have naturally many shared aspects between them. This also includes the works briefly described in the Appendix. What they also all have in common, is the focus on the creation of a space between music and language, involving audio with on-screen text, which I have defined as music-text-film. Just as I explain in the introduction, this is not an absolute or even elegant definition of a single medium, and because I still do not consider this genre as one entity, like cinema or ballet, but rather a hybrid relation of different media in constant dynamic change, its tripartite definition seems more than adequate.

Conclusion

When I embarked on this research in 2011, I was initially motivated by the idea of digging deeper into the form of music-multimedia, which I had already began exploring in music-text-film, in order to try and understand what the spectator was actually experiencing during these pieces. I was intrigued by the function of the inner voice in my music-text-film, on which I somehow had stumbled accidentally, and I wanted to examine it more closely. I was and still am fascinated by the experience of reading text to music, the way the internal voices come to the fore, and become entangled with the external voices in audio-visual work: the way the inner voice can focus the listening perspective, the way the words can hide and reveal a detail in the sound or the way this voice can hinder or assist a spectator's entry into the musical space. Out of this were born three separate issues, which I felt were within my power to examine in the thesis:

Inner Voices

One was to try and understand as well as articulate, how the voice in the head was being induced. I was convinced that this was a phenomenon experienced by audiences, and I wanted to decipher under what conditions this became stronger or weaker. It soon became clear that slowing down the text rate and leaving space around the words, while articulating them with a sound, created the clearest awareness in the listener, through the phenomenon of 'silent reading'. In this way, one could catch oneself saying the word, between the resonant gaps, where the word can be heard entangled and echoing with surrounding sounds. Putting more weight on the melodic inter-relations of the resultant sounds and using a certain level of tonal coherence or repetition, increases the possibility of the listener being able to follow the music with their voice, 'silent singing'.

Bringing these two inner voices together would occur most clearly by combining a coherent melodic pattern together with a text, leaving some gaps of notes and words in between, in order to create a change of perspective and an awareness in the listener. To trigger an awareness of the 'silent discourse', would call for a situation having more critical-distance rather than immersion. There should be enough space, distance from the music or artwork, perhaps even boredom (which in my view does not have a negative connotation) so that the mind can wander, yet still being aware of the focus of the event and able to reflect on the affects afforded by the music or artwork. A question and answer structure, as I described in Chapter 6, where there is space for the answer to be supplied by the spectator, is another such way in which 'silent discourse' comes to the fore. The above are just examples of how the conditions for awareness of the inner voice can be partially created in a musical

structure. They are certainly not exclusive to the process, and this is also not intended as a formula, as the conditions are mostly supplied by the spectator.

Perspective

The second issue I wanted to examine, is where the perspective is being created, in effect where is the voice? Looking into narrative theory was an elucidating but also complicated matter, when it came to translating the same concepts into music. What seems relatively straightforward in literature, the differences between perspective, point of view and focalization, becomes rather more complicated when translated into music. There have been bitter battles fought amongst the narrative theoreticians about the exact meaning of the three terms, which at first seem quite similar, so rather than miss-translate a term into musical use, I decided to use the word 'focus' instead of focalisation, whilst keeping the word 'perspective', because it can be applied on a more general level. 'Perspective' became a term that seemed to apply more easily to the relation between a listener and the musical space. Furthermore, the idea that a narration can take place only when a difference in ontological level was perceived, was a concept that brought many things into light, specifically about the role of perspective. The narration, or what I eventually call the 'narrational network', created between these ontological levels, are discovered through the perspectives of the viewer, which are 'focused' by specific relationships in the music or other media.

Focus

The third issue was the 'how' was this process occurring? I wanted to examine how the differences in ontology, for which I borrowed the term 'frames', where being created in the artwork. I began by analysing media relationships, which was a useful way to explore how the relation of media effected the general perspective of the spectator. This lead to the analysis system, which I call 'media correlation'. Where a theory emerged, was through looking at the metaphorical relations between the media, and asking why is one medium the 'target' and the other the 'source'. This became the answer to how 'focus' is created within the artwork, which constructs the perspective for the spectator. This focusing through metaphor hierarchy was itself a dynamic process, which I hope to have demonstrated in analysing *Subliminal: The Lucretian Picnic*, one of my first music-text-film works.

Music-text-film

As stated in the introduction, this thesis is created primarily as a 'poetics' of music-text-film, as a way of understanding the issues, dynamics, concepts and affects of

what the medium affords. With some of these theoretical tools outlined in the first three chapters, I discussed the ideas and realisations of a substantial body of work.

The research also fed into my own creative output in a way I had not expected. Most radically, in the idea of trying to open up the listening perspective, which I thought ran the danger of being too influenced by the words, especially in the earlier music-text-film. I did not want to create a situation, where the text explains or provides a fixed meaning to interpret the music through. The new directions led to more experimentation with deliberate perspective changes, interactive strategies and generative scores, which I thought would highlight some of the narrative complexity occurring between the different media.

The creative work that has resulted in over thirty works of music-text-film presents an exploration of a specific form of audio-visual work, where music, text and image (or graphics) are set up to highlight ways of listening through other media that puts emphasis on the role of the listener/spectator. In many of the works, the different silent voices that are articulated in the first chapter play an important role in the 'completion' of the media relations that are presented. This I feel is something that is often overlooked in much audio-visual and multimedia work, that aims for saturation or amplification of meaning by the multiplication of media. The recent generative and interactive work which has explored a more free approach to notation and the relation between performer and audience, creates an even more dynamic situation where the perspectives of both parties, interfacing through various forms of music-text-film are further blurred and entangled.

I have tried to highlight the different ways the idea of 'voice' can function in some of my music-text-film, as a way of articulating the dynamics of multimedia work in general. The shifting perception of what 'voice' actually is, is a compelling aspect of this form of music-text-film, as it fluctuates from a purely narrative form, to a voice as sonic expression, to the audience becoming aware of their own inner voices as they read the projected text in resonance with the music. The question of what constitutes a voice is ultimately at the heart of this, as the voice moves from being a carrier of meaning, of narrative, to determining the way our attention shifts between different layers of media.

This is a strategy I have used in the way visual information is supplied sparingly in these works – the deliberate avoidance of creating a too dominant visual field, which can monopolise the attention of the audience and push the semantic significance of the sound into the background. Achieving an 'asymmetrical balance' (if that is not an oxymoron) between the media – the text, the music and the visual form – is one of the key aspects in maintaining the dynamic shifts in perspective, necessary to avoid one of the media becoming either redundant or too dominant. As well as highlighting the built-in inequality of metaphoric relationships, the idea of

'asymmetrical balance' offers an image of something always on the precipice of change; a dynamic, volatile construction, which, throughout the course of a piece, could shift in perspective, and challenge the audience into different ways of relating to the material. Meaning will undoubtedly be constructed, especially when words are involved; but, as Derrida argued (Lechte 1994: 109), words, like music, are also inherently unstable.

In the combination of sound and projected words in my music-text-films, an environment is created, where it ultimately becomes unclear who is narrating. There is a blurring between musical structure and the inner voice of the reader, which can lead to an engaging and immersive experience, and may even reveal something about how music communicates.

Appendix: Additional Works

In the following pages, fifteen supplementary music-text-film pieces of mine are briefly described and some aspects of the media correlation discussed. These pieces relate in some of the themes and techniques to the pieces already discussed, thus providing a further viewpoint on possible approaches which can be further explored.

They are arranged in chronological order, from the first piece *Simplex* (2005) to *The Musicians of Dourgouti* (2017).

Simplex
The Queen is the Supreme Power in the Realm
Scam Spam
QFO
RE: Mad Masters
The Arrest
Circadian Surveillance
Nerve
True Histories
8'66 (or everything that is irrelevant)
Walls Have Ears
Music for Anemic Cinema
MacGuffin
Lost Border Dances
The Musicians of Dourgouti

Simplex

Simplex for ensemble and text-film. Commissioned for the 25th Anniversary of Maarten Altena Ensemble (Ensemble MAE). Premiered on the 13th November 2005 by Ensemble MAE at the Muziekgebouw, Amsterdam.



Figure 57 Stills from *Simplex*.

This work was the first of three that used material from Victorian era telegraphic code books (the others were *The Queen is the Supreme Power in the Realm* and *Telegraphic*). The codebook used in this piece was the 'Simplex Standard Telegraphic Code' by Edward W. Reiss (New York, 1911).¹⁰³ This is known as a 'commercial code', once used to save costs on telegraphic communication by compressing phrases into single words or sequence of letters or numbers.

The piece simply scans through one page of this book starting from 'If it is not a fact' and going to 'If you have no faith in'. The coded text letters are themselves encoded into five-part arpeggiated chords played by a microtonally tuned synthesizer. This forms the basis of the composition, while the rest of the instruments elaborate on the meaning of the phrases with a more ornate material. The music stays relatively static and expressionless, leaving space for the more diverse meaning coming from the text, proposing an interpretation of the music. Because the music, like the code, is largely based on a logical incremental system, it remains as 'target' in the metaphorical relation, and the text becomes a 'source' to provoke meaning from. This does not remain static, though, because the text also belongs to an alphabetic

¹⁰³ I became aware of these codes thanks to John McVey's excellent website and work: <http://www.jmcvey.net>

order, and especially during phrase variations on single words, there are shifts of hierarchy between words and music that keep the tension suspended throughout the piece. The text projection was realised using a powerpoint projection, consisting of a slide for each line of code, displaying the code number, code letters and phrases in different graphic hierarchy. These were then triggered together with the on-stage synthesizer, so as to retain the flexibility of performance.

225--IFGUW--IFTOH		
13207	IFGUW..If it is not a fact	13207
13208	IFHAS..In fact	13208
13209	IFHET..It is a fact	13209
13210	IFHIV..It is not a fact	13210
13211	IFHOW..What are the facts	13211
13212	IFHUX..FAIL(ED)	13212
13213	IFIGY..FAILING	13213
13214	IFILD..After failing to	13214
13215	IFING..Are likely to fail	13215
13216	IFIPH..Are reported to have failed	13216
13217	IFISK..Do not fail	13217
13218	IFJAT..Do not fail to	13218
13219	IFJEV..Everything has failed	13219
13220	IFJOX..Expected to fail	13220
13221	IFJUY..Failed to	13221
13222	IFKAV..Has (have) failed	13222
13223	IFKEW..If they do not fail	13223
13224	IFKIK..If they fail	13224
13225	IFKOY..If we do not fail	13225
13226	IFKUZ..If we fail	13226
13227	IFLAW..If you do not fail	13227
13228	IFLEX..If you fail	13228
13229	IFLOZ..If you have failed to	13229
13230	IFLUB..Is likely to fail	13230
13231	IFMAX..Is not likely to fail	13231
13232	IFMEY..Must be done without fail	13232
13233	IFMIZ..Must not fail to	13233
13234	IFMOB..Not expected to fail	13234
13235	IFMUC..Not to fail	13235
13236	IFNAY..Party named has failed	13236
13237	IFNEZ..To fail	13237
13238	IFNIB..We have failed in	13238
13239	IFNOC..Who failed	13239
13240	IFNUD..Without fail	13240
13241	IFOFY..FAILURE(S)	13241
13242	IFOMB..Are concerned in failure to the extent of	13242
13243	IFOLF..Are not concerned in failure of	13243
13244	IFORK..Is a failure	13244
13245	IFOWP..Is not likely to prove a failure	13245
13246	IFOYS..Is sure to prove a failure	13246
13247	IFOZT..FAIR	13247
13248	IFPAZ..Consider it fair	13248
13249	IFPEB..Do not consider it fair	13249
13250	IFPIC..Do you consider it fair	13250
13251	IFPOD..If fair	13251
13252	IFPUF..If not fair	13252
13253	IFRAB..Unless fair	13253
13254	IFREC..Will be fair	13254
13255	IFRID..Very fair	13255
13256	IFROF..FAITH	13256
13257	IFRUG..Are losing faith in	13257
13258	IFSAC..Bad faith	13258
13259	IFSED..Done in good faith	13259
13260	IFSIF..Good faith	13260
13261	IFSOB..Have faith in	13261
13262	IFSUH..Have lost faith in	13262
13263	IFTAD..Have no faith in	13263
13264	IFTEF..Have you any faith in	13264
13265	IFTIG..If you have faith in	13265
13266	IFTOH..If you have no faith in	13266

Figure 58 Page from the 'Simplex Standard Telegraphic Code'.

The Queen is the Supreme Power in the Realm

The Queen is the Supreme Power in the Realm is an improvisational environment for ensemble, live electronics and video. It was commissioned by MusikFabrik, ZKM and the Köln Triennale 2007, and was premiered in Köln in May 2007 by MusikFabrik with subsequent performances at ZKM, the Huddersfield Festival 2007, Moers Jazz Festival 2008 and Ultraschall, Berlin, 2012. Live video by HC Gilje.



Figure 59 Photos from the performance of *The Queen is the Supreme Power in the Realm* at the Köln Triennale.

The title of the work is a reference to Slater's Telegraphic Code (1870). This is a codebook of single words, indexed with numbers, which was used at the end of the 19th century as a basic form of encryption for sending telegrams in secret. Telegraphic codebooks, specifically the *ABC Telegraphic code*, 5th Edition from 1901, form the main cultural reference point in this work. As described in *Simplex*, these flourished at the height of the industrial revolution, at the turning point of the new world order when the power of the British Empire was at its peak, and were mostly used as a means for the industry to send shorter, cheaper telegrams by substituting single words or numbers for commonly used phrases. Economy of language opened up the possibility of a faster means of communication:

2 6 8 6 4.	nageklost.	natives very quiet.
2 6 8 6 5.	nagekneed.	natives very unsettled.
2 6 8 6 6.	nagelag.	natives becoming very troublesome.
2 6 8 6 7.	nagelartig.	natives rebelling.
2 6 8 6 8.	nagelbein.	natives rebelling and very excited.

2 6 8 6 9.	nagelfell.	natives becoming beyond control.
2 6 8 7 0.	nagelhout.	natives settling down quietly.
2 6 8 7 1.	nagelkram.	natives have now settled down. ¹⁰⁴

The musical material derives from various forms of coding of language into sound. These are used to create flexible structures for open scoring. Although the piece has a very clearly defined sequence of events in its definition of the material, the microstructure is open and dependant on decisions made by the musicians. Sound processing plays an important role here, and the patches created in Kyma, are in part controlled by envelope and pitch tracking of the incoming audio signal. Language is coded into music and used as a score for further possible interpretation. The composition consists of the definition of six 'zones', which are defined by their placement on the stage, and also by the type of function they have in the overall structure of the work. The 'Drones' are tutti sections of continuous tones that interact with the sound environment in different ways. The 'Feeders' play into four microphones placed in the corners of the space, and their sound is algorithmically cut up into pulse fragments. The 'Translators', using headphones, translate spoken text material into music using various rules, this modulates from pitched to un-pitched throughout the piece. 'Encoders' use a personally constructed alphabet of sounds to encode words into sound gestures. The percussion has both a percussion alphabet and a 'morse' based system to trigger samples from rhythmic patters. The piano is at the centre of the system and is the only instrument playing from a fixed score. Through its timing of phrases, and together with the computer operator the pianist controls the flow of the piece.

The Queen is the Supreme Power in the Realm is a quasi-hierarchical system (as reflected in the title), that has a passing reference both to ideas about the Victorian empire, and to bee culture, with its complex colonial structures. The rules consist of the type of material to be played, what to listen for and how to react to the 'zone's happening at the same time. The improvisational aspect of the piece is controlled by these definitions, but puts greatest priority in how the musicians respond to what is occurring around them. In this sense the image of communication, sending and receiving messages, comes to the fore in the piece, as well as the sense of a fragile order, which has the potential to disintegrate at any moment.

¹⁰⁴ A list of phrases used in the piece from the 'ABC Telegraphic Code, 5th Edition' (1901)

Scam Spam

Scam Spam is an eight minute piece for violin, electronics and text-film, first performed by Barbara Lüneburg in 2007 and later recorded by Takao Hyakutome.¹⁰⁵



Figure 84 Stills from *Scam Spam*.

Virtuoso violin playing, both real and hyperreal, underpin a video made up of spam email texts. The piece moves at a fast pace through a myriad of elaborate scams, random spam poetry and 'phishing' emails, collected by my spam filter over a few months in 2007. These are presented each with a distinct graphic form, ranging from oscillating words to large scrolling blocks of metadata.

There are two distinct layers of sound. The violin part is based largely on fast arpeggiated flourishes, crossing all four strings, with a considerable use of harmonics. The character that I wanted to communicate was one of a fast, fleeting legerdemain of a scamster's trickery taking place in the full light of day, and resonating with the folkloric image of the devil as a fiddler.

The electronics, which are synchronised tightly to the video and violin part, consists of two parts. The first layer is of sampled violin sounds: mostly sub-tones, noises, scrapes and percussive sounds projected in looping isorhythmic patterns across four speakers. These patterns are created using a special patch made in Supercollider that varies sample choice, with panning position, rhythm, filter, and volume, in complex but repeating patterns. The second layer of electronics is constructed of wave based

¹⁰⁵ Recording of *Scam Spam* by Takao Hyakutome: <https://vimeo.com/39011042>

synthesizer sounds, pulses and drones, that are at times realised in the same way as the violin samples, and at others used to create alternate polyrhythmic or harmonic layers.

An interesting observation regarding the narration about this particular music-text-film, is that it is one of the few of my pieces where the text is in the second person, addressed to the listener.¹⁰⁶ This amplifies the sense of communication between the music and the audience, because the text is constantly underlining the idea that the music is communicating the message being read. However, because we know that the text originated as spam, and it is already one step removed from a sense of direct communication, we do not take it as addressed to us personally. These are emails that hardly anybody reads, that are written in a way as to be able to get through the algorithms of spam filters, with keywords in alternative spellings, and blocks of pasted text from random literature. Towards the end of the piece, there are supposed instructions about creating your own spam emails, though this, like everything else in the piece, is revealed to be a very surreal type of 'clickbait':

Subject: Post-Impressionism

Step 1: Simply put your cursor at the beginning of this letter. Click and hold down your mouse button. Check spouse and staff.

Step 2: Investigate your own credit history. From the edit pull down menu paste. Remember to eliminate the first position and move everyone up a spot. Disappear in you own city.

Step 3: Save your notepad file.

Step 4: Go to the pull down menu entitled 'window'. Select Cezanne. After a few moments a list will show up on your server. Click on any you desire. Another crucial bit of evidence.

Step 5: Fill in the subject. Your own conduct.

Step 6: Highlight the entire contents of you .txt file. Choose a sandwich.

Step 7: Hit the send button in the upper left hand corner. You're done with your first one. Congratulations.

Have a very nice day.

¹⁰⁶ *Circadian Surveillance* is the other piece, where the text is in the second person point of view.

QFO (Queer Foreign Objects)

*QFO (Queer Foreign Objects)*¹⁰⁷ is a set of 50 slides created as an interactive Flash-based web app, which navigates through Swiss scholar Thomas Platter's description of his 1599 visit to Sir Walter Cope's 'Wunderkammer'. It consists of illustrations and design by Isabelle Vigier and a soundscape of between 1 and 15 seconds corresponding to each slide. It was commissioned and created for the new UK 'Sound and Music' website in 2008.



Figure 60 Screenshots of images by Isabelle Vigier from *QFO*.

The idea behind the use of images was to overemphasise the falsehood, underpinning the relation of text to object. The soundscapes are multi-layered, with elements, which can form associations to the object is being described, but also elements that contradict this, or pose problems.

QFO is an example of how the mind processes one piece of information through the frame of another. This is evident, as the frames that are set up in the work, including the text, the image, the harpsichord and the sound, are for the larger part constant, though they vary in the weight of information assigned to them. Whether one experiences the work from one particular perspective rather than another is highly subjective, but the work tries to bring the idea of perspective to the fore, so the viewer can challenge the credibility of what one medium tells about the other. The subject of colonialism is also brought to the fore, highlighting the way in which 'the

¹⁰⁷ The work can be accessed here: <http://earreader.nl/wp-content/uploads/2010/11/kyriakides.html>
Adobe 'Flash' is needed, so this might not work on Apple smartphones and tablets.

other', whether another medium, another culture, or nature itself, is always subject to some misunderstanding.

In the text Platter describes objects as varied as: '*A unicorn's tail*' - '*Remora. A little fish which holds up or hinders boats from sailing when it touches them*' - '*an embalmed child (mumia)*.' According to some writers (*The Collector's Voice: Critical Readings in the Practice of Collecting*, edited by Susan Mary Pearce and Alexandra Bounia), Platter seemed to be impressed by the uniqueness and diversity of this particular cabinet of curiosities, a collection undoubtedly augmented by Cope's travels in the 'Indies'. This 'Wunderkammer' presents spurious natural histories, assumed cultural artefacts, mythologies and fake folklore, which are mixed together to form a patchwork concept of the world beyond the known and 'civilised'. This way of explaining the world in uncategorised objects and words, was to form the basis of the institution of the museum. What these early cabinets reveal is how perverse and incoherent a practice it is to display cultural items like these, removed from their original context and juxtaposed with one another.

What is interesting in this text and relevant in an exploration of multimedia art, is not only the curious juxtaposition of text and object, but the fact the very object that the text refers to, is no longer there. There is an activation of fantasy at work in the viewer's act of engaging with both the cabinet and the text – how a world or a context is imagined by the symbolism of a particular object. Meaning in *QFO* is created from the discrepancies that emerge between the idea of the object that is described by the text, the illustrations, and the sound that underpins each slide.

We are constantly being fed with fabrications and yet our mind defaults to accept these without much struggle, because meaning, intended and accidental, is being created by the juxtaposition of these media, and the framework in which they take place has been clearly demarcated. This is a cabinet of curiosities, of 'queer foreign objects'.

RE: Mad Masters

RE: Mad Masters a 25-minute music-text-film for electric violin/violin, soundtrack and live electronics, was written for violinist Barbara Lüneburg and first performed at Deutschland Funk in April 2008. The subject and material are based on Jean Rouch's legendary film *Les Maîtres Fous*.¹⁰⁸

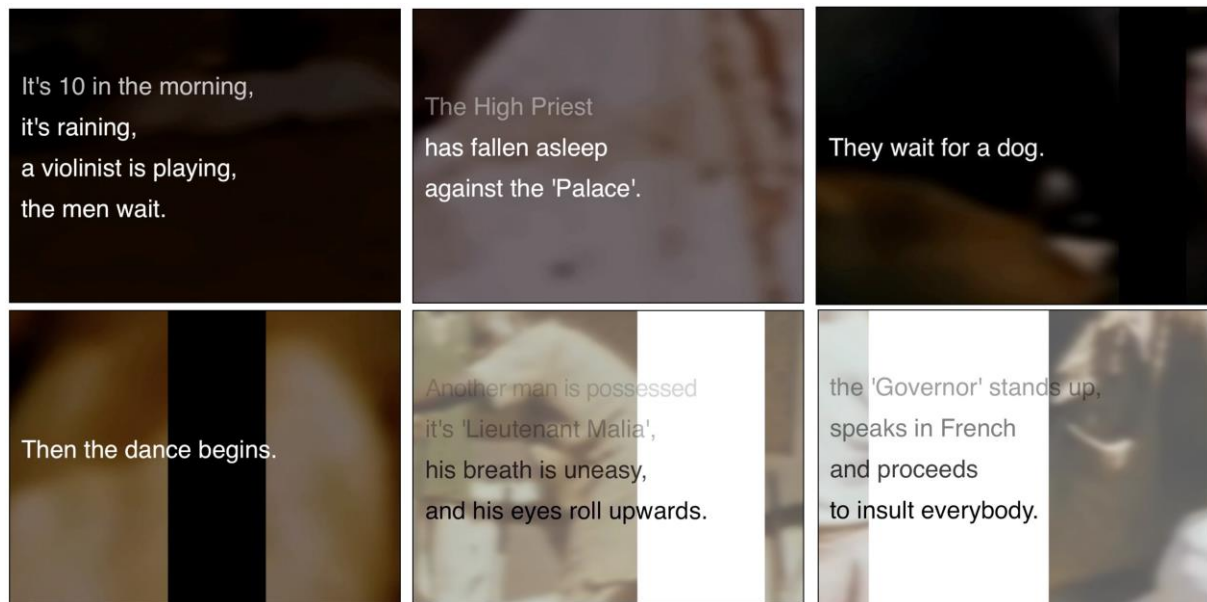


Figure 61 Stills from *RE: Mad Masters*.

RE: Mad Masters is a mash-up based on the famous 'Cinéma vérité' film *Les Maîtres Fous*, by Jean Rouch, which attempts to redraw the line between the objectivity of the camera and the themes existing beneath the surface. The dynamic between immersion and detachment is explored in how the point of view shifts between the observer and the participant in the possession ritual. The voice-over is presented as text. Music and sound act as a surrogate for the actual image; asking what role does the imagination play when it has to make up for absent narrative and media?

Rouch's original documentary is set in and around Accra, Ghana, and was filmed in 1954. He films and comments on a possession ritual, performed by the Hauka, a cult that sprung up in neighbouring Niger, and which was imported into Ghana as a result of the migration of workers into Accra, from which the social sector of the people in the film are largely drawn. The peculiar aspect of this possession ritual is that the sect members, taking on the personae of their British colonial masters, having reached a visibly possessed state, decide to kill and eat a dog as a display of the transgression of social taboos which their powerful gods can enact.

¹⁰⁸ Recording of *RE: Mad Masters* by Barbara Lüneburg: <https://vimeo.com/225816862>

To what extent are we supposed to take the possession ritual as a form of performance? Is what is been enacted by the cult members in their 'possessed' state a form of proto-theatre? One of the many things that have been said about Rouch's film of the Hauka, is that it forces the viewer to 'decolonize' his mind [Stoller 1992:160]. The viewer is forced to make sense of the confusing images he is seeing, by imagining something beyond his/her own 'European' mentality

The piece *RE: Mad Masters*, begins with texts describing what we are supposed to be seeing, text that are originally Rouch's audio commentary on what he is witnessing. They are edited only in as much as local references are taken out, names and places are removed, not so much with the idea of making it more universal, but rather in order to keep the viewer in some state of bewilderment as to what it is they are watching. Another strategy, used with this in mind, is the masking, defocusing, and scaling of the film material. The feeling that we can never be sure of what we are seeing is heightened by the manipulation of the cinematic window

The soundtrack is in part created by a handful of audio fragments sampled from the film: the plucked and bowed string music, the voices of the participants, sound of the wooden rifles slapped together, function also as a window to the diegetic space of the film, though this has to be shown by the performance of the violinist, and her enigmatic action in the concert space. At times, she plays the role of the possessed, at others the role of the master of ceremonies, or even the direct symbol, as a western musician, of colonial power. Her sound metamorphoses throughout the piece as if her violin has been possessed by another medium; the electric violin sound fades out and is used to trigger a slightly un-tuned upright piano (used as a symbol of a colonial music machine). The inspiration for this comes from the way the 'spirit' replaces its 'double' in the possession ceremony, to find a way for the electronic music to transform and displace the 'real' we hear both in the soundtrack of the original film and the acoustic sounds of the violin. This displacement of the 'real' soundtrack is a tool by which music can shift focus onto the inner rather than the outward space captured on film.

The Arrest

The Arrest is a 12 minute music-text-film for violin, clarinet(s), electric guitar, marimba, piano and contrabass, based on a dream text by writer George Perec. It was written for Ensemble MAE and premiered at Poetry International 2010.¹⁰⁹



Figure 87 Stills from *The Arrest*.

The text of *The Arrest* is from a dream narrative by Oulipian writer Georges Perec, found in a collection of 150 dream narratives he published in 1973 under the title *La Boutique Obscure*. The particular dream used in *The Arrest* is typical of one of Perec's recurring nightmares: being stopped and arrested by the police, a fear that he had been said to carry from his mother's experience, originally a Polish Jew, who was captured, deported from France and murdered during the holocaust.

Rather than articulating the narrative voice, the layers of the music are set up in a way in which the instrumental music serves as a fixed image, the ground. It acts as a voice without content, creating tension and stretching out the canvas of the narrative. The cinematic samples of found sound – such as dogs barking, a motor, street voices, a helicopter – act as a window to the dream narrative, sometimes in contradictory and sometimes in complimentary relation to the text. For instance, when we read: "the landscape is revealing itself like the background of an Italian painting", the sound of a motor comes into focus, giving us a contradictory idea about what could be an idyllic landscape. Or, when the narrator references his own guilt about the Israeli-Palestinian conflict, we hear the barking of dogs (who have

¹⁰⁹ Performance by Ensemble MAE, conducted by Bas Wiegers: <https://vimeo.com/14960327>

been in the background at several other points in the piece) coming sharply into focus, a self-conscious metaphoric relation, but one that can exist within the logic of dreaming.

The soundscape takes on a strongly metaphorical significance, because it is pushed into the foreground. The changed situation underlines the instrumental score as the inner voice, because words and musical phrases are synchronised ever more closely. Finally, an urban soundscape drifts in, perhaps highlighting the difference between inner and outer worlds, which signals the end of the dream.

The dream narrative itself hinges on the word 'copulate'. He feels that he is being hunted by the police for an inexplicable reason, which turns out in his mind to relate to his Jewish heritage while being in Tunisia, and the pretext for his arrest is having sex with his wife on the Sabbath. There is a clear in-dream pun in Perec's dreamt conflation of the slang word for police 'cop' and the word copulate. What is engaging about this dream account, and about many of the texts in *La Boutique Obscure*, is how Perec seems to be unravelling or analysing the reasons for actions and images at the same time that he is dreaming them. As if by writing these dreams, he is already distancing himself from the subconscious state he was in when he dreamt them and already interpreting them, permeating the narrative with a sense of revelation.

This way of manipulating the sound images in relation to the text has its precedence in some of the great films of Robert Bresson or Ingmar Bergman, where meaning and emotion are suggested by a carefully chosen sound object. What I found interesting to explore in this work, was how one can slip between what we perceive as sound information and what we perceive as musical information. This is most obvious in the samples that directly refer to North African music: a female voice, an ud, some Arab pop playing from a car radio. There is an obvious blurring between these samples and the instrumental music, but by their placement within the context of the musical narrative one is forced to question their meaning and function. There is a level of Orientalism at play here, partly because a sense of cultural 'otherness' is at the root of the dream narrative.

Circadian Surveillance

Circadian Surveillance was written and developed for The Electronic Hammer (Henry Vega, Juan Parra Cancino and Diego Espinoza) with Emanuel Flores (video). It was first performed at the Museum De Pont, Tilburg, November 2010.¹¹⁰

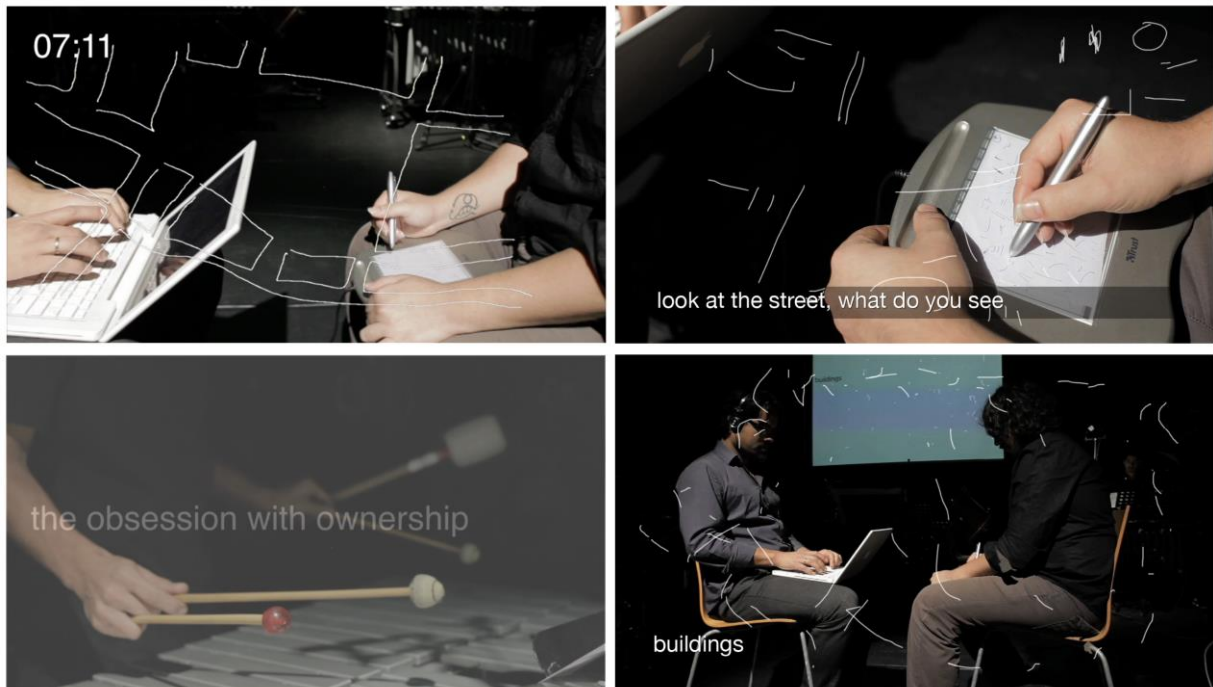


Figure 88 Photographs from the performance of *Circadian Surveillance* with screenshots overlaid.

Circadian Surveillance is a piece about the examination of the circadian time cycle, the 24 hour day. The central concept of the piece originates from an on-going field recording project, where 24 hour continuous recordings made in specific fixed locations, starting and ending at midnight, where sped up 60 times, so that the audio is time-condensed from 24 hours to 24 minutes; an audio equivalent of time-lapse photography. What tends to happen in this level of time condensation is that the acoustic trace of large scale events comes to the fore. Weather patterns, periodic cycles of machinery or patterns of road traffic become audible and perceptible; whereas momentary events, which would have caught our attention in real time - a slamming door, the barking of a dog or a person shouting - seem to almost disappear or leave only the slightest of traces, a tiny acoustic blip in a time-condensed listening.

In *Circadian Surveillance* I use a 24-hour recording made in central Nicosia, Cyprus in August 2010. The percussion acts like a clock; the music is an encoded form of time-keeping, with the values of the digital clock mapped onto the pitches of the

¹¹⁰ Performance of *Circadian Surveillance* by The Electronic Hammer:
<https://vimeo.com/202900842>

percussion's tuned metal instruments. Alongside the percussionist are two laptop players, each with a different function: the one functions as the 'writer' and the other as the 'map maker'. The 'writer' transcribes a text of what we might be hearing, based in part on a text from writer George Perec's *The Street - Practical Exercises*.

The typing of the text has a dual effect. The audience can read the text as it is typed out, while some keys (the vowels) are encoded with a live processing effect that samples the percussion part and granulates it; grabbing moments and extending them in the time domain, in order to allow them to be manipulated in the pitch domain by the 'map maker'. The 'map maker' uses a tablet with a set of 7 prints in different scales of the map of the location of the recording (in Nicosia). He re-traces this location using a Wacom table, this action having a dual function. The audience sees the lines being drawn on the screen, while at the same time the X-Y positional data from this action is mapped onto the pitch-time data of the sound files being recalled by the typist. These sound files are 'frozen in time', and through the movement on the tablet, the 'map maker' scrubs through the audio as if it is being examined under the microscope.

The musicians have the function of encoding and decoding material from one medium to another. These actions can be seen as a charting of the terrain of each medium, using the tools of another. There is a network of interconnections between these functions, which aspire to create a balanced ecosystem of actions. Time and space are under scrutiny, while observed through the lens of a clock, a map, a field recording and a text. The clock, the tempo giver, is being manipulated by the incidental typing of the text, which is also controlling the call up of field recordings, which in turn are being manipulated by the tracing of the map, which is illustrating both the space of the field recording and following the logical dramaturgy of the text in its magnification of scale from macro to micro view.

Nerve

Nerve is a 10 minute work for orchestra, piano samples and text-film commissioned by Jurjen Hempel for Jeugd Orkest Nederland, July 2012. ¹¹¹



Figure 89 Stills from *Nerve*.

Nerve deals with the phenomenon of stage fright. A projected text describes a first person account of stage fright from the point of view of a hypothetical pianist, who is about to go on stage to perform a piano concerto. The musical material of the work is loosely based on the opening eight bars of Rachmaninoff's second piano concerto. The piano is set up in front of the orchestra as if the audience were about to hear a piano concerto, and as if the soloist is about to walk onto the stage.

The text projection opens with the lines: "I love to watch tennis because they make many mistakes, in my profession there is no room for error." The text is staggered word for word and synched to every chord or timbral change in the music, reinforced by the sampled piano. In this way the 'voice' of the narrator is established by the presence of the piano sound and at the same time underlined by the absence of the pianist. The narrative is very clearly in the first person and, similarly to that of the narrator in *Mnemnonist S* and *The Arrest*, takes us into the world of the narrator's inner thoughts, which in this case transpires to be those of a concert pianist suffering performance anxiety. The narrative itself is constructed out of texts gathered from interviews about musicians' experience of this sort of anxiety. They are woven

¹¹¹ Link to performance by the Lithuanian State Symphony Orchestra, conducted by Jurjen Hempel, at Gaida Festival, 27.10.2017: <https://youtu.be/4KkveuhKv9o>

together as a single narrative in disjointed manner, as thought processes tend to be in these circumstances. These thoughts are at times directed towards the inner self, and at other it aims to describe, justify and communicate the condition to a third person; and more relevantly, to the audience at the concert.

What I tried to achieve in this work, which was somehow different to the treatment of the first person narrative voice in previous works, was to establish a direct link to a particular instrumental voice: in this case to the absent pianist reinforced by the sight of the piano, lid up waiting to be played, and the sampled piano chords heard in the electronics, which are synched to every word that appears on screen. This way of forcing a relation between the narrative 'voice' and the 'voice' in the fabric of the music, is not far removed from the narrative underlining many works in the concerto form. In *Nerve* the absence of the pianist is compensated for by the connection between the aural presence of the piano samples and the visual synchronisation of the text to these chords. The musical voicing of the text is manipulated by the harmonic intonation of the chord, the colouring, register and expressivity of the music over a given word. An example of this could be given in the part of the piece where we read: "Stage fright sends me into a state of spin". Every word before 'spin' is accompanied by a piano sound that has an increasing use of prepared or 'broken' sounds, as if the pristine sound of the instrumental voice is inexorably crumbling. This all builds up to the delayed final word of the sentence which instead of shoring up the metaphor of the damaged voice, switches to the metaphor of vertigo implied by the word 'spin', by leaving out the last piano chord and letting the suspended winds and lack of articulation imply the dizziness of a step into the void, of an ensuing black out.

The reason why Rachmaninoff's second piano concerto serves here as a reference for *Nerve*, is that it was to be programmed as an overture to this piece in the premiere and the following tour of the orchestra. There is of course the coincidence that Rachmaninoff himself was a sufferer of both performance anxiety and writer's block: the concerto is in fact dedicated to the doctor that once helped him. This gave me the opportunity to make a direct link between the narrative voice of the imagined pianist to the actual pianist appearing in the concert hall, by having her walk slowly to the piano during the last minute of *Nerve*, sit down and start playing the Rachmaninoff immediately after the closing bars of my piece, as if it was her own inner voice that we had been hearing all along.

True Histories

True Histories is a cycle of short pieces for sampled-piano and electronics initially written for Reinaldo Laddaga's *Things that a Mutant Needs to Know: More Short And Amazing Stories*. It has been performed by Saskia Lankhoorn at Korzo, Den Haag, and Reinier van Houdt at De Link, Tilburg.

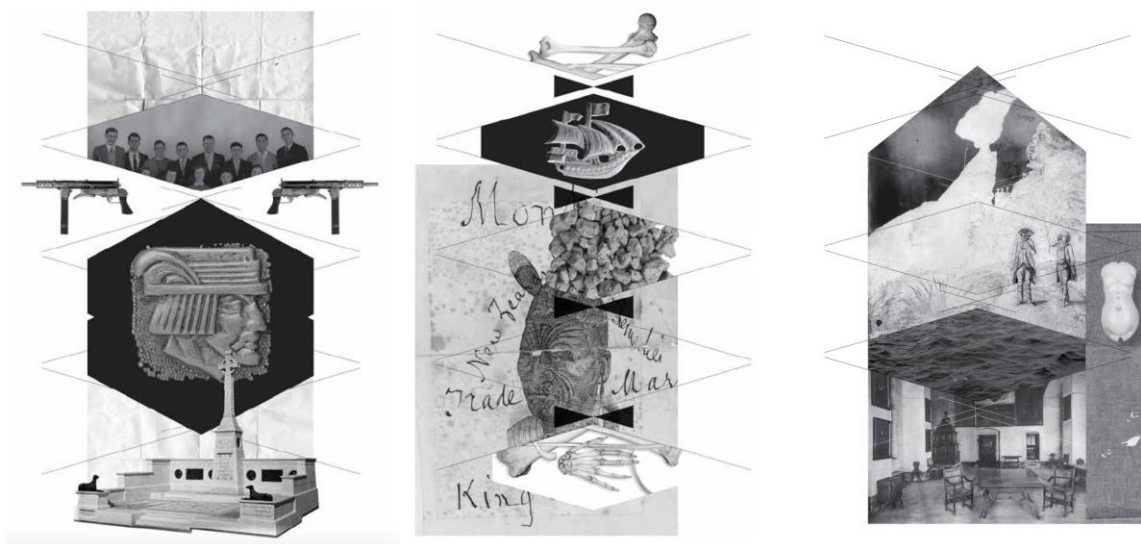


Figure 90 Illustrations by Isabelle Vigier from the book/CD release: *Things that a Mutant Needs to Know*.

True Histories explores the notion of programme music in its clear juxtaposition of story to music. The length of the narratives are short and open-ended, they are comprised of both text and music, which leaves the listener with a sense of a conundrum or a mystery that needs to be solved. One looks for an answer to the text in the music and vice-versa, yet both media are built on the idea of an untruth.

The source of many of the texts used derives from mythological accounts or travel narratives of writers from the late middle ages and the Renaissance. In fact, the origin of the project derives from a larger series of texts compiled by Argentinian writer Reinaldo Laddaga under the name *Things that a Mutant Needs to Know: More Short And Amazing Stories*.¹¹² In this project, the rule about the choice of text Laddaga made, in keeping with the concept of the original Borges/Casares anthology, was that they had to originate from books that might have been part of the private library of Jorge Luis Borges. The six texts chosen for *True Histories* are by Lucian of Samosata (from which I borrow and paraphrase the title of his best known work, *The True*

¹¹² This is an anthology (accompanied by audio tracks of 18 sound artists) conceived as a hypothetical sequel to the 1956 anthology by Jorge Luis Borges and Adolfo Bioy Casares, which was developed together with the label Unsounds and released as a book with CDs and eBook in 2013.

History), Alfred Boissier (*Selected texts relating to Assyro-Babylonian divination*), Sir John Mandéville (*Travels*), René Basset (*A thousand and one Arabic tales, accounts, and legends*) and two anonymous texts based on mythological accounts. What these texts have in common is that they require a certain suspension of disbelief (they describe 'the other' in an inadequate sense), because they are presented in a form displaced from their original culture and time:

We now crossed the river by a ford, and came to some vines of a most extraordinary kind. Out of the ground came a thick well-grown stem; but the upper part was a woman, complete from the loins upward. They were like our painters' representations of Daphne in the act of turning into a tree just as Apollo overtakes her. From the finger-tips sprang vine twigs, all loaded with grapes; the hair of their heads was tendrils, leaves, and grape-clusters. They greeted us and welcomed our approach, talking Lydian, Indian, and Greek, most of them the last. They went so far as to kiss us on the mouth; and whoever was kissed staggered like a drunken man. (Lucian of Samosata, *The True History*)

The concept adopted in the musical composition of the six pieces accompanying these texts, is one that reflects the fable-like, half-truthful nature of the narratives. The music conveys some essential illusions. The pieces are written for synthetic piano and electronics.¹¹³ Existing preludes from the piano repertoire by composers such as Bach, Chopin, Debussy, Scriabin, Alkan, and Satie are used and remapped onto an unfamiliar keyboard layout and using a peculiar individual tuning system for each piece, so that not only the tuning of the notes deviates from the original source, but the sequence of notes might be inversed or re-ordered on the keyboard. The pianist plays the original *prelude*, but what results is a fabrication, as if the original instrument has been scrambled by a cipher; the original musical vocabulary has been translated and replaced using a new mapping, that substitutes the original meaningful relation of pitches with new ones. The pianist sits down to play what would be a well-known classic, but which is conveyed through the doctored instrument as a 'fiction'.

¹¹³ In the premiere of the work at the Korzo theatre in Den Haag, a Yamaha 'silent' piano was used which sent midi data to a synthetic piano instrument (Pianoteq), giving the misconception that it was the piano itself making the sound.

8'66" (or everything that is irrelevant)

8'66" (or everything that is irrelevant) is a collaborative work created with Marko Ciciliani for Slagwerk Den Haag's Double Music project and premiered on the 11th of November 2012, during November Music in Den Bosch.

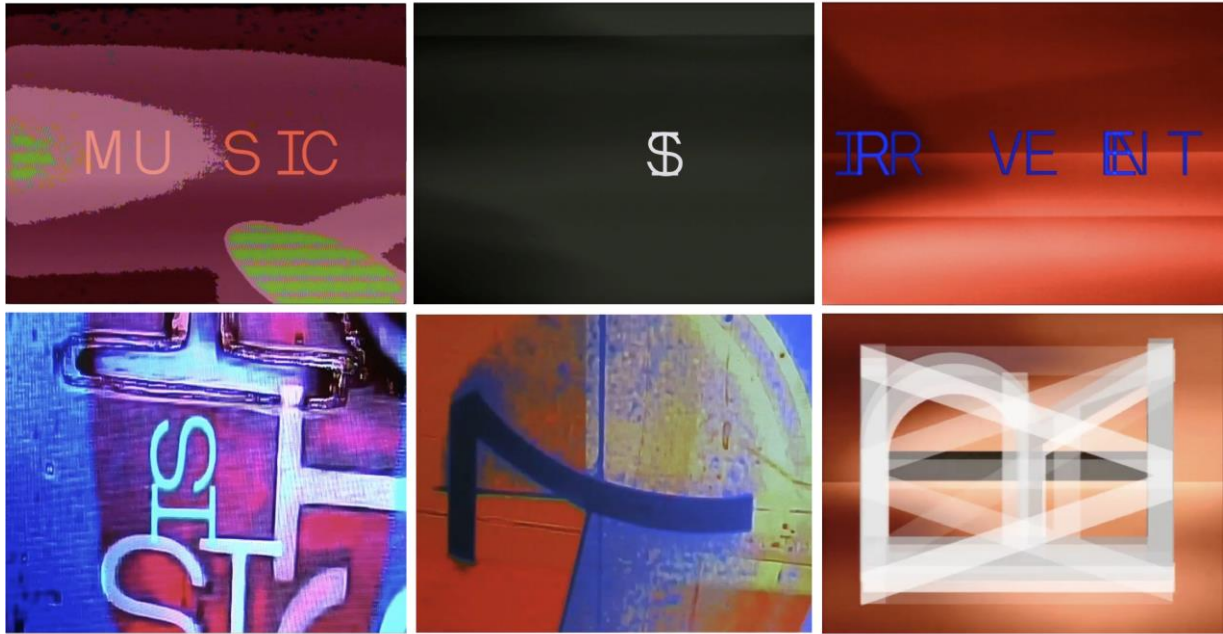


Figure 91 Stills from large projection (above) and iPads (below) from *8'66" (or everything that is irrelevant)*.

Inspired by the John Cage / Lou Harrison work of the same name, composers were asked to compose a piece with one of their closest colleagues. As the title suggest, this work has a certain relation to Cage's seminal work *4'33"*, the double of its duration. Silence, or non-performance, therefore does play a certain role in the piece. Another reference is one of Cage's final works, *One11*, the light film, which is the main source of the visual material, resampled and processed with feedback loops by Marko Ciciliani. My own input to the visual material was a Cage quote, or rather a miss-quote from one of his lectures:

Some time ago counting, patterns, tempi were dropped. Rhythm in any length of time (no-structure). Aorder. It's definitely spring - not just in the air. Take as an example of rhythm anything which seems irrelevant. (Cage 1969 : 123)

I condensed the last sentence from this quotation to: "Rhythm is everything that is irrelevant", and then gradually transformed it into various permutations throughout the piece (bold indicates the changes):

Rhythm is everything that is irrelevant
Rhythm is everything that is **irreverent**
Music is everything that is **irrelevant**
Music is **something** that is **irreverent**
Music is **anything** that is irreverent
Rhythm is anything that is irreverent
Rhythm **that is anything** is **irrelevant**

The quote is deconstructed visually as a text-film overlaid onto Ciciliani's light film, and occasionally re-enforced with gliding sine tones. This film plays for the duration of the entire piece (9 minutes and 6 seconds), over which the six players, each operating an iPad perform various actions while cueing audio-visual clips. These clips are details and variations of the main film being shown, but contain more audio material which, when played back from the various positions the musicians assume in the concert space, provide an interesting spatial experience.

At one point in the piece, the musicians also use their mouths to modulate the sound of white noise coming from the iPads, like a digital mouth harp. This piece raises interesting questions about the role of the performers, who have a somewhat passive role, but one that requires a high degree of precision in movement and body awareness. Because they become an accessory to the hand-held screen/speaker, they embody the digital material to a certain extent. This becomes an answer to Cage's silent piece, one in which there is sound, but where it is meshed onto a curious non-performance, where the visual information out-weighs that of the audio field. The image takes on the role of the music, making the latter "something that is irrelevant".

Walls Have Ears

Walls Have Ears for voice (mezzo), string quartet and text-film is a short song commissioned by Lore Lixenberg and the Brodsky String Quartet for the *Walls and Trees* project, premiered at the City of London Festival 2013. The text, which is projected rather than sung, is by Turkish Cypriot poet Mehmet Yashin.



Figure 92 Stills from *Walls Have Ears*.

This work deals with the issue of language during the time of conflict. The poet Mehmet Yashin, grew up as a Turkish Cypriot in a minority community in Cyprus of the 1960's, where speaking his mother-tongue became a complex and sensitive issue. The form of the song is based on a metaphor of the inner voice. There is a disparity developed through the piece between the text one reads in English on the video, and the incomplete Turkish phonemes that are quietly voiced by the singer.

Each note played by the quartet corresponds to a word displayed on the screen. The notes are thus enriched by an enfolding meaning that builds towards syntax completion, just as the notes add up towards melodic and harmonic coherence. The piece was originally intended to be performed with video, but because of the technical limitations of the venue of the premiere and the subsequent tour, a version was made where the quartet voice the text as they play their notes. This reinforces the sense of polyphony in the text, as each word is voiced by a different member of the quartet.

Wartime

I used to talk within myself so that no one could hear me,
and they all suspected wisdom in my silence!
Turkish was dangerous, must not be spoken,
and Greek was absolutely forbidden...
My elders who wanted to save me, were waiting,
each one trigger-ready before a machine-gun.
Anyway, everyone was then a willing soldier.
English remained right in the middle,
a slender paper-knife for cutting schoolbooks,
a tongue to be spoken at certain times
especially with the Greeks!
I was often unsure in which language to shed tears,
the life I lived wasn't foreign, but one of translation –
my mother-tongue one thing, my motherland another,
and I, again, altogether different...
Even in those days of blackouts it became obvious
I could never be the poet of any country,
because I belonged to a minority. And 'Freedom' is still
a little word uneasy in any nation's lexicon...
Then in my poems, the three languages got into a wild tangle:
Neither the Turks nor the Greeks
could hear my inner voice, nor the Others...
But I don't blame them, it was wartime.

Mehmet Yahsin (1991)

Music for Anemic Cinema

Music for Anemic Cinema is a music & video remake of Marcel Duchamp's 1926 film. Originally composed for the Ergon Ensemble and performed at the Megaron, Athens in May 2013.¹¹⁴

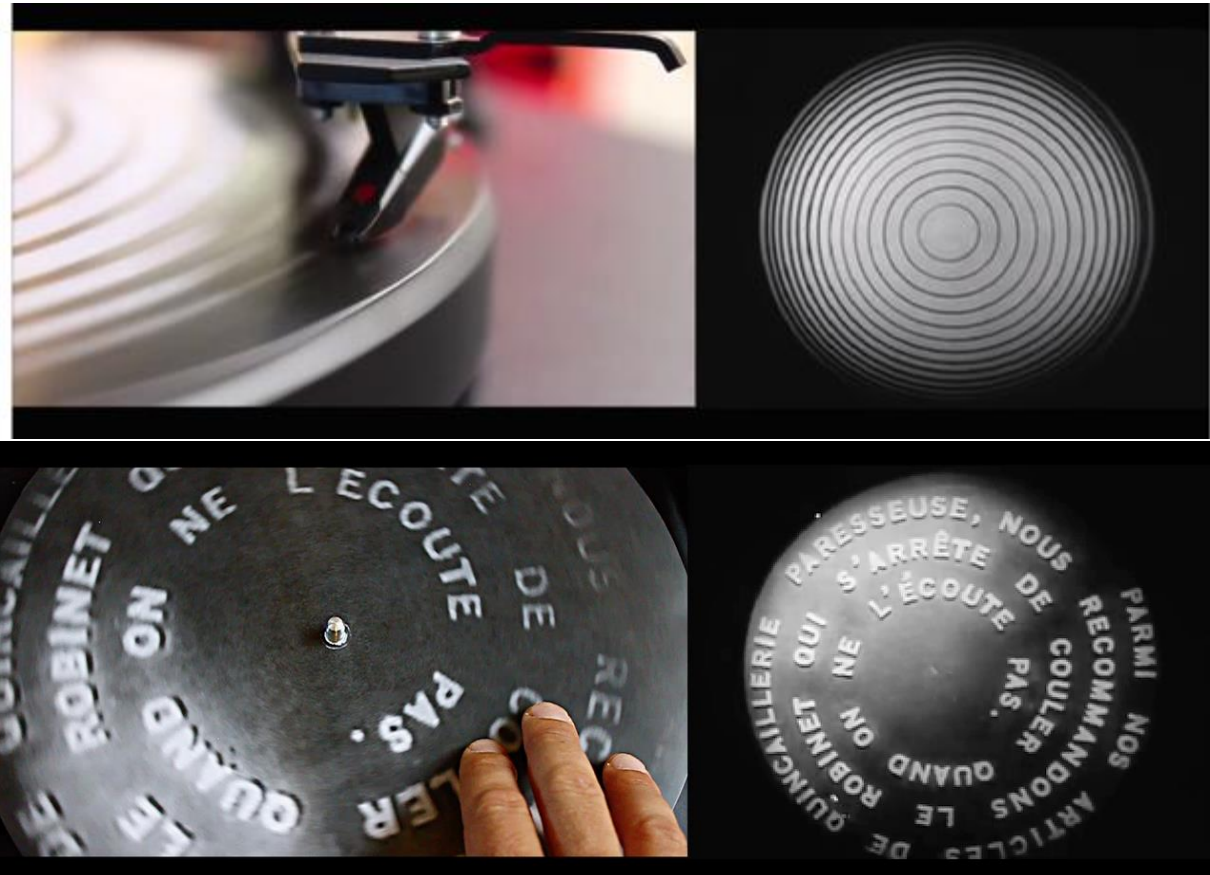


Figure 93 Stills from *Music for Anemic Cinema*.

Anemic Cinema, unlike some of the other Dadaist films of the 1920's, Satie, Clair & Picabia's *Ent'racte* and Leger & Antheil's *Ballet Mécanique*, was not intended to be shown with music (though *Ballet Mécanique* was also never shown with the intended music of George Antheil until the 1990's).

There is an almost constant flow in the film that is in itself hypnotic and creates a visual continuum which can either be highlighted or used as an element of counterpoint. This continuum is interrupted only by the alternation of the spirals with the text. The experience of the text material, even considering that the puns and language can hardly be understood, is radically different to the experience of the graphic material. One immediately feels that the juxtaposition of the new element of language creates a break, a hiatus in the hypnotic turned-in illusion of the rotor-

¹¹⁴ A realisation of *Music for Anemic Cinema*: <https://vimeo.com/105759772>

reliefs, as if one is being nudged from a reverie and given a linguistic riddle to solve a conundrum concocted by one's own inner voice.

Initially, I wished to mirror the two basic elements of the film in a dialectic form. A horizontal motion of slides and discrete pulses for the spiral section, contrasted against a vertical sequence of chords for the text sections. The horizontal aspect comprised of interference patterns between sines and an ensemble of wind and strings gliding smoothly between significant tonal centres. The piano chords were simply the words encoded into notes, spanning the entire range of the piano and heard simultaneously. The tempo or rate of chords depended purely on the number of words played at a regular pace in the given duration of the section.

Having created this initial version of the music, I wanted to go a step further in undermining the illusion of the film, and what was now the illusion of the film and music together. Knowing that the film was made by placing the rotoreliefs on a turntable, similar to the ones used for playing gramophone records, I thought it might be an interesting experiment to create the sound using the same process, to playback the graphics on a record player. I made high-quality prints of the rotorelief images from stills of the films onto photographic paper, which fitted on a modern LP player, and subsequently played them back and recorded the audio. I had no prior idea as to the result of the direct signification of the images, so I was pleasantly surprised to discover that the needle of the record player could actually sonify the differences between the darker and lighter parts of the print in a spectrum of noise. It seems that a denser perforation of ink creates a more filtered noise result, resulting in rhythms of noise born from the patterns of the rotor reliefs.

The two films, the original and the new one are played next to each other, in a manner as as synchronised as possible. The noise from the sonified rotoreliefs are mixed with the sound of the ensemble, and a translation of the text is projected in sync with its codified sonification.

MacGuffin

MacGuffin for solo electric guitar (14'), was commissioned by and dedicated to guitarist Wiek Hijmans, who premiered it in Chicago on the 11th February 2016.

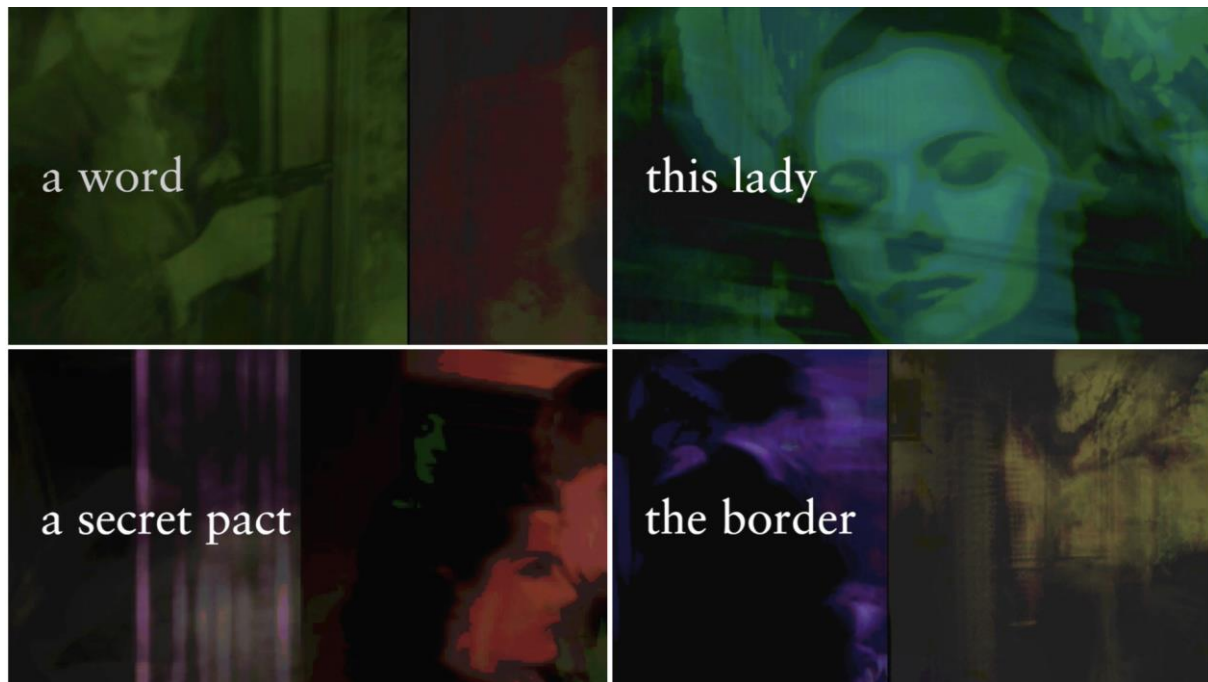


Figure 94 Stills from *MacGuffin* showing texts with processed images from *The Lady Vanishes*.

The sound world of *MacGuffin* is based on the exploration of distorted dyads and the difference tones that are thus produced. An electric guitar amplified with slight distortion can produce very clear difference tones, one or two octaves below the notated notes, especially when the dyads played are close in register. The concept of the piece is that by playing a series of the dyads (in a quarter tone tuning), a 'secret melody' is heard in these 'ghost' tones. The guitar uses scordatura: Strings I - III - V (E - G - A) are tuned down by a quarter tone (50 cents). The other three strings II - IV - VI (B - D - E) remain unchanged. This is reflected in the notation of the score. The top stave refers to the unchanged notes, the lower stave to the scordatura. Thus everything in the lower stave sounds a quarter tone lower. The piece often switches between heavy distortion, which highlights the difference tones, to non-distortion playing.

The melody that is produced by the difference tones is a quotation from Alfred Hitchcock's 1938 film *The Lady Vanishes*. In this film, Miss Froy, an unassuming governess and music teacher, has been tasked to memorise a melody performed by a folk-singer in Tirol, and bring it back to London, as the melody contains a secret message outlining the details of a treaty between two unnamed European countries.

The idea that a secret message can be contained in a melody¹¹⁵ was something in vogue in the years around the 2nd World War, and at least one other plot line using this tactic comes to mind: *Sherlock Homes - Dressed to Kill* (1946). The preposterous nature of the Hitchcock storyline, attractive as it is, does not detract from the fact that these kinds of cryptographic systems were being used for secret communication as far back as the 16th century.¹¹⁶ In the film, this was a reason for enemy forces to kidnap Miss Froy, hence the 'vanishing' of the lady. The message-in-a-melody becomes the so-called 'MacGuffin' of the plot, as it becomes the reason for the main narrative thrust of the film, i.e. finding out what happened to Miss Froy. A 'MacGuffin' is a narrative device, which triggers the main storyline of the film, yet remains unexplained or unimportant in the unfolding of the rest of plot. In this piece, just like in the film, the MacGuffin remains as the hidden motive behind what is heard and seen. The melody can only be perceived as ghost tones underneath the dissonance of the distorted dyads.

The visual part of the piece consists of 54 clips, containing text from the script of the film, an article, a noun, sometimes an adjective, all overlayed onto a processed sequence from the film. These are triggered by the musician or an assistant at various cue points in the score, usually coinciding with moments where the guitar is playing a non-distorted six note chord. The piece thus moves between moments of high intensity in sound, where no projection is seen, where the distorted melody is solely in focus, to moments of release, where words and images seem to give clues as to what the message-in-the-melody might contain.

¹¹⁵ I outline some ideas of musical cryptography in Chapter 5.2.

¹¹⁶ As detailed in *Cryptographia, oder geheime Schrifften* by Johann Balthasar Friderici, (1685).

The Lost Border Dances

The Lost Border Dances for double string quartet, electronics and text-film. Commissioned by the Holland Festival 2016 for the Kronos Quartet and Ragazze Quartet, premiered on June 23rd 2016 in the Muziekgebouw in Amsterdam.

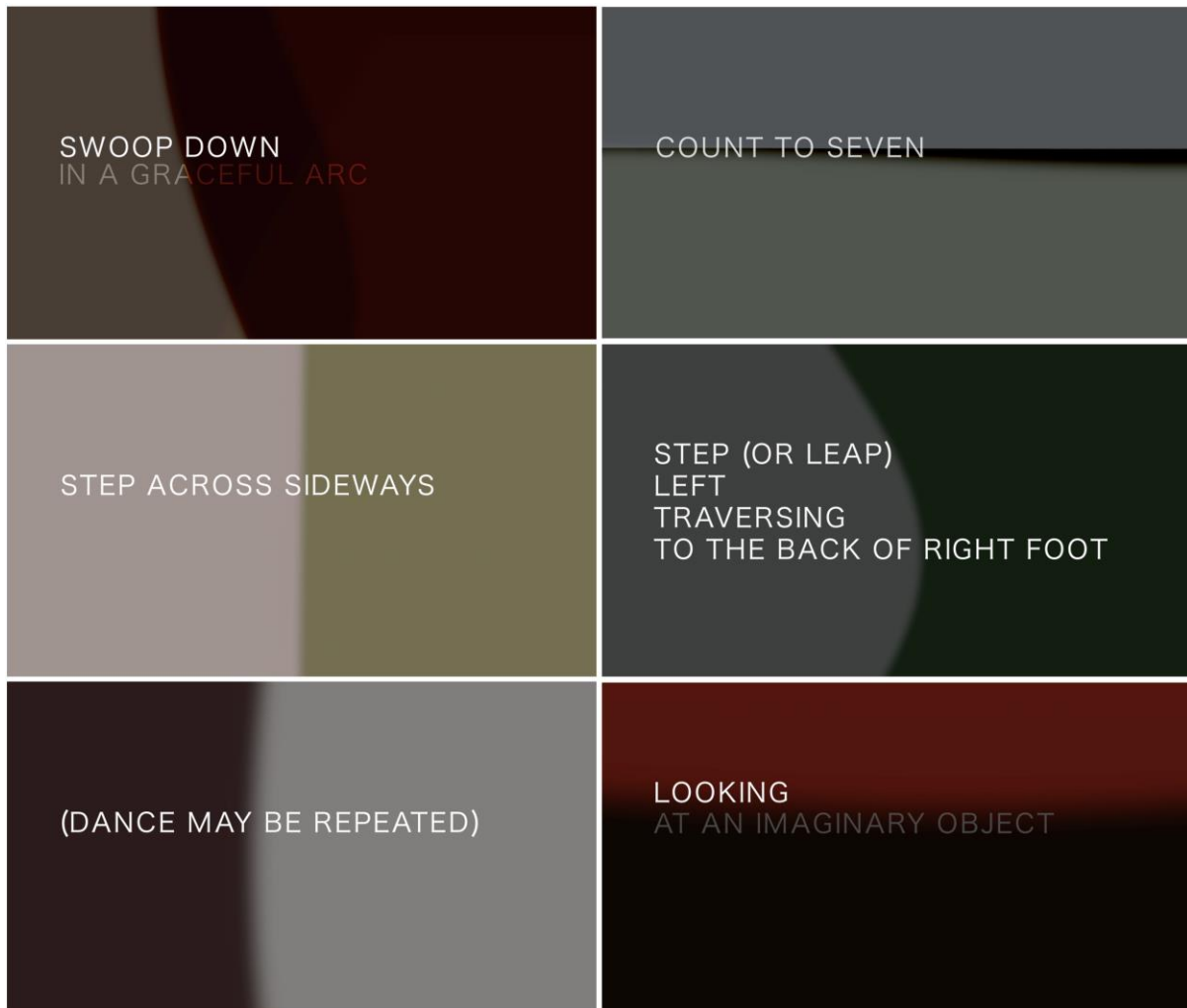


Figure 95 Stills from *The Lost Border Dances*.

The title refers to the dances performed by border guards, known as Akrites, during the height of the Byzantine Empire. The transcription of these dances and their modern equivalents found in the traditions of the Pontos, through Northern Greece to the Balkans, is used in the piece as basis for the musical material. There is no direct reference to the folk music tradition, rather the description of the movements themselves have been translated into musical gesture, while the music is re-imagined.

There are 8 dances encoded in the music, each quartet alternating in taking the lead, like a dance battle, or a dance-off. As the piece progresses, the transitions become increasingly longer, creating a shared sonic space between the re-imagined dances, an uncontested buffer zone.

An example of some of the text used:

Step in place R; pause; touch L fwd; bring L around in back of R with a circular movement;

Step L across in back of R; step R swd; step L across in front of R; pivoting on L to face RLOD, bring R around in front of L, keeping R ft close to L calf with R knee raised;

Moving to the L in RLOD, step R; step L next to R; step R; rise on ball of R, raising L knee slightly with L ft close to R calf;

Still facing to the L in RLOD, step L bwd; step R next to L; step L bwd; pivot on L to face ctr.

Each type of movement is assigned to a musical motif, formula or gesture. This is realised by the four players of each quartet at the same time. For instance, Left and Right (indicated by L and R) are always translated in either a minor or major chord respectively. Some direct movement translations are for instance: 'step' is always a step-wise melodic movement, 'circular' a circular bow movement, and 'moving' a fast run of notes. In this way the dance 'code' is sonified to produce a strange gestural music, that has the resonance of a traditional music of sorts. Furthermore, each gesture is synchronised to its description as projected in the video. The reinforcing of the movement gesture by the sound gesture, gives a clear coherence to what is being heard and described, but also underlines the fact that what is being described is never truly visualised. The visual becomes the imaginative gap, that the mind must re-imagine using words and music.

The Musicians of Dourgouti

The Musicians of Dourgouti for violin, viola, tenor saxophone, bass clarinet, piano, marimba, sound media and text-film (15') was commissioned by Ensemble Artefacts as part of their project: *Music for a New World*. It was premiered at Stegi, Athens, on the 26th of May 2017.¹¹⁷



Figure 96 Screenshots from *The Musicians of Dourgouti*.

The Musicians of Dourgouti is based on a transcription of an interview recorded by George Sachinis (UrbanDig Project) of Iosif Gevontian, a resident of the Dourgouti neighborhood of Neos Kosmos, Athens. In the interview, which begins with Gevontian singing a famous old Turkish song, *Bekledim De Gelmedim* by Yesari Asım Arsoy, he gives an account of the musical life and the many musicians, which he encountered from the 1950's onwards, while living in Dourgouti. He gives an insight into the multicultural life, and the role that music played in the everyday life of the area.

Gevontian's speaking voice, though never heard directly, is translated into the musical material of the composition. The voice is slowed down by a factor of about 1.5, and then mapped freely onto pitches and sounds of the ensemble. This is not done algorithmically, but rather manually, in order to retain some control of the

¹¹⁷ Link to performance by ARTéfacts ensemble, Onassis Cultural Centre, Athens, 26.05.17: <https://vimeo.com/243447463>

pitch and harmony. There is a polyphonic approach to the manner in which the music is arranged, each instrument contributing to the build up of harmony and sound mass. The freedom in the approach of the translation of speech to melody also enables a more unpredictable and expressive steering of the musical phrases, exploring register and timbral variations in the evolution of the narration.

The narrative is largely made up of the naming of musicians that were active in the neighbourhood at the time. Gevontian talks about the importance of the participation of the neighbourhood in musical expression through the songs, which also kept the cultural identity of the various ethnic groups alive, specifically the Armenian identity.

We had another one, Koumbonis. Lived exactly opposite my house, he played clarinet. Mr. Kostas. There was again another, Armenian, Mishak who played clarinet. There was the coal man who played ud. I remember, when it was my father's birthday they came to our house and we played. Clarinet, ud. It was crazy! crazy...

As the ensemble 'speaks', the words are projected in time with the music, in Greek with translation into English. The manner of display is kept as simple as possible. The words are grouped in phrases, and built up from one to five lines, remaining on the screen until the next utterance. The musical phrases become visually grouped on the screen in a transparent way, that enables the listener to construct a mental overview of the phrases that have just been heard. Because the text is built up from an oral interview, in which the interviewee is trying to remember names and stories from the past, there is a hesitant and unsure nature permeating the narrative. Names are repeated, order is muddled, some sentences remain unfinished.

Adding to Gevontian's own rendition of *Bekledim De Gelmedim*, we hear fragments of three other recordings of this song, by Zeki Müren, Nevin Demirdöven and Stelios Kazantzidis (alluded to in the text). These are also used in the fabric of the frozen voices heard in the background of the piece. All these renditions are time stretched throughout the piece and filtered into different ranges, which evolve over time, providing a harmonic backdrop to the melodic contours created by the ensemble. This reinforces a 'figure and ground' perspective, and like in many of my pieces, where the electronics provide some kind of background for the instrumental foreground, it is unstable and volatile, and has the potential (which is realised a few times in this piece) to overwhelm the acoustic instruments, and flip the perspective around.

Bibliography

- Abbate, C. (1991). *Unsung voices: opera and musical narrative in the nineteenth century*. Princeton, NJ: Princeton University Press.
- Adorno, T. W. (1992). *Music and Language: A Fragment*, in: *Quasi una Fantasia: Essays on Modern Music*, trans. Rodney Livingstone. London: Verso.
- Agamben, G. (1999). *The man without content*. Stanford, CA: Stanford University Press.
- Archer, M. S. (2003). *Structure, agency and the internal conversation*. Cambridge: Cambridge University Press.
- Ashley, R., & Dietrich, R. (2009). *Outside of time: Ideas about music/Ausserhalb der Zeit: Gedanken u ber Musik*. K ln: Musiktexte.
- Barsalou, L. (1992). Frames, Concepts, and Conceptual Fields. In: Lehrer, A., & Kittay, E. F. (1992). *Frames, fields, and contrasts new essays in semantic and lexical organization* (pp.21). Hillsdale (N.J.): L. Erlbaum Associates.
- Barthes, R. (1974). *S/Z* (R. Miller, Trans.). New York: Farrar, Straus & Giroux.
- Barthes, R. (1977). *Image-music-text: Roland Barthes* (S. Heath, Trans.). Glasgow: Collins.
- Barthes, R. (1989). *The Rustle of Language* (R. Howard, Trans.). New York: Farrar, Straus and Giroux.
- Bataille, G. (1979). *Oeuvres Completes: Lascaux: La Naissance de l' Art*. Paris: Gallimard.
- Bates, M. (1995). Models of natural language understanding. *Proceedings of the National Academy of Sciences*, 92(22), 9977-9982. doi:10.1073/pnas.92.22.9977
- Bauer, P. J. (2008). *Amnesia, Infantile in Language, Memory, and Cognition in Infancy and Early Childhood*. Retrieved from <https://scholarblogs.emory.edu/bauerlab/selected-publications/>
- Besson, M. & Sch n, D. (2003). Comparison Between Language and music in: Peretz, I. & Zatorre, R. (editors). *The Cognitive Neuroscience of Music*. Oxford: Oxford University Press.
- Blom, D., Bennett, D., Stevenson, I. (2016). The Composer's Program Note for Newly Written Classical Music: Content and Intentions, *Frontiers in Psychology* Nov. 2016, Vol.7, Article 1707.
- Bourriaud, N. (2007). *Postproduction: Culture as screenplay: How art reprograms the world*. New York: Lukas et Sternberg.
- Brakhage, S. (1963). *Metaphors on Vision*. S.l.: Film Culture.
- Brown, S. The "Musilanguage" Model of Music Evolution in: Wallin, N. L., Merker, B., & Brown, S. (2005). *The Origins of Music*. Cambridge: MIT Press.
- Caba as, K. M. (2014). *Off-screen cinema: Isidore Isou and the Lettrist avant-garde*. Chicago: University of Chicago Press.
- Cage, J. (1946). "The East in the West". *Modern Music*, 23, pp. 111-115.
- Cage, J. (1969). *A Year from Monday: New Lectures and Writings*. Middletown: Wesleyan University Press.

- Campbell, P. J. (1978). The origin of "Zorn's Lemma" in: *Historia Mathematica*, 5(1), 77-89. doi:10.1016/0315-0860(78)90136-2
- Chang, V. (2004). *Melos, Opsis, Lexis*. Keywords Glossary. University of Chicago. Retrieved from:
<https://lucian.uchicago.edu/blogs/mediatheory/keywords/melosopsislexis/>
- Chisholm, B. (1987). Reading Intertitles. *Journal of Popular Film and Television*, 15(3), 137-142. doi:10.1080/01956051.1987.9944095
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Cohen, A. J. (2014). Congruence-Association Model of music and multimedia in: Origin and evolution in: Tan, S., Cohen, A. J., Lipscomb, S. D., & Kendall, R. A. *The psychology of music in multimedia*. Oxford: Oxford University Press.
- Cook, N. (1998). *Analysing musical multimedia*. Oxford: Oxford University Press.
- Cooke, D. (1959). *The Language of Music*. Oxford: Oxford University Press.
- Copland, D. (2010). *Marshall McLuhan: You Know Nothing of My Work!*. London: Atlas Press.
- Cox, A. (2011). *Embodying Music, Principles of the Mimetic Hypothesis*, *Journal of the Society for Music Theory*, Volume 17, Number 2. Retrieved from:
<http://www.mtosmt.org/issues/mto.11.17.2/mto.11.17.2.cox.html>
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: HarperPerennial.
- Deleuze, G. (1990). *The Logic of Sense* (M. Lester, Trans.). New York, NY: Columbia University Press.
- Derrida, J. (1987). *The truth in painting*. Chicago; London: The University of Chicago Press.
- Derrida, J. (2004). *Dissemination* (B. Johnson, Trans.). London: Continuum.
- Dimick, H. T. (1915). *Photoplay making: A handbook devoted to the application of dramatic principles to the writing of plays for picture production*. Ridgewood: Editor Company.
- Dinhut, Charlene. (2011). Experience of Ruins and Ghosts in: *Suspended Spaces. 1. Famagusta*. Montreuil: Black Jack Editions.
- Dooley, J. (2016). *Codes, ciphers and spies: Tales of military intelligence in World War I*. New York, NY: Copernicus Books, an imprint of Springer Nature.
- Dunbar, B. (2004). Subvocal Speech Demo. Retrieved from
<https://www.nasa.gov/centers/ames/news/releases/2004/subvocal/subvocal.html>
- Eidinow, E. (2007). *Oracles, curses, and risk among the ancient Greeks*. Oxford: Oxford University Press.
- Farness, J. (1991). *Missing Socrates: Problems of Plato's writing*. Pennsylvania Park (PA): Pennsylvania State University Press.
- Feldman, J., & Narayanan, S. (2004). Embodied meaning in a neural theory of language. *Brain and Language*, 89(2), 385-392. doi:10.1016/s0093-934x(03)00355-9
- Fernyhough, C. (2016). *The voices within: The history and science of how we talk to ourselves*. London: Profile Books.

- Fleischman, S. (1990). *Tense and narrativity from Medieval performance to modern fiction*. London: Routledge.
- Fludernik, M. (2009). *An introduction to narratology*. London: Routledge.
- Frampton, H. (1983). *Circles of Confusion*. Visual Studies Workshop Press.
- Fritz, J., et al. (2013). The Neurobiology of Language, Speech, and Music in: Arbib, M. A. *Language, music, and the brain: A mysterious relationship*. Cambridge, MA: The MIT Press.
- Frye, N. (1957). *Anatomy of Criticism: Four Essays*. Princeton, NJ: Princeton University Press.
- Gallagher, S. (2014). Phenomenology and Embodied Cognition. In: Shapiro, L. A. *The Routledge handbook of embodied cognition*. London: Routledge.
- Gell, A. (1998). *Art and agency: An anthropological theory*. Oxford: Clarendon Press.
- Genette, G. (1980). *Narrative discourse: An essay in method* (J. E. Lewin, Trans.). Ithaca: Cornell University Press.
- Genette, G. (1997). *Paratexts: Thresholds of interpretation* (J. E. Lewin, Trans.). Cambridge: Cambridge Univ. Press.
- Genette, G. (2005). *Narrative discourse revisited* (J. E. Lewin, Trans.). Ithaca: Cornell University Press.
- Gibson, James, J. (1977). The Theory of Affordances in: Shaw, R., & Bransford, J. D. *Perceiving, acting and knowing: Toward an ecological psychology*. Hillsdale, NJ: Lawrence Erlbaum ass.
- Gidal, Peter. (1976). *Structural Film Anthology*. London: British Film Institute.
- Gilmore, Bob. (2011). *The Ear of the Voice of the Eye*. Yannis Kyriakides. Tilburg: teleXpress
- Girard, Rene. (1987). *Things Hidden Since the Foundation of the World* (S. Bann, M. Metteer, Trans.). Stanford, CA: Stanford University Press.
- Goebel, G. (2014). *Codes, Ciphers, & Codebreaking*. Retrieved from: <http://vc.airvectors.net/ttcode.html>
- Goethe, J. W. (2016). *Goethe's Theory of Colours* (C. L. Eeastlaee, Trans.) Fairford: The Echo library.
- Grau, O. (2003). *Virtual art: From illusion to immersion* (G. Custance, Trans.). Cambridge MA: The MIT Press.
- Greenberg, C. (1986). Towards a Newer Laocoon in: *Clement Greenberg: The collected essays and criticism Vol.1*. Chicago: The University of Chicago Press.
- Greenberg, C. (1993). Modernist Painting in: *Clement Greenberg: The collected essays and criticism Vol.4*. Chicago: The University of Chicago Press.
- Grice, H. P. (1991). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Honderich, T. (2005). *The Oxford companion to philosophy*. New York: Oxford University Press.
- Hühn, P. (editor) (2009). *Point of view, perspective, and focalization: Modeling mediation in narrative*. Berlin: Walter de Gruyter.

- Hurlburt, R. T., & Akhter, S. A. (2006). The Descriptive Experience Sampling method. *Phenomenology and the Cognitive Sciences*, 5(3-4), 271-301.
doi:10.1007/s11097-006-9024-0
- Husserl, E. (1980). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy*. La Haye: M. Nijhoff.
- Isou, I. (1947). Qu'est-ce que le lettrisme? in: *Fontaine*, no.62, October 1947 (529-550)
- Isou, I. (1947). *Introduction à une nouvelle poésie et à une nouvelle musique*. Paris: Gallimard.
- Jorgensen C. Chief Scientist for Neuro Engineering, Ames Research Center, Moffett Field, CA. Retrieved from
<http://www.techbriefs.com/component/content/article/24-ntb/features/whos-who/15620-chuck-jorgensen-chief-scientist-for-neuro-engineering-ames-research-center-moffett-field-ca?limitstart=0>
- Kahn, D. (1996). *The codebreakers the story of secret writing*. New York: Scribner.
- Kim-Cohen, S. (2009). *In the blink of an ear: Toward a non-cochlear sonic art*. New York: Bloomsbury.
- Knox, B. M. W. (1968). Silent Reading in Antiquity. *Greek, Roman, and Byzantine Studies* 9.4 (1968): 421–35.
- Kövecses, Z. (2010). *Metaphor: A practical introduction*. New York: Oxford University Press.
- Kramer, L. (2002). *Musical meaning: Toward a critical history*. Berkeley: University of California Press.
- Kramer, L. (2014). On Voice: An Introduction in: *On Voice*, W. Bernhart (ed.) Amsterdam: Rodopi.
- Kwastek, K. (2015). Immersed in Reflection? The Aesthetic Experience of Interactive Media Art. In: F. Liptay. *Immersion in the Visual Arts and Media*. S.l.: Brill.
- Kyriakides, Y. (2007). Voices in Limbo. In: *A Fearsome Heritage: Diverse Legacies of the Cold War*, J. Schofield & W. Ccroft (Editors). London: Routledge.
- LaBelle, B. (2010). *Acoustic territories: Sound culture and everyday life*. New York: Continuum.
- LaBelle, B. (2014). *Lexicon of the mouth poetics and politics of voice and the oral imaginary*. London: Bloomsbury.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Latinus, M., & Belin, P. (2011). Human voice perception. *Current Biology*, 21(4).
doi:10.1016/j.cub.2010.12.033
- Lechte, J. (1994). *Fifty key contemporary thinkers: From structuralism to postmodernity*. London: Routledge.
- Leech, Geoffrey (1974). *Semantics*. London: Penguin Books.
- Lemaître, M. (1952). *Le Film est déjà commencé?* Paris: Éditions André Bonne.
- Levinson, J. (1984). Hybrid Art Forms. *Journal of Aesthetic Education*, 18(4), 5.
doi:10.2307/3332623

- Liptay, F., & Dogramaci, B. (2016). *Immersion in the visual arts and media*. Leiden: Brill Rodopi.
- López, C. R. 2006. *What kind of affordances are musical affordances? A semiotic approach*. Paper presented at L'ascolto musicale: condotte, pratiche, grammatiche. Terzo Simposio Internazionale sulle Scienze del Linguaggio Musicale. Bologna 23-25 February 2006. Versión on-line: www.lopezcano.net
Accessed [13.07.2017]
- Lucretius (1968). *De Rerum Natura, The way things are*. Indianapolis, IN: Indiana University Press.
- Luria, A. R. (1987). *The mind of a mnemonist: a little book about a vast memory*. Cambridge: Harvard University Press.
- MacCaffery, S. & Rasula, J. (2001). *Imagining language: An anthology*. Cambridge, MA: The MIT Press.
- MacDonald, S. (1988). *A critical cinema*. Berkeley: Univ. of California Press.
- MacDonald, S. (1995). *Screen writings: Scripts and texts by independent filmmakers*. Berkeley: University of California Press.
- Manguel, A. (1996). *A history of reading*. New York: Penguin Books.
- Manovich, L. (2001). *The Language of new media*. Cambridge: MIT Press.
- Margulis, E. H. (2010). When program notes don't help: Music descriptions and enjoyment. *Psychology of Music*, 38(3), 285-302. doi:10.1177/0305735609351921
- Markie, Peter. "Rationalism vs. Empiricism." *Stanford Encyclopedia of Philosophy*, Stanford University, 6 July 2017, plato.stanford.edu/entries/rationalism-empiricism/.
- Marshall, S. K., & Cohen, A. J. (1988). Effects of Musical Soundtracks on Attitudes toward Animated Geometric Figures. *Music Perception: An Interdisciplinary Journal*, 6(1), 95-112. doi:10.2307/40285417
- Martin, K. (1975). *Marcel Duchamp's Anemic Cinema*, Studio International 189, no.973.
- McCarthy-Jones, S. (2012). *Hearing voices: The histories, causes, and meanings of auditory verbal hallucinations*. Cambridge: Cambridge Univ. Press.
- McLuhan, M. (1962). *The Gutenberg galaxy: The making of typographic man*. Toronto: University of Toronto Press.
- McLuhan, M. (2013). *Understanding media: The extensions of man*. Cambridge (Mass.): The MIT Press.
- Meelberg, V. (2006). *New sounds, new stories narrativity in contemporary music*. Leiden: Leiden University Press.
- Mendelssohn, F. (1946). *Felix Mendelssohn: Letters*. (G. Selden-Goth, Trans.). London: Paul Elek.
- Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90(2), 227-234. doi:10.1037/h0031564
- Mitchell, W. J. T & Hansen, M. B. N (2010). *Critical terms for media studies*. Chicago: The University of Chicago Press.

- Mithen, S. J. (2006). *The singing Neanderthals: The origins of music, language, mind, and body*. Cambridge, MA: Harvard University Press.
- Monaco, J. (2004). *The new wave: Truffaut, Godard, Chabrol, Rohmer, Rivette*. New York: Harbor Electronic Pub.
- Morris, Charles (1938). *Foundations of the Theory of Signs: Volume 1. Number 2*, Chicago: The University of Chicago Press.
- Nattiez, J. (1990). *Music and discourse: Toward a semiology of music* (C. Abbate, Trans.). Princeton: Princeton University Press.
- Nietzsche, F. (2002). *Friedrich Nietzsche: Beyond good and evil: Prelude to a philosophy of the future* (J. Norman, Trans.). U.K: Cambridge University Press.
- Nikolić, D. (2009). *Is synaesthesia actually ideaesthesia? An inquiry into the nature of the phenomenon*. Retrieved July 7, 2016, from <http://www.danko-nikolic.com/wp-content/uploads/2011/09/Synesthesia2009-Nikolic-Ideaesthesia.pdf>
- Nussbaum, C. O. (2007). *The musical representation: Meaning, ontology, and emotion*. Cambridge, MA: MIT.
- Oberhelman, S. M. (2008). *Dreambooks in Byzantium: Six "Oneirocritica" in translation, with commentary and introduction*. Aldershot, England: Ashgate.
- O'Barr, W. M. (2005). "Subliminal" Advertising. *Advertising & Society Review*, 6(4). doi:10.1353/asr.2006.0014
- Oers, R. V. (2014). *Deserving citizenship citizenship tests in Germany, the Netherlands and the United Kingdom*. Leiden: M. Nijhoff.
- Packard, V. (1957). *The Hidden Persuaders*. New York: David McKay Co., Inc.
- Parnin, C. (2011). *Subvocalization - Toward Hearing the Inner Thoughts of Developers*. 2011 IEEE 19th International Conference on Program Comprehension. doi:10.1109/icpc.2011.49
- Peirce, C. S. (1994). *The Collected Papers of Charles Sanders Peirce. Electronic edition. Volume 4: The Simplest Mathematics* (C. Hartshorne & P. Weiss, Eds.). Charlottesville, VA: IntelLex Corporation.
- Perec, G. (1973). *La boutique obscure*. Paris: Éditions Denoël.
- Perec, G. (1998). *Species of spaces and other pieces*: (J. Sturrock, Trans.). London: Penguin.
- Pier, J. (2010). *Metalepsis: The living handbook of narratology*. Hamburg University Press. Retrieved September 5, 2012, <http://www.lhn.uni-hamburg.de/>
- Plato (2009). *Complete works* (J. M. Cooper, Ed.). Indianapolis: Hackett.
- Potalsky, M. (2006). *Mimesis*. London: Routledge.
- Poyatos, F. (1993). *Paralanguage: A linguistic and interdisciplinary approach to interactive speech and sound*. Amsterdam: J. Benjamins.
- Poyatos, F. (2002). *Nonverbal communication across disciplines: Paralanguage, kinesics, silence, personal and environmental interaction*. Amsterdam: John Benjamins Publishing Company.
- Raaijmakers, D. (1978). *De Kunst van het Machinelezen*, in: J. Brouwer, A. Mulder, D. Raaijmakers, (2008). *Dick Raaymakers: A monograph*. Rotterdam: V2_Instituut voor de instabiele media.

- Raaijmakers, D. (1998) Guidebook to 'The Complete Tape Music of Dick Raaijmakers', Amsterdam: Donemus.
- Ramachandran, V.S. & Hubbard, E.M. (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies* 8(12): 3-34.
- Ricoeur, P. (1983). *Time and Narrative Vol.1*, (K. MacLaughlin & D. Pellauer, Trans.). Chicago: University of Chicago Press.
- Rogers, J., & Pullum, G. K. (2011). Aural Pattern Recognition Experiments and the Subregular Hierarchy. *Journal of Logic, Language and Information*, 20(3), 329-342. doi:10.1007/s10849-011-9140-2
- Salzman, E., & Desi, T. (2008). *The new music theater: Seeing the voice, hearing the body*. New York: Oxford Univ. Press.
- Sams, E. 'Cryptography, musical' in: Sadie, Stanley (ed.), *The New Grove dictionary of music and musicians*, Macmillan, 1980, (6th ed. of the Grove dictionary), vol.5, p. 80.
- Scheffer, B. (2013). *Typemotion: Type as image in motion*. Ostfildern: Hatje Cantz.
- Shaw-Miller, S. (2014). Opsi Melos Lexis in: J. H. Rubin. *Rival sisters, art and music at the birth of modernism, 1815 - 1915*. Farnham: Ashgate.
- Simonney, D. (2012). Lalangue en questions. *Essaim*, 29(2), 7. doi:10.3917/ess.029.0007
- Small, C. (1998). *Musicking: The meanings of performing and listening*. Middletown: Wesleyan University Press.
- Smalley, D. (1997). Spectromorphology: Explaining sound-shapes. *Organised Sound*, 2(2), 107-126. Cambridge University Press. doi:10.1017/s1355771897009059
- Smith, R. G. (2010). *The Baudrillard dictionary*. Edinburgh: Edinburgh University Press.
- Spielmann, Y. (2008). *Video: The reflexive medium*. Cambridge: The MIT Press.
- Spitzer, M. (2004). *Metaphor and musical thought*. Chicago, London: The University of Chicago.
- Sternberg, E. J. (2015). *Neurologic: The brain's hidden rationale behind our irrational behavior*. New York: Pantheon Books.
- Sterne, J. (2008). Enemy Voice. *Social Text*, 26(3 96), 79-100. Duke University Press. doi:10.1215/01642472-2008-005
- Stoller, P. (1992). *The cinematic griot: The ethnography of Jean Rouch*. Chicago: University of Chicago Press.
- Sudre, F. (1866). *Langue musicale universelle*. Paris: Chez la veuve de l'auteur ... et Chez G. Flaxland.
- Testa, B. (2002). *Early Cinema and the Avant-Garde*. Retrieved from http://www.sixpackfilm.com/archive/veranstaltung/festivals/earlycinema/symposion/symposion_gunning.html
- Ugresic, D. (2011). *Karaoke culture* (D. Williams, Trans.). Rochester, NY: Open Letter.
- Vervaeck, B., & Herman, L. (2001). *Handbook of narrative analysis*. Lincoln: University of Nebraska Press.

- Viola, B. (2002). *Going forth by Day*. Catalogue of Exhibition. The Solomon R. Guggenheim Foundation.
- Visch, V. T., Tan, E. S., & Molenaar, D. (2010). The emotional and cognitive effect of immersion in film viewing. *Cognition & Emotion*, 24(8), 1439-1445.
doi:10.1080/02699930903498186
- Vitu, F. (2011). On the role of visual and oculomotor processes in reading, in: S. P. Liversedge, I. D. Gilchrist & S. Everling (eds.), *The Oxford handbook of eye movements*. Oxford: Oxford University Press.
- Weidle, R. (2009). Organizing the Perspectives: Focalization and the Superordinate Narrative System in Drama and Theater in: P. Hühn (editor) *Point of view, perspective, and focalization: Modeling mediation in narrative*. Berlin: Walter de Gruyter.
- Whalley, G. (1997). *Aristotle's Poetics*. Montreal: McGill-Queen's University Press.
- Wilkins, J. (1970). *The mathematical and philosophical works of the Right Rev. John Wilkins*. London: Frank Cass.
- Windhausen, F. (2004). *Words into Film: Toward a Genealogical Understanding of Hollis Frampton's Theory and Practice*. October, 109, 76-95. Boston: MIT Press.
doi:10.1162/0162287041886494
- Woolhouse, R. S. (1994). *Gottfried Wilhelm Leibniz. critical assessments*. London: Routledge.

Audiovisual Media Citations

- Ablinger, P. (2008) *A Letter from Schoenberg* [Music Multimedia]. Retrieved from <http://www.youtube.com/watch?v=BBsXovEWBGo>
- Allen, W. (Director). (1985). *The Purple Rose of Cairo* [Film]. Orion/Jack Rollins-Charles Joffe.
- Ashley, Robert. (2006). *Perfect lives: An opera for television* [Video Opera]. Lovely Music.
- Björk. (2010, August 10). *All is Full of Love* [Music Video]. Retrieved from <http://www.youtube.com/watch?v=AjI2J2SQ528>
- Blonk, J., & Levin, G. (2017, August 24). *Ursonography*, Jaap Blonk & Golan Levin, 2007 [Audiovisual Performance]. Retrieved from <https://vimeo.com/2687898>
- Brakhage, S. (1988). *I Dreaming in: By Brakhage, an anthology: Volume One* (2010). [DVD] Irvington, NY: Criterion Collection.
- Buñuel, L., & Dali, S. (1928). *Un chien Andalou* [Film]. France.
- Cage, J. (1961). *Variations II for any number of players and any sound producing means* ; NY: Peters.
- Clair, R. (1924). Erik Satie/René Clair: *Entr'Acte* [Film]. Retrieved from <http://www.youtube.com/watch?v=mpr8mXcX80Q>
- Coppola, F. F. (1979). *Apocalypse now* [Film]. Universal pictures.
- Cunningham, M. (1973). *Walkaround Time* - Merce Cunningham Dance Company [Dance Performance] Retrieved from <http://www.youtube.com/watch?v=dVpF7qZPavU>
- Curtis, A. (2009). *It Felt Like a Kiss* [Film]. Retrieved from <https://vimeo.com/22589118>
- Debord, G. (1973). *La Société du Spectacle* (1973) [Film]. Retrieved from <http://www.youtube.com/watch?v=IaHMgToJlJA>
- Duchamp, M. (1926). *Anemic Cinema* [Film]. Retrieved from: http://ubu.com/film/duchamp_anemic.html
- Dylan, B. (1965). *Bob Dylan - Subterranean Homesick Blues*. [Music Video]. Retrieved from <http://www.youtube.com/watch?v=MGxjIBEZvx0>
- Ferrari, L. (1995). *Luc Ferrari Presque rien* [CD]. INA/GRM.
- Frampton, H. (1979) *Gloria* in: *A Hollis Frampton Odyssey* [Film] USA: Criterion Collection.
- Frampton, H. (1972) *Poetic Justice* in: *A Hollis Frampton Odyssey* [Film] USA: Criterion Collection.
- Frampton, H. (1968) *Surface Tension* in: *A Hollis Frampton Odyssey* [Film] USA: Criterion Collection.
- Frampton, H. (1970) *Zorn's Lemma* in: *A Hollis Frampton Odyssey* [Film] USA: Criterion Collection.
- Godard, J. L. (1987). *Armide* in: *Aria* [Film]. Great Britain: Virgin Vision.

- Herzog, W. (1992). *Lessons of Darkness* [Film]. Germany: Werner Herzog Filmproduktion.
- Isou, I. (1951) *Traité de bave et d'éternité. Treatise on Venom and Eternity* [Film]. France. Retrieved from <https://vimeo.com/ondemand/isidore>
- Léger, F. (1924) *Ballet Mécanique* [Film]. France: Synchro-Cine.
- Lemaître, M. (1951). *Le Film est déjà commencé? (Has the film already started?)* [Film]. France. Retrieved from http://ubu.com/film/lemaitre_film.html
- Logan, J. (1955). *Picnic* [Film]. United States: Columbia Pictures.
- Manuva, R. (2006). *Roots Manuva, Too Cold*. [Music Video]. Retrieved from http://www.youtube.com/watch?v=lAeCx5_4L3o
- Marker, C. (1983). *Sans Soleil* [Film]. France: Argos Film.
- Michael, G. (1990). *Praying for Time*. [Music Video]. Retrieved from <http://www.youtube.com/watch?v=goroyZbVdlo>
- Oswald, J. (1998). *Homonymy* performed by Eve Egoyan. Retrieved from <https://vimeo.com/167620526>
- Perry, L.S. & Warrior Queen (2011). *Two Edged Sword* [EP] on: *Profit, Have-A-Break-Recordings*.
- Prince (1987). *Sign o' the Times* [Music Video]. Retrieved from <https://www.youtube.com/watch?v=8EdxM72EZ94>
- Pulp (1998) *This Is Hardcore* [Music Video]. Retrieved from <http://www.youtube.com/watch?v=JXbLyi5wgeg>
- Raaijmakers, D. (1966) *Ballade Erbkönig* on: *The Complete Tape Music of Dick Raaijmakers*, NEAR/Donemus, Amsterdam 1998.
- Reich, S., & Corot, B. (1995). *The Cave*. [CD] US: Elektra / Nonesuch
- Serra, R. (1973). *Television Delivers People*. Retrieved from <http://www.youtube.com/watch?v=nbvzbj4Nhtk>
- Sherwin, G. (1972) on: *Guy Sherwin: Optical Sound Films 1971-2007* (2008) [DVD] UK: LUX
- Snow, M. (1982). *So Is This*. [Film] Retrieved from <http://www.youtube.com/watch?v=8i6H1KDj9Ic>
- Snow, M. (1967). *Wavelength* [Film]. Canadian Filmmakers Distribution Centre.
- Snow, M. (1975). *Musics for Piano, Whistling, Microphone and Tape Recorder* [Vinyl] US: Chatham Square Productions
- Viola, B. (1976). *He Weeps for You* [Installation]. US: MoMA.
- Walshe, J. (2016). *Everything is Important* [Music Multimedia]
- Wilson, R., & Waits, T. (1990). *The Black Rider: The Casting of the Magic Bullet* [Music Theatre] Retrieved from <http://www.youtube.com/watch?v=lbQkzAbCjio>

Links to Online Media of Music-Text-Film

(In order of appearance in thesis)

Introduction

Words and Song Without Words:

Recording by Francesco Dillon: <https://vimeo.com/54731855>

Performance by Karolina Öhman:

<https://www.youtube.com/watch?v=D4dT7WPfoOs>

Performance by Larissa Groeneveld:

<https://www.youtube.com/watch?v=F7J94yFpFaQ>

Chapter 3

Subliminal: The Lucretian Picnic

Performed by ASKO|Schoenberg: <https://www.youtube.com/watch?v=ZJDC-gBb9rQ>

Chapter 5

Dreams of the Blind

Performed by Ensemble MAE: <https://vimeo.com/226624063>

Mnemonist S

Performed by ASKO|Schoenberg: <https://vimeo.com/13766483>

Memoryscape

Performed by MusikFabrik: <https://vimeo.com/226623184>

Chapter 6

Machine Read

<http://tijdschriftterras.nl/a-reflection-ideas-dick-raaijmakers/>

Chapter 7

Wordless

Complete Suite: <https://youtu.be/UHg6-hEn8Ms>

Varosha

Video with recording from Resorts and Ruins CD: <https://vimeo.com/192369559>

Der Komponist

Live audio recording from premiere performance by Philharmonie Zuidnederland, conducted by Bas Wiegers: <https://soundcloud.com/yannisky/der-komponist-for-orchestra-an-electronics>

Chapter 8

Karaoke Etudes

Video scores only:

<https://vimeo.com/191127009>

Performed by Thin Edge New Music Collective and Ensemble Paramirabo:

<https://www.youtube.com/watch?v=HEPaAtrnZag>

Performed by Seattle Chamber Players:

<https://www.youtube.com/watch?v=F0T6DVhyedw>

Trench Code

Performed by MAZE: <https://vimeo.com/226869061>

Oneiricon

Performed by MAZE: <https://www.youtube.com/watch?v=GroHLc9QXTk>

iOS app at App Store: <https://itunes.apple.com/us/app/oneiricon/id1293741939?mt=8>

Appendix

Scam Spam

Performance by Takao Hyakutome: <https://vimeo.com/39011042>

QFO

<http://earreader.nl/wp-content/uploads/2010/11/kyriakides.html>

Adobe 'Flash' is needed, so this might not work on Apple smartphones and tablets.

RE: Mad Masters

Performance by Barbara Lüneburg: <https://vimeo.com/225816862>

The Arrest

Performance by Ensemble MAE: <https://vimeo.com/14960327>

Circadian Surveillance

Performed by The Electronic Hammer: <https://vimeo.com/202900842>

Nerve

Performance by the Lithuanian State Symphony Orchestra:

<https://youtu.be/4KkveuhKv9o>

Music for Anemic Cinema

Computer realisation: <https://vimeo.com/105759772>

The Musicians of Dourgouti

Performance by ARTéfacts ensemble: <https://vimeo.com/243447463>

Summary

Over the past years, I have developed a form of composition, music with on-screen text, which I define as 'music-text-film', in which I explore the dynamics between sound, words and visuals. In this thesis, I explore the ideas around these pieces, and attempt to explain how meaning is constructed in the interplay between the different layers of media.

The issues that initially arose out of the research, was related to the question of 'voice': Who is narrating? And where is the voice located? These questions became more pertinent after I noticed a strange phenomenon occurring during performances of these works: that when we read text synchronised to music, we become very aware of an inner voice silently reading along. This effect of hearing one's own voice in the music, becomes an added phenomenon that I had not initially predicted. It was a discovery that had many consequences for the ways in which I subsequently approached composition and ideas about listening. In my music-text-films a state of limbo is created between the narrative voice of the text and the implied voice of the music, due to the absence of a conventional focal point to pin it on - an actor or a singer. In the thesis I suggest that because of this vacancy and the way the projected word takes the place of the sung or spoken voice, the inner voice of the audience becomes activated. This then becomes a vital immersive dimension in the performance, as the inner voice of the audience finds its place within the space of the composition.

I have chosen to call this thesis a 'poetics.' There exist two thoughts behind this: first, I wanted to place the focus primarily on the form, meaning and implications of music-text-film. Rather than deal with theoretical, aesthetic or other philosophical questions discretely, I wanted to approach them as they arise out of, or through commenting on, this particular artistic practice. Secondly, the themes that I have chosen to structure the theory around come directly from the Aristotle's poetics: The keywords for the first three chapters are based around terms strongly associated with the *Poetics*: 'mimesis', 'diegesis', and the trichotomy of media: 'melos, lexis, opsis'.

The first two chapters elaborate on Plato's binary distinctions of art: mimesis and diegesis (imitation and narration). In Chapter 1, I begin with the basic definitions, centred on the idea that art is by nature imitative, and develop the idea of mimesis, not in terms of how art mirrors the world, but how the spectator mirrors the artwork. The question of what extent the spectator is implicated in the artwork, the relation of immersion versus critical distance involved in music-text-film, is defined as an intermediate state of 'cognitive immersion', not fully immersed but engaged on a certain cognitive level, where the spectator is projected into the artwork.

Out of this I define three forms of inner vocality that, I argue, are activated by music-text-film: 'silent reading' as in the reading of the text; 'silent singing' as in the tracing of melodic contours with the inner voice; and 'silent discourse' the hidden dialogue of thought that occasionally surfaces during overt self-reflexive moments in the works, or when the half-completed syntax of words triggers a myriad of possible answers.

Chapter 2 develops Aristotle's conception of diegesis, the art of narration, elaborated into questions about how narrative operates in a musical context and specifically a multimedia form such as music-text-film. The necessary conditions for narration are discussed, specifically highlighting the relation between narration and voice: the focus of the narrative that is then given over as perspective to the spectator. The idea that for narration to exist there have to be two distinct ontological levels is one of the conclusions that drawn from this. One of the concluding observations is that ontological levels are also demarcated by differences of media.

Multimedia art is the principal focus of Chapter 3. Aristotle's trichotomy of media, melos, lexis and opsis, forms the basis of a discussion of the history, hierarchy and opacity of media, as well as notions of what in fact constitutes a medium. I go on to propose two different models of analysing multimedia: the first based on the correlation of six different aspects of media, and the second model, by looking at how hierarchies of media are manifested in the artwork.

In Chapter 4, I trace a history of text-film organised not in chronological order but in terms of metaphoric relations between the two dominant media. This, again, demonstrates the way in which perspective is dependant on the particular art practice these works emerge from, as well the cultural context. These include pieces that have had a significant influence on my own work: Marcel Duchamp's *Anemic Cinema*, Hollis Frampton's *Zorn's Lemma*, Michael Snow's *So Is This*, Dick Raaijmakers' *Ballade Erbkönig*, Robert Ashley's *Perfect Lives* and Isidore Isou's *Traité de bave et d'éternité*.

The second part of the thesis is devoted to the discussion of my own music-text-film pieces. In recent years I have written about 30 works which use projected text in some form, in a music or sound art context. I have charted this progress in the four chapters that make up this part of the thesis: 'Internal Monologues', deals with three ensemble works that highlight first-person narratives derived from conscious or semi-conscious discourse, 'Unanswered Questions', deals with a video and two installations that explore question and answer structures across media, 'Voiceprints' concerns work where the material is based on the manipulation of spoken voices,

and 'Interactive Scores', looks at my recent work dealing with algorithmic app-scores that extend the idea of music-text-film towards interactive musical notation.

I have tried to highlight the different ways the idea of 'voice' can function in some of my music-text-film, as a way of articulating the dynamics of multimedia work in general. The shifting perception of what 'voice' can be, is a compelling aspect of this form of music-text-film, as it fluctuates from a purely narrative form, to a voice as sonic expression, to the audience becoming aware of their own inner voices as they read the projected text in resonance with the music. The question of what constitutes a voice is ultimately at the heart of this research, as the voice moves from being a carrier of meaning, of narrative, to determining the way our attention shifts between the many layers of different media.

Samenvatting

In de afgelopen jaren heb ik een vorm van compositie ontwikkeld 'muziek met tekst op scherm' die ik definieer als 'music-text-film', en waarin ik de dynamiek tussen geluid, woorden en beelden onderzoek. In dit proefschrift onderzoek ik de ideeën rondom deze stukken en probeer ik uit te leggen hoe betekenis wordt geconstrueerd in het samenspel tussen de verschillende lagen van media.

De kwesties die in eerste instantie voortkwamen uit het onderzoek, hadden te maken met de kwestie van 'stem': wie vertelt er? En waar bevindt de stem zich? Deze vragen werden relevanter nadat er een vreemd fenomeen optrad tijdens uitvoeringen van deze werken: wanneer we tekst lezen die met muziek is gesynchroniseerd, blijken we ons heel goed bewust te worden van een innerlijke stem die stil leest. Dit effect van het horen van een eigen stem in de muziek, werd een toegevoegd fenomeen dat ik aanvankelijk niet had voorspeld. Het was een ontdekking die veel gevolgen had voor de manier waarop ik vervolgens componeren en ideeën over luisteren benaderde. In mijn 'music-text-film' ontstaat er een limbo-situatie tussen de verhalende stem van de tekst en de impliciete stem van de muziek, vanwege de afwezigheid van een conventioneel brandpunt om deze vast te pinnen - een acteur of een zanger. In het proefschrift stel ik voor dat vanwege deze vacature en de manier waarop het geprojecteerde woord de plaats inneemt van de gezongen of gesproken stem, de innerlijke stem van het publiek wordt geactiveerd. Dit wordt dan een vitale meeslepende dimensie in de uitvoering, aangezien de innerlijke stem van het publiek zijn plaats vindt binnen de ruimte van de compositie.

Ik heb ervoor gekozen om dit proefschrift een 'poëtica' te noemen. Daar zitten twee gedachten achter: ten eerste wilde ik de aandacht vooral richten op de vorm, betekenis en implicaties van muziek-tekst-film. In plaats van discreet om te gaan met theoretische, esthetische of andere filosofische vragen, wilde ik ze benaderen als voortkomend uit, of commentaar gevend op deze specifieke artistieke praktijk. Ten tweede zijn de thema's die ik heb gekozen om de theorie te structureren, rechtstreeks afkomstig van de poëtica van Aristoteles: de sleutelwoorden voor de eerste drie hoofdstukken zijn gebaseerd op termen die sterk verbonden zijn met de poëtica: 'mimesis', 'diegesis' en de trichotomie van media: 'melos, lexis, opsis'.

De eerste twee hoofdstukken gaan in op Plato's binaire kunstverscheidenheid: mimesis en diegesis (imitatie en vertelling). In hoofdstuk 1 begin ik met de basisdefinities, gecentreerd rond het idee dat kunst van nature imitatief is, en het idee van het ontwikkelen van mimesis, niet in termen van hoe kunst de wereld weerspiegelt, maar veeleer van hoe de toeschouwer het kunstwerk weerspiegelt. De vraag in hoeverre de toeschouwer betrokken is bij het kunstwerk, de relatie tussen immersie en het op kritische afstand betrokken zijn bij muziek-tekst-film, wordt gedefinieerd als een tussentijdse staat van 'cognitieve onderdompeling', niet volledig ondergedompeld maar betrokken op een bepaald cognitief niveau, waar de toeschouwer in het kunstwerk wordt geprojecteerd.

Op basis hiervan definieer ik drie vormen van innerlijke vocaliteit die, naar ik betoog, worden geactiveerd door muziek-tekst-film: 'silent reading' zoals bij het lezen van een tekst; 'silent singing' zoals bij het opsporen van melodische contouren met de innerlijke stem; en 'silent discourse' de verborgen gedachtendialoog die af en toe opduikt tijdens openlijke

zelfreflexieve momenten in de composities, of wanneer de halfvoltooide syntaxis van woorden een groot aantal mogelijke antwoorden triggert.

Hoofdstuk 2 ontwikkelt Aristoteles' conceptie van diegesis, de kunst van het vertellen, uitgewerkt in vragen over hoe de verhaalstructuur werkt in een muzikale context en specifiek in een multimedia-vorm zoals 'music-text-film'. De noodzakelijke voorwaarden voor 'het verhaal' worden besproken, waarbij met name aandacht wordt besteed aan de relatie tussen vertelling en stem: de focus van het verhaal dat vervolgens als perspectief aan de toeschouwer wordt gegeven. Het idee dat er voor het bestaan van een vertelling twee verschillende ontologische niveaus moeten zijn, is een van de conclusies die hieruit is getrokken. Een van de concluderende observaties is dat ontologische niveaus ook worden afgebakend door verschillen in media.

Multimediakunst is de belangrijkste focus van hoofdstuk 3. Aristoteles' trichotomie van media – 'melos, lexis en opsis' – vormt de basis van de bespreking van de geschiedenis, hiërarchie en transparantie van media, evenals van de noties van wat in feite een medium vormt. Ik ga verder met het voorstellen van twee verschillende modellen voor het analyseren van multimedia: de eerste gebaseerd op de correlatie van zes verschillende aspecten van media, en het tweede model, door te kijken naar hoe hiërarchieën van media zich in het kunstwerk kunnen manifesteren.

In hoofdstuk 4 presenteer ik een geschiedenis van tekst-film die niet in chronologische volgorde is georganiseerd, maar in termen van metaforische relaties tussen de twee dominante media. Dit demonstreert opnieuw de manier waarop perspectief afhankelijk is van de specifieke kunstpraktijk waar deze werken uit voortkomen, evenals de culturele context. Dit zijn producties die een belangrijke invloed op mijn eigen werk hebben gehad: *Anemic Cinema* van Marcel Duchamp, *Zorn's Lemma* van Hollis Frampton, Michael Snow's *So Is This*, *Ballade Erlkönig* van Dick Raaijmakers, Robert Ashley's *Perfect Lives* and Isidore Isou's *Traité de bave et d'éternité*.

Het tweede deel van het proefschrift is gewijd aan de bespreking van mijn eigen stukken 'music-text-film' composities. In de afgelopen jaren heb ik ongeveer 30 werken geschreven die in een of andere vorm geprojecteerde tekst gebruiken, in een context van muziek of geluidskunst. Ik heb de ontwikkeling van deze composities in kaart gebracht in de eerste vier hoofdstukken die deel uitmaken van dit proefschrift: 'Internal Monologues', behandelt drie ensemblewerken die verhalen vanuit de eerste persoon belichten, die voortgekomen zijn uit een bewust of halfbewust discours, 'Unanswered Questions' behandelt een video en twee installaties die vraag- en antwoordstructuren door media heen onderzoeken, 'Voiceprints' gaat over werk waarbij het materiaal is gebaseerd op de manipulatie van gesproken stemmen, en 'Interactive Scores' kijkt naar mijn recente werk over algoritmische app-scores die de ideeën van 'music-text-film' uitwerken naar de interactieve muzieknotatie.

Ik heb geprobeerd om de verschillende manieren te benadrukken waarop het idee van 'stem' kan functioneren in sommige van mijn 'music-text-film' als een manier om de dynamiek van multimedia-werk in het algemeen te verwoorden. De verschuivende perceptie van wat 'stem' kan zijn is een dwingend aspect van deze vorm van muziek-tekst-film, omdat het van een puur verhalende vorm verandert naar 'stem' als sonische expressie, naar het publiek dat zich bewust wordt van zijn eigen innerlijke stemmen terwijl de geprojecteerde tekst leest in resonantie met de muziek. De vraag wat een stem is, vormt uiteindelijk de kern van dit onderzoek, omdat de stem zich verplaatst van drager van betekenissen en verhaalstructuren

naar de wegwijzer in het systeem waarin onze aandacht zich verplaatst tussen de vele lagen van verschillende media.

Biography

Yannis Kyriakides was born in Limassol, Cyprus in 1969, emigrated to Britain in 1975, and has been living in the Netherlands since 1992. He studied musicology at York University (BMus) and later composition at the Royal Conservatory in The Hague (MA) with Louis Andriessen and Dick Raaijmakers. He currently lives in Amsterdam with his wife and two sons.

As a composer and sound artist he looks for ways of creating new forms and hybrids of media that problematize the act of listening. The question as to what music is actually communicating is a recurring theme in his work and he is often drawn to the relation between perception, emotion and language and how that defines our experience of sound. In the last years, his work has been exploring different relations between words and music, both in concert compositions and installations through the use of systems of encoding information into sound, synthesizing voices and projected text.

He has written over 150 compositions, consisting mostly of music theatre, multimedia and electroacoustic works for chamber groups and large ensembles. His work has been performed worldwide at many of the prominent music festivals, and by many leading contemporary music ensembles. His opera *An Ocean of Rain*, opened the Aldeburgh Music Festival in 2008. He has been featured composer at both Huddersfield Contemporary Music Festival in 2007 and November Music in 2011. In the last years his sound installation work has been receiving more exposure and he contributed two works for the Dutch Pavilion at the Venice Biennale in 2011.

Prizes have included the Gaudeamus prize in 2000 for *a conSPIracy cantata*, a French Quartz award for the CD *Antichamber*, the Dutch Toonzetters prize for *Paramyth*, the Willem Pijper prize for *Dreams of the Blind*, an honorary mention at the Prix Ars Electronica for the CD *Wordless* and in 2014 the first prize in the International Rostrum of Composers for *Words and Song Without Words*.

Together with Andy Moor and Isabelle Vigier he founded and runs the record label for new electronic music 'Unsounds'. He is a founding member of the electroacoustic ensemble MAZE, and teaches composition at the Royal Conservatory in The Hague. His scores are published by Donemus, NL. More information and a full list of works can be found on the website: <http://www.kyriakides.com/>

