DAVID WILLIAM NOVACK

SOUND DESIGN FOR FILM: LIBERATING FROM IMAGE, SEDUCING SURRENDER

Orientador/Advisor Prof. Doutor Luís Cláudio Dos Santos Ribeiro
Co-Orientador/Co-Advisor Prof. Doutor Manuel José Carvalho Almeida Damásio

Lusófona University of Humanities and Technology
School of Communications, Architecture, Arts and Information Technology

Lisbon
2022

DAVID WILLIAM NOVACK

SOUND DESIGN FOR FILM: LIBERATING FROM IMAGE, SEDUCING SURRENDER

Thesis publicly defended at Lusofona University of Humanities and Technology on the 22 of July, 2022, nomination dispatch number 242/2022 for achieving the Doctorate Degree in Media Arts before a jury with the following composition:

President: Doctor Victor Manuel Esteves Flores, Lusófona University of Humanities and Technology (ULHT);
Members: Doctor Deirdre O'Toole, Dun Laoghaire, Instituto de Arte - Fac. Dublin (IATD Ireland);
Doctor Elena Rusinova, Instituto Gerasimov de Cinematografia - Moscovo (VGIK Russia);
Doctor Gonçalo Alves Gato Lopes, Lusófona University of Humanities and Technology (ULHT);
Doctor João Manuel Marques Carrilho, Lusófona University of Humanities and Technology (ULHT);
Advisor: Doctor Luís Cláudio dos Santos Ribeiro, Lusófona University of Humanities and Technology (ULHT).

Lusófona University of Humanities and Technology

School of Communications, Architecture, Arts and Information Technology

Lisbon

2022

The knower of the mystery of sound
knows the mystery
of the whole universe.

(Inayat Khan, 1996)

DEDICATION

This research and thesis are dedicated to my students over the years with whom sharing my knowledge has been a joyful journey, and from whom I have learned so much.

ACKNOWLEDGEMENTS

I offer my deepest thanks to a number of people throughout my long career, which has led me to this point of intersection, a collision of seemingly disparate pathways that find their way to a singular point, a land where, as the narrator in *Velvet Goldmine*'s mellifluous alto entices us to ponder, "all things are perfect. And poisonous."

To have the gift of music and the double gift of parents who support its development is indeed a deep privilege. My parents, Lee and Sy Novack are due my first and sadly posthumous gratitude, along with all my music teachers along the way, from early piano and bass at home to orchestration and arranging at Berklee College of Music where I first studied recording technology, for there is no ear training better than a musical education of theory and practice starting at a young age. The sadly-shuttered Sound One in New York was fundamental to completing my practical education as a film sound re-recording mixer, with further ear and technology development provided by Rocky "T-o-r, another T-o-r…" Tortorella and Orin Portnoy, who taught me how to use my "fuzzy little ears", and a career shepherded by the formidable Bill Nisselson who could manage clients, talent and a buzzing post-production facility inimitably, intuitively, grumpily and lovingly. Many directors and supervising sound editors along the way, far too many to list, gave me hours upon hours in front of the screen and the mixing board, where the creative, narrative, and emotive rhythms of film become ingrained through repetition, trial and error. It is in this room, with these talented people to my side and behind me, screen always in front and enveloped in speakers, where aesthetics are learned, developed, and subverted, and the intuitive power of sound in film manifests. Todd Haynes, however, will always stand out as having provided the greatest open palette, with films whose scripts, visuals and edits demanded phenomenological experimentation rooted in sound.

On the academic front, I owe special thanks first to the Cinema & Media Studies program at the University of Pennsylvania, particularly Nicola Gentili, Peter Decherney, and Tim Corrigan, for welcoming me into their sphere, trusting me with their students, and allowing sound to become a thread in their program. Secondly, Lusófona University, and in particular Manuel Damásio and Luis Cláudio Ribeiro, are due all my gratitude for bringing me into the very core of outstanding academic teaching and research in sound and film, in an international setting where I thrive, and with colleagues who are equally supportive and endlessly interesting. Luis Cláudio Ribeiro additionally is owed special thanks for shepherding me through the

process of formulating this thesis, with calm guidance and an open mind, and an endlessly deep and poetic intellect. Manuel Damásio also earns my gratitude for his page-by-page nitpicking of my first draft, pushing me forward and challenging me along the way. Lastly to my colleague and professor, José Bragança de Miranda for sharing discussions in aesthetics, throwing a massive reading library my way, taking me out for good southern Chinese food, and leading me to Genette early in my process, as if he intuited Genette's resonance with my explorations.

Special thanks go to CICANT, the Center for Research in Applied Communications and New Technologies at Universidade Lusófona and my colleagues and professors for their support throughout this process, including a grant for the completion of the content analysis, a critical component of this project. Many colleagues contributed with thoughts, comments and time, but in particular I offer thanks to Diogo Morais and Manuel Pita for assistance with quantitative and qualitative statistics, and to PhD student and my former Master's student, Tarun Madupu, who worked with me on processing qualitative data.

Finally my wife and extraordinary film editor Nancy Novack is due special thanks not only for introducing me to film in the first place (at the young age of 20) and continuously being a sounding board for my work, but even more for riding the sinuous curves of my career with me, supportively and always with chocolate in the cabinet.

ABSTRACT

Sound Design for film is a subject of increasing critical studies in recent decades. Michel Chion's foundational theories provide strong frameworks for discussing film sound, however most of his work and the work of others have theorized sound in constricted paradigms rooted in image theory, greatly limiting the discourse around the functions of sound in film. This doctorate thesis is a work of artistic research by a film sound practitioner and posits a path to liberate sound design from image first by critically examining sound theory, then articulating sound functions with broader cinematic, linguistic and philosophical writings. We then create an artistic output that sheds light on specific functions of sound design that are not well explained by existing theory, and interrogate those functions via statistically-based content analysis. With these functions established, we turn to Genette's narrative discourse as a possible avenue for establishing a rhetorical construction for sound design as a direct means of communication, and we apply an extrapolated model to an analysis of Todd Haynes' *Velvet Goldmine*. In conclusion, we find that sound subverts image paradigms altogether via a hybrid engagement we call *surrender*, and offers the listener-viewer a path, beyond representation, to experience.

Keywords: Film Sound, Diegesis, Rhetoric, Hermeneutics, Phenomenology

ABSTRATO

O design de som para cinema é um tema de muitos estudos críticos nas últimas décadas. As teorias fundamentais de Michel Chion fornecem estruturas fortes para discutir o som de um filme, no entanto, a maioria de seu trabalho e o trabalho de outros analisam o som em paradigmas restritivos, enraizados na teoria da imagem que limitam muito o discurso em torno das funções do som no cinema. Esta tese de doutorado é um trabalho de pesquisa e investigação de um profissional que propõe uma reflexão para libertar o design de som da imagem. Primeiro, examinando criticamente a teoria do som, depois articulando as funções do som com escritos cinematográficos, linguísticos e filosóficos mais amplos. Em seguida, criamos uma produção artística que lança luz sobre funções específicas do design de som que não estão bem enquadradas pela teoria existente e interrogamos essas funções por meio de análise de conteúdo com base estatística. Com essas funções estabelecidas, recorremos ao discurso narrativo de Genette como um caminho possível para estabelecer uma construção retórica para o design de som como meio de comunicação, e aplicamos um modelo extrapolado para uma análise do Velvet Goldmine, de Todd Haynes. Em conclusão, descobrimos que o som subverte os paradigmas da imagem por meio de um "engajamento" híbrido a que chamamos de rendição, oferecendo ao ouvinte-espectador uma via, além da representação, para a experiência sonora.

Palavras-chave: Som, Diegese, Retórica, Hermenêutica, Fenomenologia

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS OR TABLES

INTRODUCTION

PERSONAL BACKGROUND AND MOTIVATION

In 1986, with an engineering degree completed, a solid interest in film, a deep background in music, and no idea what to pursue for a career, I found myself listening to, not watching, a music video of Phil Collins. This sparked the clear image that behind the musicality was a myriad of technical engineering and creative genius that made his pounding drum sound like the forward march of something coming "In The Air Tonight" (Collins, composer, Padgahm, producer/engineer 1981). The audio engineering infused the natural sounds with a new spirit, transforming Collins' drums from well-performed but ordinary sounds to extraordinary ones that took on notes of a human character, with immense power that can be difficult to describe. And yet, with this power came enough wisdom and humility to get out of the way of vocals and other instruments. The band made the music, but the engineers made the sound – the sound that would emerge from my speakers, send energy across the room, and create periodic motion in the air molecules beside my eardrums. The singular revelation that day was that those waves of sound were not analogous to the ones in the recording room with the band, rather they were analogous to those in the control room, where mixer/producer Hugh Padgham touched knobs, buttons, dials, and faders to breathe a heightened, hyper-real transformative breath of life into the music (Flans, 2005).[1] With an engineering degree followed by a degree in music, I found myself sitting at the mixing board for 15 years of mixing sound for films, in a constant dance between the overlapping spheres of technology and art. I came to understand that my ears and fingers worked on recorded sound day after day, sculpting it into something intangibly powerful and oftentimes subversive. When whispered dialogue can be leaned into and understood, when background crickets can intensify at just the right moment of narrative tension, when music can drive not only emotion but cognition, and when the impression of a projected image on film can become rendered real through hyper-realistic synchronized sound from a different source, subversive work is being done that engages a

[1] I would later learn that engineer Hugh Padgham accidently came upon a technique utilizing the talkback microphone, then added dramatic gating, compression, and reverb to create the signature drum sound (Flans, 2005).

listener-viewer[2] in a narrative and emotional conversation with the film itself. These powers, effects, impacts or functions come from aesthetic brushstrokes that film sound designers and mixers frequently make intuitively and that audiences experience emotionally, physically and cognitively.[3] Yet the canon of film theory pays disproportionately little attention to the sensory field that is half the audiovisual experience – sound. Now, well into a career of mixing and filmmaking, and with over a decade of teaching film sound theory at the University of Pennsylvania and Lusófona University, new questions have emerged from discourse in the classroom and the mixing studio that beg further analysis. This project represents an exploration in questions that have continued to beguile me as I reconcile film and sound theory with the intuitively artistic and technical processes of sound ideation and realization.

As noted, sound studies for film have been few and far between in the canon of film theory, creating what Elisabeth Weis and John Belton called "frustrations…from inadequacies and gaps in the critical literature," due to publications that mostly deal only with film music or sound's initial development out of the silent era (Weis & Belton, 1985, p. ix). In the 1980s and 1990s, with an explosion of writing on film in the digital age, much theory and criticism "focused on the image rather than sound" while "the soundtrack quietly made its way into the digital era with little notice, somewhat lost in the critical conversation" (Whittington, 2013, p. 62). Sound for film, therefore, has been and continues to be largely a wide open field. When Michel Chion's fourth book on sound theory, *Audio-Vision: Sound on Screen*, enjoyed its English-language debut in 1994, sound designer/mixer, picture editor and lecturer Walter Murch noted in the introduction that "it is symptomatic of the elusive and shadowy nature of film sound that Chion's four books stand relatively alone in the landscape of film criticism, representing as they do a significant portion of everything that has ever been published about film sound from a theoretical point of view" (Chion, 1994, p. ix). Thus regarded as a kind of father of film sound theory with an impressive body of published works that have continued to the present, Chion has provided researchers and theorists fertile ground for further development. While Chion's work has been groundbreaking and deeply insightful, of course

[2] Note that we do not consider an audience member only a "viewer" and rather use the term "listener-viewer". We also recognize that in current theories of embodiment, audiences also smell, taste and feel in the filmic experience. However, sound and visual stimuli in the theater are real (light and sound waves) whereas the other senses are triggered primarily through phantom mechanisms.

[3] These words - powers, effects, impacts, functions – are weighty words in academic writing, loaded with assumptions that I seek to address later in this work.

it cannot be considered conclusive, and as a sound designer and teacher there are times accepted theories leave me feeling as if much has been either overly simplified or not yet addressed. On deeper examination, and as will be shown throughout this work, it seems that most of Chion's output, and the works by other writers of film sound, most often reconcile sound theory into existing paradigms of image-oriented film theory, from screen-based diegesis to a simplified semiology of sound, from synchronization to rhetorical relations between dialogue, which is essentially screen-based, and everything else. While this may result in an attraction for contemporary film criticism to pay more attention to sound in the near future, as some film theory moves towards cognitive studies, embodiment theory, and a post-modern psychoanalytic or narrative viewpoint, as a sound artist I have become intrigued by rooting an exploration in artistic research as a means of sharing artistic knowledge through a constructed sound design, that can then by applied as a case study that exposes some of the difficulties in applying current theory. The artistic output could then further be used as an object for content analysis to interrogate the assumed powers and functions of sound that serve as desired outcomes for filmmakers and sound designers, something that I have not yet found in any other literature.[4] Then that same output can be used as a jumping-off point to liberate film sound theory from image-based paradigms by adapting literary rhetorical theory to it, and then testing the approach as a case study applied to a modern film, a work previously done by me and therefore a prime work to utilize for further transfer of artistic knowledge. It is my hope that the journey embodied in this thesis opens the field for more such explorations.

THOUGHTS ON ARTISTIC RESEARCH

This research pivots iteratively on a central artistic output created by me specifically for this exploration and, as such, may be regarded as artistic research or practice-led research while the work itself, informed by research, may be termed research-led practice. The work is an audiovisual piece for which sound design, in the form of manipulated sound effects, has

[4] I have read a substantial portion of theoretical content available on film sound, whether single-authored or anthologies, and none contain or refer to statistical content analysis of the functions of sound. Searches in academic databases result in some cognitive studies around sound perception for film and other media, but none that I have encountered examine impacts or effects of sound effects / sound design through content analysis in a statistically-based study.

been added to an excerpt of the film *Koyaanisqatsi* (Reggio, 1982) that, in its original form, only had music and image. The resulting transformation of the original imparts embedded knowledge that emerges from my tacit understanding of sound design and mixing, acquired over years of artistic work. What I have created is not meant to be an improvement of the original nor a groundbreaking reimagining of it. Rather, I have approached its creation as an artist who wishes to materialize an audiovisual text that is more of an illustrative template, embodying many functions of sound design so that the piece can be used in three articulations throughout this exploration, articulations from which additional knowledge emerges: (i) as a case study for interrogating existing sound theories; (ii) as an audiovisual object for qualitative and quantitative analysis that explores reception of sound design; and (iii) as an audiovisual text in conversation with literary rhetorical theory in an effort to adapt the theory to film sound. These methods are applied rigorously, but as artistic research there is a discourse that emerges from the inquiry that is, in itself, a fundamental output of the research. In his introduction to Art Practice as Research: Inquiry in The Visual Arts (2005), Sullivan contrasts research by artists with traditional research, highlighting similar yet complementary paths. "What is common is the attention given to rigor and systematic inquiry, yet [for artistic research] in a way that privileges the role imagination and intellect plays in constructing knowledge that is not only new but has the capacity to transform human understanding" (p. xi-xii). This approach, we argue, provides for us fundamental scaffolding that supports an inquiry rich in descriptive output; from insights drawn by our content analysis to the audiovisual text that has been created for this research. Our text has the ability to "transmit knowledge in non-verbal and non-numerical terms" (H. Smith & Dean, 2009, p. 3). For audiences engaging with this text, even as subjects of content analysis testing, knowledge is transmitted via experience and, most certainly as we will show, differently than it would have been without sound design. Moreover, my artistic output address all three of the primary functions of creative work in research as set forth by Smith & Dean: (i) as a site of application of basic noncreative research via quantitative and qualitative analysis; (ii) as an artwork whose creation is directly informed by research and tacit knowledge, and (iii) as a work that becomes a vehicle of research through "the documentation, theorization and contextualisation of an artwork – and the process of making it – by its creator" (pg 3). This last function is realized most elaborately through the discussion of qualitative results, and also as a case study both for critical analysis of sound theory and extrapolation to sound theory from literary theory. Sullivan articulates this triad

differently in his chapter Making Space: The Purpose and Place of Practice-led Research (Sullivan, 2009) referencing a "triangle of practices" – conceptual, dialectical, and contextual (p. 48). Through conceptual practices "artists give form to thoughts in creating artefacts that become part of the research process." Certainly our creation of the sound-designed version of *Koyaanisqatsi* with the tacit knowledge of the artist is a conceptual practice. An exploration of meaning-making is at the heart of dialectical practices and this is the core of our qualitative analysis as we interrogate audience reception of the artistic work. It is important to note here that these experiences are not just "felt" and "lived" but also "reconstructed and reinterpreted" recognizing that "images and objects are texts that carry forms of cultural coding that require analysis and dialogue to create and communicate meaning." Lastly Sullivan positions contextual art practices as those that make use of "cognitive processes that are best described as 'thinking in a setting'", a primary characteristic of the creation of our piece and the handling of data that might be "local in focus but global in reach." Each of these paradigms of artistic research set out by Smith & Dean and Sullivan has a strong role in our process and methodology and will be enunciated over the course of this thesis.

Not only do we therefore have an iterative model in terms of revisiting the artistic work in several paradigms of research, but for me the process of creating the work in itself is an iterative one. Sound selection and editing, sound placement, and the unlimited sonically-transformative techniques of mixing all require endless repetition as I work with the elements and tools, resulting in what Heidegger's refers to as "the particular form of knowledge that arises from our handling of materials and processes"[5] (H. Smith & Dean, 2009, p. 6). Smith & Dean further cite Bolt who refers to this as "praxical knowledge" that might bring about "a shift in thought". Our exploration seeks to materialize my own praxical knowledge via the audiovisual object and the conceptual research that both informs it and is fueled by it.

Beyond conceptual research, however, we include both quantitative and qualitative analysis. Our goal to utilize content analysis stems from the aforementioned experienced understanding, or praxical knowledge, that sound design and mixing have goals beyond subservience to image. Yet, my readings from Chion to Weis & Belton, Bordwell & Thompson to Beck & Grajeda, and Altman to Sider have only suggested the relationship

[5] Bolt, Barbara (2007), 'The magic is in handling', in E. Barrett and B. Bolt (eds), Practice as Research: Approaches to Creative Arts Enquiry, London and New York: I. B. Tauris, pp.29-34.

between sound design and the desired outcomes of filmmakers through analysis, never pointedly positioning desired outcomes as artistic goals. There is logic to this, for one cannot assume authorial intent without directly questioning practitioners, seldom done in the academic arena, or by testing in some way. But I am a practitioner as well, which offers an exciting possibility to juxtapose artistic authorship and praxical knowledge with reception. My content analysis seeks to materialize this connection by demonstrating that sound design does work to support desired outcomes. Smith & Dean note that "qualitative, quantitative and conceptual research are all approaches to research which creative practitioners will benefit from" and that the "unique combination of creative practice and research can sometimes result in distinctive methodological approaches, as well as exhilarating findings and artworks" (2009, p. 5). What we do is engage in a discourse between qualitative, quantitative and conceptual research, not only over the course of the entire journey but within the discussion of content analysis results. Our combined approach to inquiry does, we believe, result in challenging and interesting findings as well as a work of art that engages in a unique way.

The final chapter of this thesis is a departure from the *Koyaanisquatsi* piece but, nonetheless, can be regarded as artistic research and it is through this prism that we offer it to our readers. It is an analysis of a film, *Velvet Goldmine* (Haynes, 1998) mixed by me over 20 years ago. It still holds within it my praxical knowledge, ripe for critical evaluation through the application of our modified literary rhetorical theory. As an artist of the piece, the brushstrokes of creation are visible by me, the notes and phrasing and orchestration of the sound (dialogue, music, and sound effects) are audible by me, and through this knowledge I am able to transfer a more rich understanding from the intersection of theory and artistic practice. Perhaps artistic research and the somewhat complex methodologies that have arisen out of this exploration are a bit unconventional. However, as Sullivan notes,

> It is evident from past discipline histories that imaginative investigations that breach accepted practices and challenge assumed canons contribute in a profound way to the core of our understanding. This is the legacy of what artist-researchers have to offer (Sullivan, 2009, p. 48).

It is our goal to make such a contribution and so, moving forward, we organize this thesis around addressing a fundamental exploratory ambition via interrogation in four areas of inquiry.

THE OVERALL AMBITION – LIBERATING SOUND DESIGN

Can we theoretically liberate film sound – dialogue, sound effects, music and mixing – from strict image paradigms, and then bring it back into conversation with image on equal ground?

This overall ambition begs a strategy, a path to liberation. In order to achieve this, we have mapped out four areas of inquiry and approaches for addressing them. Below are the inquiries which define aims and objectives, followed by an iterative process that expands approaches to theorizing, springing from artistic research that serves as the core of this thesis.

1) Are current, ubiquitous, structural techniques of discussing film sound adequate?
2) Are basic assumptions in the literature regarding intentional *desired outcomes* of filmmakers via sound, demonstratable, and therefore defendable?
3) Can literary rhetorical theory be adapted for the sound film?
4) Can such an adapted theoretical framework be tested against a complex sound design via a case study?[6]

Ultimately, these interrelated inquiries seek to draw connections that address a fundamental concept: that the strengths of sound for film go beyond typical notions of discursive style, emotional signaling and diegetic positioning that are common to film sound analysis in the contemporary canon, and therefore warrant an exploration in other fields of discourse. To accomplish this, our method draws on a combination of artistic research (both in the form of a new sound design and a reflection on a prior artistic work), quantitative and qualitive content analysis, and literature review as laid out below. Let us revisit each of these inquiries with a summary of the methodology applied to each.

[6] We considered a fifth inquiry, whether we can define sound design elements as narrative rhetoric and then look at such rhetoric as a transactional actor with a network, that network defined by the work of art that is the complete film. Since this seems to not ultimately deal with an aesthetic look at sound design, we have chosen to save this for later research.

INQUIRIES AND METHODOLOGY

Each inquiry warrants its own methodology, outlined here, charted below in Table 1, and according to the visual research roadmap offered below in Figure 1. Additionally a brief discussion of methods not used is provided.

Inquiry 1:

Are current, ubiquitous, structural techniques of discussing film sound adequate?

Methods: Desk Research in the form of literature review and case study, Artistic Research with self-reflective case study.

To examine this area, we will first apply literature review to discuss *sound design* itself as a term, then briefly frame the technological developments that accompanied the rise of the use of the term, and finally lay out relevant theories to our work, especially those of Chion. Here we look at Chion's primary theories of sound diegesis through various sources and the rhetorical structure of the "said/shown" relationship (Chion, 2009).[7] Diegesis is first articulated through a more extensive historical understanding of the term. For rhetoric as well, we first position film analysis in a rhetorical construct. Lucy Fischer's sound editing devices are helpful to define terms and show sound editing as neutralized montage according to Eisenstein (Fischer, 1985). The discussion is then expanded beyond sound theory to include immersion, embodiment, hermeneutics and phenomenology. The relevance of this discourse is rooted in establishing language and concepts that will be used throughout the report. We then move to artistic research to create a sound design by adding sound effects (SFX) to a segment of a film that has none, *Koyaanisqatsi* (Reggio, 1982). This version with sound effects is examined rigorously as a self-reflective case study, using the relevant theories and exposing unanswered questions, contradictions and general shortcomings of the application of these theories. The use of this constructed sound design offers the reader of this thesis a single audiovisual source for engaging in the discussion which is then further used in content analysis in inquiry two, and moreover provides a common audiovisual reference of sound design that echoes throughout the thesis via an iterative use of the artistic output.

[7] See sections 1.3.1 and 1.3.2, pages 33 to 63.

Inquiry 2:

Are basic assumptions in the literature regarding intentional *desired outcomes* of filmmakers via sound, demonstratable, and therefore defendable?

Methods: Desk Research in the form of literature review, Artistic Research via utilization of artistic output as an object of study, Content Analysis (quantitative and qualitative).

This area of examination is fundamental to putting sound on an equal playing field with image, for it diffuses criticism based on claims of biased assumptions. We cannot talk about sound as, for example, engaging attention, contributing to meaning-making, or impacting emotion without first demonstrating this to be true. In order to interrogate this question, we begin with literature review around the notion of *desired outcomes* as a requisite to technological innovation and diffusion, first examining through discursive reading of technology theory whether sound design satisfies other requirements of a technology. We then define testable desired outcomes thereby pinning sound design to narrative clarity, engagement, and thematic interpretation. While Brian Winston (1998) and Everett Rogers (1983) form the basis of defining the terms and requirements of a diffusible technological innovation, which we apply to sound design, additional discourse is required to inform our choices of desired outcomes to test. For this, we connect to the broader sound engineering field where Schreger (1985) posits a continuum of sound-fidelity obsession dating to the wax cylinder and taking us through Dolby. Kittler asks us to regard sound as the *real* and a number of writings on psychoacoustics are employed to defend that position, particularly Sonnenschein (2001). One neuroscience study by Hughes (2001) demonstrates the way imagined sound is evoked through anticipated but unrealized sound omission, supporting the notion of interactive engagement and cognitive sound processing.[8] From here, we are able to define our desired outcomes for survey design and test them through quantitative and qualitative content analysis, specifically with respect to sound effects (SFX). To perform this test, we use our artistic output, the SFX-

[8] Note that interactivity in this sense does not imply being an actor in narrative or "story" decision-making, as it does in many discourses on virtual interactivity. We are particularly drawn to this notion of interactivity as it posits a much less passive experience in media listening-viewing than is typically supposed.

designed *Koyaanisqatsi* segment. We screen each version, with and without sound design, to separate audiences and provide a survey with both quantitative and qualitative questions. We begin our discussion with quantitative analysis which seeks to measure audience response with respect to the desired outcomes. We then move to qualitative analysis. Analyses and conclusions cross-pollinate the quantitative and qualitative results, bearing fruit that gives deeper insights and raises additional questions, exploring new knowledge via methods that align with artistic research as earlier stated. Addressing this fundamental inquiry hopes to establish if sound design does, in fact, achieve predictable benefits, putting to rest criticisms in research regarding certain assumptions of the impacts of sound design as well as authorial intentions. Yet interestingly, the full discussion, particularly of qualitative results, also supports the need for new paradigms of film sound analysis, further justifying our third area of inquiry which follows.

Inquiry 3:

Can literary rhetorical theory be adapted for the sound film?

Methods: Desk Research in the form of literature review, Content Analysis results as insightful input, and Case Study on Artistic Research output.

The discussion of sound as it relates to traditional models of film diegesis has thus far resulted in an inadequate depiction of sound relative to both the visual frame and narrative, as will have been discussed and illustrated in inquiry one and elaborated upon in the qualitative analysis performed in inquiry two. Turning to another source, we examine sound as an actor within the broadened scope of the term *narrative* as defined by Genette, applying sound design to his detailed work on tense, mood and voice as well as to the Proustian scene (Genette, 1983) in which, for us, sound editing and mixing techniques act upon the narrative in Proustian terms. Note that via Genette, “narrative” is redefined much more broadly than film theory typically offers and it is in his richer definition that sound as narrative finds particularly strong relevance. Kittler (1999)further expands the conversation to the dance between the real and symbolic since, unlike with written literature, we are addressing the very real way in which sound functions in a screening environment. That environmental scaffolding moreover suggests that sound design de-territorializes the screen (Deleuze & Guattari, 1987), carrying narrative through physical space and expanding to the audience

through various mechanisms, most notably narrative counterpoint. Film is shown to be a Deleuzian audiovisual map which can break the diegetic construct or perhaps subvert it by exploiting the tension between, on one side, the audience's sense of boundary given visually by the frame and, on the other, the lack of auditory boundary offered by sonic immersion. This exploration circles back to both the *Koyaanisqatsi* artistic research case study and the insights drawn via content analysis, and also brings many other cases into conversation, suggesting that Genette may function as a meaningful model for discussing sound in film. This brings us to the test case addressed in area of inquiry four.

Inquiry 4:

Can such a theoretical framework be tested against a complex sound design via a case study?

Methods: Desk Research in the form of literature review, Content Analysis, and Case Study engaging with an artistic output from this sound artist's past works.

For this we turn to a case study, utilizing Genette's analytical framework as explored in the previous inquiry, and applying it to a sound design mixed by this author over 20 years ago, *Velvet Goldmine* (Haynes, 1998), leveraging praxical knowledge in an iterative process that re-engages the artist with his work via fresh analytical methods and insights. The choice of film is discussed and analyses are conducted of the opening 15 minutes and a short sequence towards the end of the film, bringing back into discourse not only Genette, but theories and terminologies discussed throughout this thesis, as well as insights from our content analysis. Here we not only take an applied theoretical approach with respect to Genette but, by bringing into discourse other linguistic and theoretical concepts – semiotics, phenomenology, embodiment, hermeneutics – we seek to articulate a much more open-ended dialectic of sound design that is truly liberated from image and brought back on equal grounds.

Fundamentally, our research is exploratory, a journey by which we seek to uncover not only ways in which film sound functions but also an expanded discourse with which to examine, discuss and teach film sound that more holistically embodies the experience of sound in cinema. It is this exploration that drives our artistic research, a personal search where we wish to interrogate existing theory against our experience as artists, researches and educators who feel intuitively that, as has been discussed, sound is something of an orphan child in the

world of film theory. To liberate the child and usher them into adulthood, we feel, requires the hybrid approach we have described, with areas of inquiry and methods that mark small milestones along the journey. We offer Table 1, below, as a summary of methods used, a snapshot of our "imaginative investigation" that "challenge[s] assumed canons", hopefully honoring "what artistic-researchers have to offer" (Sullivan, 2009, p. 48); in the realm of film studies and, specifically sound design for film.

Exploratory Research to address our overall ambition: Can we theoretically liberate film sound from strict image paradigms, and then bring it back into conversation with image on equal ground?	
Areas of Inquiry	Methods
1) Are present theories adequate?	• literature review (existing film, linguistic, philosophical theory) • case studies (citing cases throughout theoretical discussion) • artistic research with case study (sound design with reflective case study applied to sound design)
2) Are desired outcomes defendable?	• literature review (existing theories of technology applied to Dolby and sound design) • content analysis (quantitative and qualitative study to interrogate desired outcomes of sound design, applied to artistic research output)
3) Can literary theory be adapted as a possible framework?	• literature review (Genette's narrative discourse explained, adapted, and in discourse with earlier-mentioned theories) • content analysis (integration of insights from content analysis into discussion) • case studies (using artistic research output and others as cases throughout theoretical discussion)
4) Can an adapted literary theory work in a case study?	• case study from artist's repertoire (specific case of Haynes' *Velvet Goldmine* (1998)) • literature review (Genette's narrative discourse in conversation with earlier-mentioned theories throughout case study) • content analysis (integration of insights from content analysis)

Table 1: Areas of Inquiry & Methods

Beyond those approaches listed above, there are a number of methods that had also been considered and are worth a brief discussion here since each informs possible future work.

With respect the relationship between desired outcomes and audience response, Janice Radway's 1984 *Reading the Romance*, (Radway, 2006) presented compelling ideas in applying

ethnographic studies to sound theory. However, we determined that this would not result in data that address the area of inquiry. We are not concerned with the texture and composition of the audience, but rather with their immediate response to stimuli. This could be an interesting approach to a question that looks at the marketing of new sound technologies, i.e. Dolby Atmos, and the demographics of audiences to whom these innovations in sound design are most consciously appealing. This however is not that study.

Critical discourse analysis could be a potential vehicle to examine inquiry one, film sound theory as overly rooted in diegesis and other image-based constructs. However, this would necessitate a deep analysis of the power dynamics within film theorists over time – dynamics that certainly exist but would take immense research to expose and would form a thesis of its own. Certainly, interesting questions would emerge from such research.

We also thought to potentially expand this to look at sound theory as a minor literature within film criticism, suggesting further that sound design itself is equally a minor literature with subversive power within the audiovisual sphere. However, a study such as this is reflective of the kind of research we hope this project might inspire.

In addition, since one of the points of discussion that repeatedly emerges in this study deals with the power of sound itself to work in the intersection of phenomenology and reception without the receiver's conscious awareness, critical discourse analysis could be used to examine the dominant structure of film sound manipulators, thereby empowering audiences with a greater understanding of those tools of manipulation, and reducing susceptibility. For instance, in 1928 Sergei Eisenstein sought to neutralize the image, liberating it from the dominant power structure. In fact, he resisted synchronized sound for film altogether for fear it would undermine the neutralizing of image (Eisenstein et al., 1988). Now, sound is arguably neutralized when used metaphorically or as pure montage, and yet the tools of manipulation, the sound apparatus, remain securely behind the wizard's curtain. Critical discourse analysis could expose that apparatus if we could define an appropriate object of such a study. To do this, we explored Van Dyke to examine, in particular, the notion of ideology embedded in sound design rhetoric via lexicon, propositions and semantics (Dijk, 1993). We came to the conclusion, however, that this would be beyond the scope of this thesis, although it may be an interesting avenue to pursue in the future, notably in a study rooted purely in an elaboration of sound design as rhetoric.

In order to collect all the literature review and data in a meaningful manner, we designed to following shape- and color-coded research roadmap to offer us a visualization of our overall exploratory journey that can be easily shared with our readers.

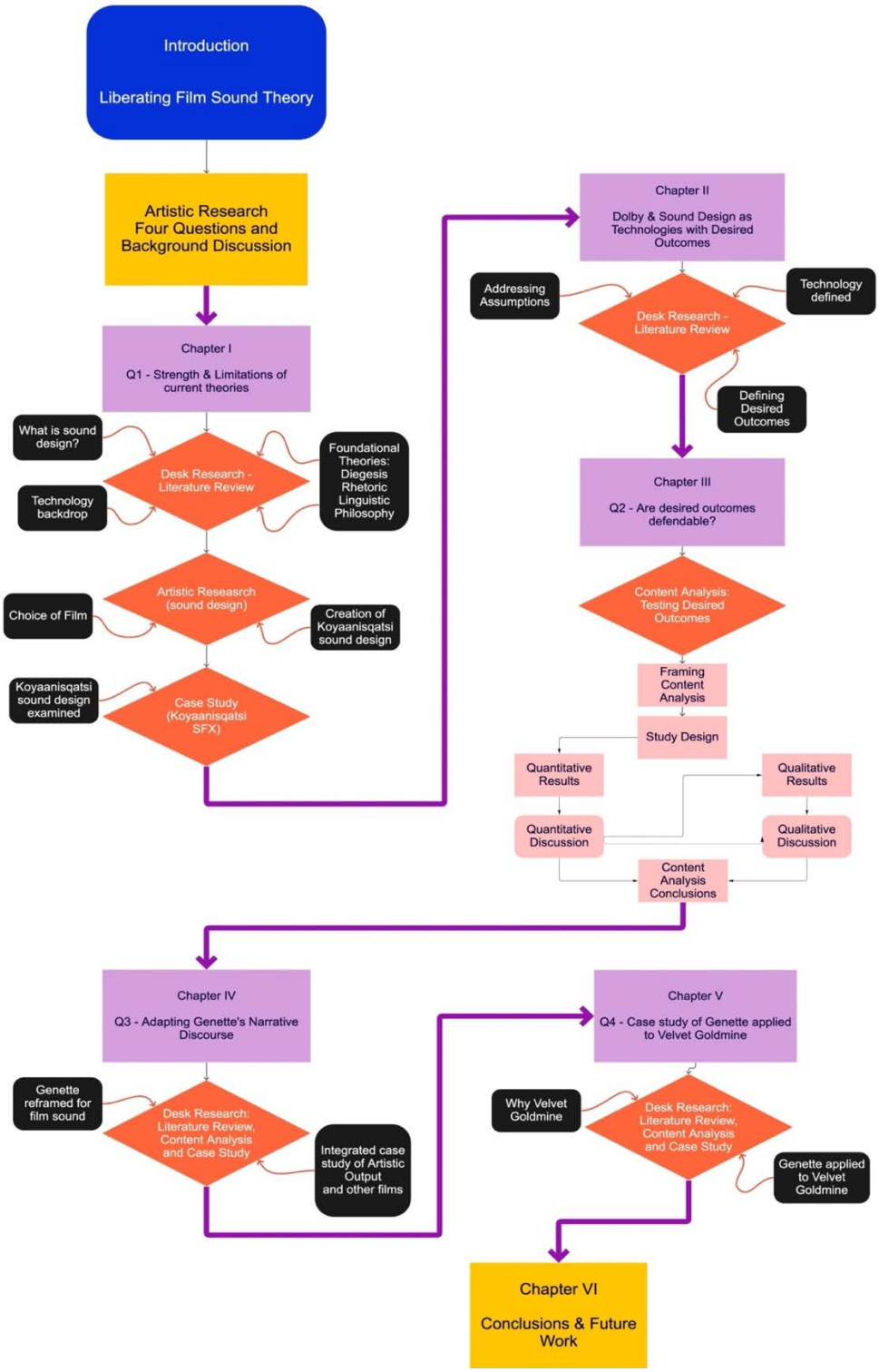


Figure 1: Methodological Roadmap

CONTRIBUTION OF WORK

This research has a number of ambitions that represent significant contributions to the literature and research currently available on sound for film. Firstly, the most fundamental inquiry of this thesis aligns with the primary ambition; to liberate film sound from image paradigms and bring it back into discourse on equal ground. This represents a major shift in how film sound is most frequently critiqued and taught within film schools, hopefully resulting in increased interest in making sound a more fundamental part of film research and education.

With regard to specific critical applications, this thesis makes a contribution by looking to theories outside film studies in general, in our case literary narrative discourse, as a means to position sound on equal rhetorical footing as image in the universe of film theory.[9] Our hope is that narrative discourse will be applied to other films by other scholars, and that other types of communications theory outside current film-theory paradigms will be explored by other scholars in their application to film sound as well, encouraged by this work. We additionally articulate our research with embodiment, phenomenology, hermeneutics and immersion, all of which are areas that hold promising possibilities with sound studies and, above all, highlight the interdisciplinary nature of fertile sound criticism.

The content analysis within this thesis seems to be the first of its kind, testing sound effects for targeted outcomes in a statistically rigorous manner with some anticipated and some unexpected outcomes. We believe this forms a basis for further experimentation in this area with methodologies modeled on and modified from our content analysis, including additional studies by our team.

The qualitative portion of the content analysis exposes notions and questions related to audience studies, psychology, cognitive studies, psychoacoustics and neuroscience, and therefore can be used by other researchers in these areas as a springboard for additional research. Additional hypotheses emerge from the discourse herein that can be leveraged for further content analysis or deeper discourse via desk research. That these questions emerge from our analysis we consider to be a major contribution as it exposes the need for critical work in this field.

[9] Note that the positioning of "rhetoric" in this discussion will be deeply explored. It refers more to literary argument than persuasion, in our sense.

Finally, artistic research is at the forefront of debate in both research and pedagogical circles. The artistic research that frames the entirety of this report via multiple iteratations of creation and examination, that is, the creation and use of a sound design specifically to examine theoretical functions, is a major contribution in the area of sound studies for film largely due to the unusual nature of a combined film-sound practitioner/researcher at the helm of this thesis. While this type of work has been done in other areas of film and in other areas of sound studies, it is quite unusual in film sound studies, likely due to the rarity with which practitioners, including sound designers and mixers, are also academics. Our artistic research approach can further inform other areas of artistic research, particular in terms of grounding content analysis in artistic creation to interrogate specific functions of the artist's stroke, no matter the artistic field.

CHAPTER I. ARE CURRENT THEORIES OF FILM SOUND ADEQUATE?

This chapter deals with our first area of inquiry which examines current ubiquitous structural techniques of discussing film sound, particularly diegesis and its rhetorical relationship with image, and seeks to expose ways in which this discourse is at times inadequate. We begin with a look at the term *sound design*, and then provide the technology backdrop that is relevant to understanding the explosion of the use of the term sound design in recent decades, as well as providing an understanding of some of the artistic choices made in artistic research later in this chapter. We then move on to foundational theories that are particularly relevant to this study, framing them in their historical development and aesthetic implications. Lastly, using these underpinnings, we then apply artistic research to the *Koyaanisqatsi* excerpt by creating a sound design and utilizing it as a reflective case study in order to illustrate both strengths and weaknesses of existing theories.

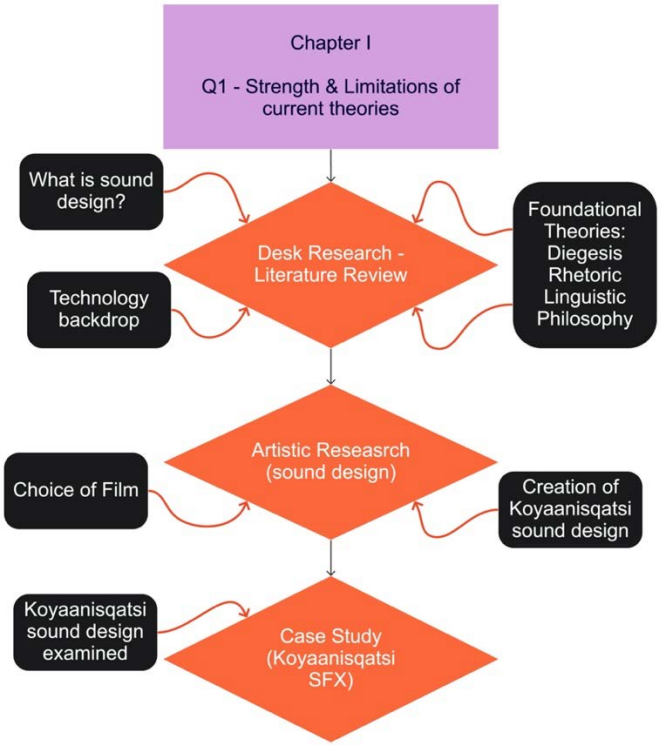


Figure 2: Chapter I Roadmap

1.1 WHAT IS SOUND DESIGN?

Sound design is a term that can often have different meanings in different settings. In live theater and opera, it extends from sound reinforcement techniques like microphone and speaker placement under the definition of "sound systems" to sound effects and/or musical sounds triggered live under the direction of a stage manager (Illuminated Integration, 2020). In music and sound synthesis, it is often thought of as literally the designing of new sounds, either from raw electronic sound signals generated by a signal generator or the manipulation of sounds derived from electronic synthesizers or sampled sounds,[10] as exemplified even in the title of a ubiquitous guide to synthesis, *Computer Sound Design: Synthesis, Techniques and Programming* (Miranda, 2005). For the purposes of film sound, it is thought of by practitioners, filmmakers and film producers as the conceiving, planning, creation, recording and/or selection of all sounds that comprise the sound of a film plus their manipulation through editing, mixing and other technologies to form the final sound served to audiences. One often-used glossary entry notes:

> Sound design [is] the process of selecting sounds, editing them, and mixing them together to form a desired artistic effect, usually used as a term when sound is accompanying a picture (Holman, 2010, p. 224).

Twenty-five years earlier, closer to the origin of the use of the title, lecturer and critic Marc Mancini, took a much broader view in his article entitled "The Sound Designer", where he states:

> [Sound Designers] guide the sound of a motion picture from beginning to end, interpreting the director's expectations, 'hearing' the script and storyboards, coordinating with the composer and sound editor, contributing to the mixing process, even ensuring that what is heard in the theater is of optimum quality. Decades ago, they might have been called supervising sound editors (they often still are), but that title has a craftsy connotation that downplays the true nature of their job: they are aural artists (Mancini, 1985, p. 361).

[10] A sampled sound is a sound that is recorded and then used as a sound object that can be manipulated. For instance, a modern electronic keyboard can sound very much like a piano or organ because of sampled sounds that are being played.

Historically, with the advent of sound film, the role of Sound Designer was in the hands of the picture editor first, and later the Supervising Sound Editor and or Mixer. But it is only with the bold sound work of editor/mixer Walter Murch for Francis Ford Coppola's *The Rain People* (1969) that the title of Sound Designer was first given mention in credits, a title that was again bestowed by Coppola to Murch with the groundbreaking sound design for *Apocalypse Now* (Francis Ford Coppola, 1979) of which much has been written and that will be referenced later in this thesis. It is of no surprise that this shift was accompanied by technological advances, which will be discussed in section 1.2. Of note, this final sound produced by the accumulated and aggregated functions of sound design is often referred to as a film's soundtrack,[11] not to be confused with a music publication also called a soundtrack, which is typically a compilation of music from a film or musical theater piece. To further clarify from personal experience, while sound design for film does at times incorporate strategies for recording on a film or television stage or on location, a process known as production sound recording, which calls for considerations around microphone choice and positioning for actors, the term most frequently comprises the post-production portion of sound, that is, the work done after image capture. For dialogue, this means selection, editing and mixing while for sound effects and music it may include their recording as well.[12] Furthermore, sound effects (SFX) are frequently divided into subcategories that line up most directly with the division of labor among traditional sound-editing staff prior to the digital era, where now less labor is needed: hard effects which typically come from sound effects libraries, Foleys which are SFX recorded in a studio synchronously to image projection, and background sounds or ambiences. Mixing is a critical component to sound design, but a film's mixers, beyond some low-budget scenarios, are not generally considered to be sound designers. And yet, it is through the hands and ears of a mixer that sound design is fully and finally elaborated and executed. Therefore, for the purposes of this writing and based on years

[11] In fact, Michel Chion argues that "there is no soundtrack", taking issue with the term "track" which implies, to him, some sort of container. To Chion sound for film has no container that corresponds to an image frame, and prefers the term "sound channel" (Chion, 2009, p. 226). He also notes that the "track" or "channel" that corresponds to the physical placement of sound's analogue on the side of the printed frame of film, usually an optical track, does not take into account what is excluded, removed, or not chosen in the sound of a film, the lack of which is distinctly part of the film's sound design and therefore not captured by the term "soundtrack".

[12] Note that ADR (automated dialogue replacement), when actors perform their lines in a studio during post-production, is typically not thought of as sound design or sound editing work, even though creative microphone placement strategies, selection of performances and editing are involved.

of professional experience, film sound design will be considered to include all aspects of post-production sound, including mixing, regardless of the breakdown of labor for a particular film. The overwhelming analyses herein focus on the selection, editing and mixing of sound effects (SFX). It is here where most of the mysteries, and the least amount of prior research, has been done.

1.2 TECHNOLOGICAL BACKDROP TO SOUND DESIGN

A complete history of sound technology in general and specifically for film would be beyond the scope of this work, however the critical developments beginning in the 1970s are quite relevant to sound design, our artistic research, and our aesthetic discussions. From the mid-1970s and into the 1980s, commercial cinema experienced a global change in the performative aspects of the theatrical experience rooted in Dolby surround speaker installations and noise reduction systems (Beck, 2008, p. 68). Theaters across the world were provisioned with Dolby Surround Sound and Dolby Spectral Recording noise reduction, in part due to a push by United Artists to elevate the sound environment for the release of *Apocalypse Now* (Francis Ford Coppola, 1979) (Dienstfrey, 2016, p. 167).[13] To movie-goers, this represented the first dramatic technological change in the exhibition space in a very long time. Arguably, for image on a mass scale, the jump to 35mm film projection in 1909 was the most dramatic visual change through the entire 20th century (Alsobrook, 2000)[14]. For sound, the introduction of optical track recording on film in the 1920s had been the most comparable change in movie-hall technology (Kellogg, 1955). Experiments in stereo, quadraphonic and other multichannel sound continued through the years (Dienstfrey, 2016, pp. 174–176), but it is with the industry-propelled marketing of Dolby and possibly, as Dienstfrey argues, the self-promotion of filmmakers and sound practitioners, that multichannel sound became the theatrical standard (p. 186).

[13] Dienstfrey notes that only 70mm prints of *Apocalypse Now* had surround capabilities, and this initial push only happened in 20 or so theaters, but this was the beginning of positioning Dolby as a brand that ultimately became ubiquitous in theaters while elevating the status of sound technicians through a period of economic uncertainty.

[14] Color film was not as dramatic a change as is often thought. Commercial experiments in color began as early as 1908, even though it did not become common until the late 1930s.

In brief, Dolby's surround matrix, which is still in operation in some theaters, allows four discrete signals that carry information for the left, center, right and surround speakers respectively, encoded onto the two preexisting optical tracks of a 35mm motion picture print. In the theater, the two optical tracks are decoded, back to four signals and sent to the four sets of speakers. This made earlier experiments with immersive surround sound feasible on a global scale, allowing for the dissemination of surround as an object of innovation. While the matrix enabled surround sound, Dolby Spectral Recording noise reduction, adapted from Dolby's music recording technology, was utilized throughout the post-production chain and in the theater to nearly eliminate noise from both magnetic recording and the optical track played in theaters. The result was a much quieter track with a greatly expanded dynamic range[15], a wider frequency range[16], and the ability to almost get to silence.

Along with Dolby's technical innovations came an explosion of creative output in sound from the late 1970s to the early 1990s, from *Apocalypse Now* (Francis Ford Coppola, 1979) to *Silence of the Lambs* (Demme, 1991), that had previously been constrained by technical limitations: the cumulative noise floor, a limited frequency response, and a single-point source of mono sound emission in most theatrical environments. Suddenly, the concept of a Sound Design was born.[17] In the 1990s and beyond, with the advent of digital sound for film, surround was made readily available in a 5.1 speaker configuration that allows for three speakers at the screen, left/center/right, two channels for surround speakers, surround left/right, and a subwoofer,[18] able to generate very low frequencies even below audible pitches – all with no noise added by the printing process due to the digital format. The earlier entrance of the subwoofer in 1977 marked the expansion of the use of low frequencies to "whip up emotions", a phrase cited by Beck from film reviewer Ursula K. Le Guin's review of a 70mm print of *Close Encounters of the Third Kind* (Spielberg, 1977) (Beck, 2008, p. 70).

[15] The decibel range from quietest to loudest that a medium can provide.
[16] The range from low to high frequencies that a medium can faithfully reproduce.
[17] Dienestfrey argues that Sound Design as a title was born out of self-promotion, and that its use decreased into the 1990s (Dienstfrey, 2016, p. 181). Beck notes that pressure exerted by the labor union IATSE to protect the job categories of Supervising Sound Editor and Re-recording Mixer contributed to this decline (Beck, 2008, p. 74). We posit that this may also have been driven by the Academy of Motion Picture Arts & Sciences' classification of these two job functions as awardable Oscar categories, where there is none for a Sound Designer.
[18] The subwoofer was not new here, for 70mm Dolby prints already had provisioned a magnetic (not optical) subwoofer track since the late 1970s. But because 70mm theaters were few and far between, this was rarely experienced by audiences until the development of the digital 5.1 optical system.

But it is with the advent of digital 5.1 and the release of *Batman Returns* (Burton, 1992) that the digital format with subwoofer is first made available in widely-distributed 35mm prints (Miller, 2004, p. np). In her chapter entitled "Dirty Sound", Coulthard notes that this format provides low-frequency sounds at the same volume as higher frequencies, which enables "spectacular effects that impact the body of the spectator" (Coulthard, 2013, p. 119). She further cites Kerins who credits the subwoofer with providing "a rhythm that is felt as much as heard [and] a pure gut […] straight-to-the-brainstem physical response" (Kerins, 2011, p. 83). Coulthard further shares that it provisions the "drone, a sound that makes emphatic the haptic, gut-stirring abilities of low frequency and makes acute the presence of noise as an aesthetic choice, one stressing disturbance and disorientation" (p. 119). Thus we have introduce a new and potent sound element that functions in ways that are beyond pure listening and linguistic signification. Eventually, the system is expanded to 7.1, an adaptation made by splitting the surround and rear speakers from one stereo pair to two pairs, and Dolby Atmos characterized by the addition of an array of side and ceiling speakers. We believe, however, the changes offered by these later developments, while relevant to sound design, are less dramatic than the initial industry-wide implementation of noise reduction, surround sound, and the subwoofer. We make special note of the subwoofer and surround, as they are critical to certain moments in our artistic research as well as to an overall aesthetic approach to sound design that highlights embodiment, items that will be discussed later in this work.

1.3 Relevant Foundational Theories

Theories of film sound have been generally sparse when compared with theories of image, especially before Michel Chion began to tackle sound in the 1970s. As previously noted, when his fourth book on sound theory, *Audio-Vision: Sound on Screen*, was released in English in 1994, Walter Murch noted in the introduction that Chion's previous four books were "a significant portion of everything that has ever been published about film sound from a theoretical point of view" (Chion, 1994, p. ix). In the same preface, Chion himself considers this volume to be a "definitive book on film as an art of sound," and notes that when he began down this path in 1978, "few works existed on the subject aside from a small number of articles that analyzed particular films" (Chion, 2009, p. vii). As further elaborated in Chion's preface, through the years some filmmakers pushed boundaries with the sound components of this

audiovisual medium, of note Vertov, Hitchcock, Resnais, Lang, Bresson, Kubrick, Welles, Duras, and many more. Some critics did take notice, including Raymond Bellour and Christian Metz, among others. But this small scattering meant that sound lagged behind the dominant positioning of image theory (pp. vii-viii). Rick Altman notes in 1985 that "more than half a century after the coming of sound, film criticism and theory still remain resolutely image-bound," and that with "with each new image-inspired theory, film study's exclusive image orientation gains ground" (Altman, 1985, p. 44). He further argues that this is a natural evolution originating in the Western world's "privileging of vision over all other senses" since the Renaissance and extending to Baudry's adoption of Lacan's visual metaphors, particularly the "mirror stage", as well as to the image dominance in Metz's semiotics (p. 45).

Indeed semiotics, which concerns itself with language, has most often been applied to the visual language of the film experience rather than the sonic. Christian Metz himself notes that the hierarchical placement of sound below image is a general cultural stance (Metz, 1977, p. 154). Furthermore, in his essay "Aural Objects", he puts forth that the words we use for sounds function more or less adjectivally; they describe rather than name (Metz, 1985, p. 156). We hear a buzzing – of what? – or a crash – of what? And yet it suffices to say "the sound of a bee" or "the sound of a car crash." The source of the sound, something we can see, is an object, whereas the sound is a descriptive of that object. Metz further positions this in "primitive substantialism," rooted partly in the "subject-predicate structure" of Indo-European language, thereby influencing Western philosophy. We further posit that visual culture, from which the graphic sign and writing emerged, prevailed over other senses in part due the recordability of the visual sign, whether on a cave wall or parchment, and eventually in the photograph and moving image. Sound, however, is an invisible, temporal, ephemeral phenomenon. Where visual culture and the visual object date back many millennia, the arrival of sound recording, thereby creating the sound object, is a relatively new event in human existence. Even with the sound object taking hold in the 20th century, a number of reasons persist that impede semiotic analysis, some related to the natural differences between visual and sonic perception which, for example, are rooted in sourcing, spatiality and temporality; concepts that are not easily accounted for in the semiotics of Saussure or Peirce. In Chapter III, "The Audible Semiotically Considered", of Andrea Valle's *Towards a Semiotics of the Audible*, we find a discussion of these key differences between sound and image when applying semiotics (Valle, 2015, pp. 74–77). In terms of actoriality or source, Valle cites Bregman, noting that,

> The auditory world is like the visual world would be if all objects were very, very transparent and glowed in sputters and starts by their own light, as well as reflecting the light of their neighbors. This would be a hard world for the visual system to deal with (Bregman, 1990, p. 37).

Here Valle addresses the notion that we hear from all directions and can hear multiple sounds from the same general direction as well, which blend, always dynamically, on our sonic canvas. Valle concludes that "the audible seems to be characterized by a sort of actorial competition" (p. 76). In other words, the relations between things simultaneously in the audio spectrum may or may not bear syntagmatic or diagrammatic relevance to one another. Even when they do, the parameters of sound that do play into a communicated event are far more diverse and less apparent than with image; things like timbre, reverb, pitch, direction, delay, or in summary the variety of acoustic and production characteristics of an array of simultaneous sounds. In terms of spaciality [Valle's spelling], Valle notes,

> If the visible space immediately presents itself as a unity that coordinates the mutual relations of the figures, the audible space is a field of multiplicity, a set of figures that precede their relations (p. 76).

This speaks to a component of sound communication that visual communication does not have; in the sound field, we hear and process spatially-related sounds without first being cognizant of the spatial relation. Transposed to semiotics, both syntagmatic and diagrammatic conventions, formal and contextual, are challenged when dealing with sound reception and, as will be discussed later, the line between denotation and connotation is greatly blurred. With respect to temporality, time is a key aspect of all sound and is not a requirement of visual language. Of course film language is temporal but for the visual component, time can be frozen to examine the image frame and its composition while no such parallel for sound exists. The sound field has an envelope, a graph *over time*, that represents its shape in terms of volume. No sound is temporally static nor can it be made to be static. Even an unchanging pitch is a product of a sound pressure wave in repeated periodic motion *over time*; without time, there is no pitch, there is no sound. Citing Dufourt, Valle shares that "on comprend à présent que l'évanescence de l'information acoustique soit la condition de son intelligibilité et qu'elle détermine sa

qualité"[19] (Dufourt, 1999, p. 77), summarizing with "presence is turned into presentation" (Valle, 2015, p. 76). Sound has no static presence, it is presented, over time.[20] These realms highlight the difficulty in summarily assigning categories of Peircean icon, symbol or index to sonic signifiers over the course of analysis, let alone bringing the diagrammatic element into play. It is into this field of an historically small body of work on film sound as a component of film language that Chion boldly steps with his project.

While Chion has embarked on both formal and critical work in many areas of film sound, here we will focus on two areas that are most relevant in that they exist in direct relation to image and are ubiquitous in academic teaching of sound theory for film: diegesis and the rhetorical relationship between what is said – dialogue – and what is shown – everything we see plus non-dialogue sound. For each, we will begin with an historical and linguistic orientation that informs our assessment of Chion and our journey to discover alternate paths of discourse.

1.3.1 DIEGESIS

Diegesis, of course, has a history of its own in terms of its definition and its relationship to narrative, and before we delve into Chion, it is relevant to look at the term. *Off Screen*, a long-running popular journal for film buffs now online, offered a forum on the subject in 2007, *Discourses on Diegesis* (Taylor et al., 2007). In it, Henry Taylor at University of Zurich offers a brief history, making the case that the term's use in film is a misnomer. According to Taylor, Plato noted two types of narrative, "simple narrative" or haple diegesis, which is characterized by a narrator in their own voice, and "mimesis" in which the voice of the telling has switched to a representational one, as in a character's voice in the first person. For Plato, the term diegesis is therefore reserved for referencing a narrator who tells a story, rather than coming from it. Taylor looks to Paul Ricœur in his *Time and Narrative* (Ricœur, 1990) to argue that Aristotle reassigns the term *mimesis* to include an external narrator, arguing that such storytelling is still representational, and calls this narration *diegetic mimesis*, as opposed to *mimetic mimesis*. Thus

[19] One immediately understands that evanescence (or unfolding over time) of acoustic information is a function of its intelligibility and determines its quality (trans. Novack).
[20] The "performative" nature of sound for film is of particular interest to us because of the hands of the recordist, editor and mixer at work over time. Every sound and the composite mixture of multiple sound is, in essence, a recorded performance.

with Aristotle, the term *diegesis* has been joined to representational storytelling, as for its current use, however it still refers to a third-person storyteller's voice, not the internal world of the story.

Eventually, the term diegesis comes to replace the term mimesis in some disciplines, particularly with reference to film, and always with some degree of ambiguity. This was initially part of a broader project by the early filmologists, including Gilbert Cohen-Séat and Etienne Souriau from whom, according to Guido Kirsten in his analysis of the evolution of Christian Metz's theories, Metz drew vocabulary and inspiration (Kirsten, 2018, p. 127). Kirsten cites Alain Boillat, who examines Metz's terminology and states that,

> le concept de « diégèse» et les questions soulevées par sa définition prennent place parmi les préoccupations originelles de la sémiologie du cinéma, que la filmologie semble bien avoir « préfigurée» (Boillat, 2009, p. 229).

We translate: "the concept of diegesis and the questions underlined by its definition take their place among the original concerns of the semiotics of the cinema, which filmology seems to have prefigured well." Boillat therefore connects the ambiguities of the term diegesis to the concerns of both the filmologists and the development of cinematic semiology. What sets filmology apart, for Metz, is its status as being outside of film criticism and the cinematic institution, for, according to Metz, it "was the scientific study conducted from outside by psychologists, psychiatrists, aestheticians, sociologists, educators, and biologists" (Metz, 1974b, p. 90). Kirsten notes that this scientific distance provided Metz with "research methods and a technical vocabulary" that was outside of what filmmakers and film critics were using at the time (p. 131). Metz is therefore deeply influenced by the shift from mimesis to diegesis conducted by the filmologists, who Taylor asserts introduced the terms *diégèse* and *diégétique* around 1950 in the writings of Etienne and Anne Souriau (Taylor et al., 2007). Taylor states that the English terms *diegesis* and *diegetic* are translations from French and refer to "the spatial story worlds primarily of fictional texts/films" and further reflects that this "highly successful term…is really a misnomer", having therefore come to be associated with the represented story world (or mimesis) rather than its Greek origins, which was more in line with the storytelling, even if such storytelling is, according to Aristotle, mimetic. Stefano Castelvecchi, in his article "On 'Diegesis' and 'Diegetic': Words and Concepts", goes as far as to assert that the

filmologists were simply in error, "they were misconstruing its meaning by claiming that it could have something to do with the 'story,' the content of a narrative" (Castelvecchi, 2020, p. 20). Nonetheless, this was the seed of our current use of the term diegesis.

As we shall see, the filmologists' influence extended beyond the domain of film into literary theory and, as a result, literary theory has contributed both to the ambiguities and shift in meaning of the term *diegesis*, particularly via its use by Gérard Genette. This serves as an important elaboration because we will return to Genette's narrative theories in Chapter IV as a possible vehicle for exploring sound discourse in film. Castelvecchi devotes a detailed analysis to assert that Genette made a considerably important contribution to disseminating the filmologists' use of the term (Castelvecchi, 2020, pp. 14–22). Summarizing the argument, Castelvecchi notes that Genette uses the term in 1966 and 1969 in a similar manner to the Greeks,[21] even referring to Aristotle, and in line with centuries of literary theory. Yet in an essay first published in 1969,[22] Genette utilizes the filmological notion by defining diegesis as the "spacio-temporal universe to which the primary narrative refers." Castelvecchi further notes that by the time of Genette's *Discours du récit* (1972), Genette acknowledged in a footnote that this recent use of *diégèse* "'comes to us from the theorists of film narrative,' but he provided no reference for that claim" (Castelvecchi, 2020, p. 16). Moreover, throughout his works, Genette uses both versions of diegesis, that of the narrator and that of the story-world, depending on his writings. At one point he engages in an explanatory work that seeks to justify a parallelism using two separate terms with a minor variation in spelling, diégése and diégèse, which unfortunately translate into the same word in English. Ironically, according to Castelvecchi, Genette in his 2009 *Codicille* puts the final nail in the coffin, admitting to have abandoned the Greeks. Castelvecchi notes,

> Here, a near-octogenarian Genette wrote that he considered his contribution to neologism to have been greatly overestimated, and provided the example of the word "diégèse," which, after all, is "borrowed from the theorists of film" and "barely gallicizes the Greek diegesis" (Castelvecchi, 2020, p. 22).[23]

While both Taylor and Castelvecchi note that we have now moved beyond the filmologists'

[21] In Genette's essay "Frontières du Récit" (1966), reprinted in his book Figures II (1969)
[22] In Genette's Figures of Literary Discourse (1969)
[23] Castelvecchi cites Genette, "Diégèse, qui francise à peine le diégésis grec, est emprunté aux théoriciens du cinéma" (Genette, 2009, p. 194).

error and latched onto the *story world* as the defining element of diegesis, a further examination of Christian Metz's use of the term offers a broader concept than either of them assert, or rather, a concept they both skip in their examinations of the lineage over time. According to Kirsten (2018), Christian Metz himself first uses the term *diegetic* in 1965 and later elaborates in his article "Some Points in the Semiotics of The Cinema" published first in French in the 1966 edition of *Essais Sur La Signification au Cinéma* and in the 1974 English addition, *Film Language: A Semiotics of the Cinema* cited here:

> The concept of *diegesis* is as important for the film semiologist as the idea of art. […] It designates the film's *represented* instance […] – that is to say, the sum of a film's denotation: the narration itself, but also the fictional space and time dimensions implied in and by the narrative, and consequently the characters, the landscapes, the events, and other narrative elements, in so far as they are considered in their denoted aspect (Metz, 1974a, p. 98).[24]

Kirsten elaborates that Metz considers "diegesis as equivalent to everything denoted in the film" and "very close to Souriau's original concept" (Kirsten, 2018, p. 132). However, and most relevent to us, a *represented instance* can go beyond the story time and world, for other elements are representational as well. For instance, musical score can be representational of a time period or a character's emotion, denoting Metz's fictional time dimension or character attributes. And particularly with sound effects, a *represented instance* can be quite broad and the line between denotation and connotation can be fuzzy, especially when examining symbolic or indexical signifiers. For example, a sound of an unseen hawk screech in flight is, iconically, a hawk screech; it denotes the sound of a hawk and connotes a hawk in flight. Symbolically, it may connote a predator or a warning of death. However, indexically, if relating perhaps a dead body nearby depending on the context, it only denotes, for as Francesco Bellucci shares in his discussion of "Peirce on Symbols", an index "has denotation but lacks an adequate connotation" (Bellucci, 2021, p. np). We take issue with Bellucci here because our example, in fact, does both. One can see that with sound, and with this example, we illuminate further complications in applying non-dialogue sound to semiotics, in part due to sound's ability to engage simultaneously in iconic, symbolic and indexical signification, blurring the distinction between denotation and connotation. Furthermore, when sound effects join images together to support a

[24] We lean in strongly to the idea of *represented instance*, as will be discussed later.

unified interpretation of those images, is the resultant interpretation, meaning or theme denoted or connoted by the juxtaposition of sound and image? For example, imagine a sequence in which a protagonist and an antagonist are intercut in different outdoor spaces, both preparing for a confrontation. That hawk sound effect heard over both characters joins them, not in physical diegesis, but in a kind of thematic *represented instance*, forewarning death. The combined effect of image and sound, it can be argued, is at once symbolic and indexical; indexical because it is the juxtaposition itself that is the signifier and therefore denotes according to Peirce and Belluci. Or we can look at it as a syntactical arrangement of sound and image symbols, an arrangement that only connotes, since they are all symbols. Semiotics finds it challenging to organize this space where the sound-image in tandem generates meaning.

In the creation and discussion of our artistic research and again in our content analysis, we will see this phenomenon many times. To address it, we wish to use the term *thematic diegesis*[25] or at times *thematic continuity* when specifically referring to a series of shots or sequences, to represent a thematic part of a story-world that emerges from story via denoted meaning in the relation between image and sound. Back to Metz, we regard these thematic diegetic moments, as they unfold, as "represented instances", a sum of the film's denotated elements of sound and image. One last item within Metz's definition to address is that he cites the *narration itself*, as part of denotation, which leads us to wonder what of narration he includes, for perhaps it is more broad, and more akin to Genette, as we will see in Chapter IV. Therefore, is Metz's definition of diegesis perhaps more broad than where we have landed in the film school teachings? Does the idea of his *represented instance* open more possibilities when sound is brought into analysis? We shall explore further via case study throughout.

Regardless, it seems that we have institutionally drifted afield from this notion of the denotated or represented instance, and this drift is important with respect to our research because within sound theory, particularly as formulated by Chion, but also as has been ubiquitously taught in film studies and found in criticism, *diegesis* has fully adopted an oversimplified scaffolding. In film studies at the level taught in schools, *diegesis* is typically thought of as a construct that divides the "spatial story worlds primarily of fictional texts/films,"

[25] Chion uses the term "audio-visiogenic effects" to capture ideas like our thematic diegesis (meaning) and spatial diegesis (scenography of imaginary film space) along with other effects like temporality and rendering of material qualities. All fall under his general category of "added value" (Chion, 2009, p. 468). We interrogate these separately with our own terminology more tightly defined.

from everything else related to storytelling (Taylor et al., 2007). In other words, *diegetic* elements of cinema belong to the narrative or narrated world of the film, whereas *non-diegetic elements* are not and are therefore the storytelling means. One only needs to turn to Susan Hayward's ubiquitous entry-level guide on film theory and practice, *Cinema Studies, The Key Concepts*, first published in 1996 and now in its 5th edition, to note that a common definition reads confusingly and locks the construct to the screen, to the image. As she states,

> Diegesis refers to narration, the content of the narrative, the fictional world as described inside the story. In film it refers to all that is really going on on-screen, that is, to fictional reality. […] Hence the term diegetic **sound**, which is sound that 'naturally' occurs within the screen space (Hayward, 2000, pp. 84–85).

To begin with, syntactically it is unclear if Hayward equates "narration" with "narrative content". If so, perhaps this might align more with Metz and even Aristotle. But she goes on to point to "the fictional world described", which aligns with the more narrow perspective of the story world. Most importantly, Hayward ties diegesis to the screen, calling what is on-screen the "fictional reality". Everything seen "on-screen" in a traditional narrative film is certainly part of the world of the story, and fully representational, therefore mimetic. However, it is notable that this screen-based definition makes no mention of any offscreen presence that can be denoted by an identifiably-offscreen sound, or by any other acousmatic [26] sound. Additionally, we find it problematic to equate Hayward's "fictional world described" with Metz's "represented instance", especially with respect to sound. As mentioned, does musical score, for instance, not have the ability to denote? Is a voice-of-god voice-over not a represented instance? Neither one is part of Hayward's universe of fiction nor onscreen, yet both are Metz's represented instances, the latter even conforming to Aristotle's *diegetic mimesis*. And what about metaphorical use of sound, for instance a heartbeat sounding in a moment of tension as is found in our artistic output? It is not onscreen, not of the world, yet clearly a represented instance as a signifier; we might argue, and will surely later discuss, that an audible metaphorical heartbeat is not diegetic as an icon since no character can hear it, but it is diegetic as an index or symbol because it directly denotes – i.e. *heartbeat*, *anxiety*, or *life* – depending on the paradigmatic context. This process aligns with Metz, regardless of the particular signified

[26] Acousmatic: A term coined by Pierre Schaeffer, 1952. Pertaining to the auditory situation in which we hear sounds without seeing their cause or source (Chion, 2009, p. 465).

provided by the index or symbol. When examining this discourse around the evolution of the term diegesis and its implications, therefore, diegesis with sound seems to take on greater dimensions than is typically discussed in basic film courses.

We do find an expansion beyond the frame by Bordwell and Thompson, albeit restricted to narrative film, with a definition that "includes events that are presumed to have occurred and actions and spaces not shown onscreen" (Bordwell & Thompson, 2010, p. 491). Arguably, these additions are most frequently communicated to audiences via sound; a space or action we don't see can only come into the story-world if we hear it,[27] like audible nighttime crickets outside during an indoor scene, or a character who speaks, recounting a walk in the park taken earlier in the day that we do not see, or even a huge reverberation on sounds in a space to denote the inside of a cathedral when no such cathedral exists visibly in the shot. Offscreen diegesis, therefore, is the domain of sound, albeit with one exception. Very infrequently, we can have an image that Bordwell & Thompson call a "nondiegetic insert", one in which visual onscreen objects "are represented as being outside the world of the narrative" (p. 493). This is a rare event, done for metaphorical effect, and one of two "devices of discontinuity" cited by Bordwell & Thompson, the other being the jump cut. The nondiegetic insert is used, for instance, by Godard in *Breathless* (1960), and also frequently by Eisenstein. They note that these nondiegetic inserts "construct a running, often ironic, commentary on the action" and prompt the listener-viewer to construct implicit meaning.[28] And yet, to offer a comparison of how filmic non-diegesis is a concept mostly used for sound, Bordwell & Thompson offer only one short paragraph dedicated to the rare, visual, nondiegetic insert, whereas eight pages are dedicated to an in-depth discussion of the binary system of diegetic/nondiegetic sound (pp. 284-292), in which nondiegetic sound is the only alternative to diegetic sound (p. 284). They further elaborate on the binary system of onscreen vs. offscreen diegetic sound, and that of external vs. internal diegetic sound, highlighting that diegetic sound can come from within the mind – internal – of a character and not heard by other characters. At one point, they summarize that "sound may be diegetic (in the story world) or nondiegetic (outside the story world). If it is diegetic, it may be onscreen or offscreen, and internal

[27] A rare exception might be if the audience knows through another source, be it literature in the case of an adaptation, or through reading criticism of a film, about spaces, actions, or expositional elements prior to watching a film.

[28] Metz's narrative instance would include this in the diegesis, and we will see later that Genette also does not make such a distinction, finding metaphor to be part of narrative, or diegesis.

(subjective) or external (objective)" (p. 291). Bordwell & Thompson have liberated the notion of diegesis beyond Hayward's screen, but still anchored it into the story world. Back to Taylor, he finally notes that while he takes issue with the term on historical grounds – asserting that *mimesis* would be a better term than *diegesis* – he notes that diegesis is "well established" and "particularly useful in designating aspects and features of filmic sound as it relates to the relatively closed story-worlds of fiction" (Taylor et al., 2007). Thus, he resigns himself to a utilitarian reason to stick with the binary formulation despite his own reservations from the term being a "misnomer".

If we are to accept the term *diegesis*, there are those who challenge its binary construction, pointing to the in-between or sounds' ability to transverse a diegetic boundary, and there are those who lean into it. Returning to, *Discourses on Diegesis* (Taylor et al., 2007), other academics and sound designers/mixers weigh in on the history and debate over the functionality of the binary construct. Professor Martin Nordan at University of Massachusetts-Amherst finds "critical and pedagogical value in maintaining the distinction between diegetic and non-diegetic sound, as blurry as that line often is." For him diegesis works as a tool, again utilitarian, but brings challenges because "assumptions about these two broad variants of sound usage sometimes interfere with our understanding of what filmmakers can actually do with them."[29] There is at least, in Nordan, an acknowledgement, therefore, that sound *functions* at the will of filmmakers in a way that diegesis does not capture. We regard this as a step in the right direction and it informs our artistic research and the creation of our new audiovisual text, as well as the development in Chapter II of a way to define and test certain functions rooted in filmmakers' intentions. It is important to note that neither Taylor nor Nordan are sound specialists in their academic disciplines, but rather are film and film-studies professors in Europe and North America. Bridging into sound specialists, into the artists in the field, we find even more concern around the binary diegetic construction.

Mark Kerins, who teaches film, writes about film, and also does some sound work for film notes that the diegetic/nondiegetic "distinction has some value" but warns that focusing on it as a determining factor of sound has "overshadowed the question…of how diegetic sounds are used." He is largely concerned here with spatiality and immersion.

[29] We will refer back to this important point in our discussion of desired outcomes of filmmakers in chapter two.

> This is not to diminish the use value of the diegetic / non-diegetic distinction, but to point out that today there is more to the relationship than mere nomenclature. In a multi-channel world, many (though not all) films are willing to let diegetic sounds spread out into the theater to create a more "complete" space (Taylor et al., 2007).

Kerins does not here expand upon the implications of what constitutes a "complete space" nor how that alters or influences reception of narrative, emotion or other communicative aspects. He is, however, hinting at a certain fallacy of the notion of diegetic sound *tied to the screen* in a viewing/listening world where sound comes from offscreen sources and offers an embodied experience. This will be further explored in this thesis, especially with respect to embodiment and phenomenology, but in summary here, Kerins speaks to two additional dimensions - immersion and the real of sound - where the dichotomic construction of diegesis undermines functional analysis.

Continuing in the *Discourses on Diegesis* forum, Randolph Jordan, an academic and acoustic ecologist who crosses the realms of film and sound theory, discusses the "hearability" of a diegetic sound by the characters in a film. We find that in contrast to Kerins' concerns, Jordan concludes that the term diegesis itself, and the binary distinction therefore, is necessary for the discussion to be had (Taylor et al., 2007). He sees it, therefore, as a requisite tool in the hands of those analyzing film sound.[30] Jordan is followed by Randy Thom, one of the most widely-awarded sound designers and mixers in Hollywood. Quite forcefully, he notes that none of his colleagues have ever used the term "except to deride it as an academic term of little practical use" (Taylor et al., 2007). In general, this is certainly verifiable from my personal experience overall as a mixer. I, however, must add that many times in mixes I have heard the term come up and, in fact, it is while mixing that I first heard it. There are directors and editors who are versed in the academics, either through film school or self-taught, and do use the terms. Yet I agree with Randy Thom's conclusion, where he confesses that he has "another, more radical suspicion: I think the question of whether a sound in a given scene is diegetic or not is often irrelevant to the effect the story has on its audience" (Taylor et al., 2007). Once again, the practitioners lean into the idea of the "effect" as the dominating reason for making sonic choices, further supporting our artistic research and content analysis design that seek to

[30] At times, we will use "diegetic" or "nondiegetic" in this paper for the purpose of brevity, simply meaning, "of the story world" or "not of the story world".

illustrate some of these effects and expose correlation between sound choices and outcomes.

Perhaps this speaks to Castelvecchi's final ruminations. He takes issue with the use of the term *diegesis* altogether. He acknowledges that the term is used by a tiny minority of the speakers of natural languages such as English or French and is, rather, of a specialized language that, as such, not only can be described but also proscribed. To those who defend it based on utilitarian grounds as an established practice, he cautions,

> It is the practice of scholarly communities to continually scrutinize and refine their terminologies, thus refining the attendant conceptual apparatus (Castelvecchi, 2020, p. 25).

It is in the spirit of "refining the conceptual apparatus" the we continue on our journey here. The question of relevance of the term diegesis and the diegetic sound construct is something that is further explored throughout the course of this report and even embedded in our artistic output and conclusions drawn. But before we can further challenge it, we must examine the most well-structured look at sound diegesis that is foundational to modern film sound studies, and this of course takes us back to Michel Chion.

As noted earlier by Walter Murch, Michel Chion's œuvre represented, by 1994, the overwhelming body of theoretical work in film sound (Chion, 1994, p. ix). More than a decade later, in the introduction to *Lowering The Boom: Critical Studies in Film Sound* (Beck & Grajeda, 2008), Chion[31] is credited with helping "to define the field of film sound studies" (p.20) via his 1994 publication Audio-Vision , the objective of which was "to demonstrate the reality of audiovisual combination—that one perception influences the other and transforms it" (Chion, 1994, p. xxvi). Only nine years later, this volume was followed by Chion's 2003 work, *Film: A Sound Art* which spans over 500 pages and further advances the field. Beginning in the 1980s, Michel Chion became interested in examining the diegetic space. But we must ask, what definition of diegesis forms the basis of his examinations? Chion offers no such definition in 1994, rather, he only refers to the use of the term nondiegetic "to designate sound whose supposed source is not only absent from the image but is also external to the story world" (Chion, 1994, p. 73). But in the glossary of his 2009 English-language version of *Film, A Sound Art* he gives a clear definition of *diegetic sound* as follows:

[31] Along with Rick Altman and Weis & Belton

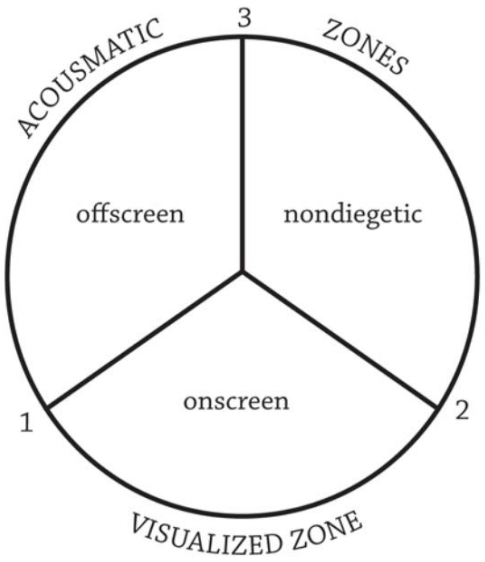
> Sound whose apparent source is in the space-time continuum of the scene onscreen. Diegetic sound is sound that the film leads us to believe the characters can (or could) hear (Chion, 2009, p. 474).

As with Hayward cited earlier, while we can go beyond the screen, we are tied to it in that the screen defines the possibilities of diegesis. Furthermore, Chion provides the following definition for *nondiegetic sound*:

> In the audiovisual relation, sound whose apparent source is not the space-time of the scene depicted onscreen. The most frequently encountered examples of nondiegetic sounds are voice-over narrators or commentators and musical accompaniment (p. 480).

In both these definitions, Chion leans strongly into the "onscreen" scene as rooting diegesis. Thus, we begin with a reductive construct tied to image and more in line with the textbooks of Hayward or Bordwell & Thompson than with the ruminations of Metz, the filmologists, Genette, or even Aristotle. We will address Chion's position later in the subchapter, but first, let us look at his structured theory with an understanding that, for him, diegesis is defined as a story world, one which is *visually determined*.

In his early work, Chion was concerned with bringing to light the ways in which sound either crosses diegetic space between the world of the characters and outside that world, crosses between the on-camera and off-camera diegetic space, or lives somewhere that is difficult to define along diegetic lines. In these "three borders," as he calls them, sound holds most of its power as a storytelling rhetorical devise. Let us examine Michel Chion's "tricircle" which has become a staple in all pedagogy of film sound theory. It was first published in an edition of *Cahiers du Cinéma* (Chion, 1985) and later appeared as a part of a chapter entitled "The Audio Visual Scene" in his book *Audio-Vision* (Chion, 1994, pp. 66–94), where it is introduced along with a number of sound concepts. Fifteen years later, the tricircle and diegetic theory form the content of an entire, albeit shorter, chapter entitled "The Three Borders" in his book *Film, A Sound Art* (Chion, 2009, pp. 247–261). Below is Chion's "tricircle," the structure of which does not change across these publications through the years.



1. Onscreen/Offscreen border
2. Onscreen/nondiegetic border
3. Offscreen/nondiegetic border

Figure 3: Chion's Tricircle (Chion, 2009, p 251)

To summarize Chion's tricircle, he has categorized sounds as living in one of three spaces. First, with synchresis32 to image, a sound is in the "visualized zone" at the bottom of the circle: onscreen and diegetic. For example, a vacuum cleaner whines as we see it being used. To further expand upon this, even if we hear dialogue from the center speaker, for instance, yet spoken by a character far to one side, we perceive that dialogue to come from that character's mouth. This, he calls "magnetization" which is in this case a function of synchresis (Chion, 1994, pp. 69–71) and speaks to the strength of our willingness to overlook directionality when presented with visual sync. Then, in the upper 2/3 of the pie, we have two zones that are not visualized, but rather "acousmatic", a term frequently used by Chion but, as he notes, coined by Pierre Schaeffer in 1952, simply meaning sounds that can be heard "without seeing their cause or source" (p. 465). On the left and within the acousmatic area is our second zone, offscreen diegetic sound, like that same vacuum cleaner but heard from a shot in another room, or that same dialogue but with the character offscreen, in other words with no synchresis offered. On the right is our third area, the only place for nondiegetic sound,

[32] Perception of a visual event and sound event as happening concurrently, where the visual action appears to cause the sound, due to placing them synchronously (Chion, 2009, p 492).

like the triumphant music score for Rocky as he runs up the steps of the Philadelphia Art Museum, for certainly there is no 50-piece orchestra playing on the steps (Avildsen, 1976). With this triplet of positions, we can begin to assign to sonic signifiers or sound objects a position with respect to the image. But what interests Chion the most are the borders between these spaces, as suggested by the title of the chapter: "The Three Borders" (Chion, 2009, pp. 247–261). Here Chion describes how sounds pass from one zone of the circle into another. For instance, music that begins as "source" music, like someone on screen turning on a stereo, can be played first looking at the stereo speaker rendering it onscreen diegetic, then continue as the camera follows the character into another room to become offscreen diegetic, and continue further over the next scene as nondiegetic musical score. Thus it can cross all three borders and frequently this shift is a function of mixing – the frequency and spatial treatment of the music and the locational placement in a speaker array – which allow the shift to work.

Tackling the range of diegetic relationships between sound and image is a formidable task and Chion's tricircle goes very far to establish a formalistic mechanism for evaluating sound, albeit locked to an image paradigm, and we will give a few such examples here. With regard to the border between onscreen sound and nondiegetic sound, Chion states that this crossing is rarely done beyond the common source/score music trick when a song moves from being "sourced" in the scene to becoming score, or vice versa, or the switch of voice-over from narrative to onscreen. We find this to be much more common than Chion states. One example of the voice-over switch can be found with virtually any talking-head or witness-based documentary film, for instance in Spike Lee's Emmy-award winning series about Hurricane Katrina, *When The Levees Broke* (Lee, 2006). Here, interview subjects speak while, at times, the filmmakers use associative editing to show either the damage of which they speak or another visual idea that complements their testimony. As the image lands on the speaker, the force of either witnessing or victimization is strengthened. Other documentaries use this technique around archival footage, enunciating what we see via the offscreen talking head and landing on the speaker visually as an "expert" with a title, to solidify the testimony as coming from a trusted source who appears to be speaking to us, as in the film *Finding Babel* about the life of Soviet writer Isaac Babel, directed by the author of this report (Novack, 2015). Here, like with music shifting between source and score, Chion's border maintains its integrity for the words are either onscreen or nondiegetic, never somewhere else. However, and tellingly, both these examples expose the hegemony of image-based thinking as

they presume the original disembodied voice to be nondiegetic, and therefore diegesis is completely determined visually. Imagine the opposite: that it is the audio testimony that is diegetic, especially since it is non-representational sound, and therefore the associative visual shots used to illustrate would be nondiegetic images, those "very rare" things to which Bordwell & Thompson earlier referred, and which are actually not rare at all in the documentary film context, and which ultimately cut to the diegetic shot of the speaker. The traditional application of the vocabulary is, again, visually-based but here we offer a subversion, a new perspective that places sound in the driver's seat.

At the border between onscreen sound and offscreen diegetic sound, numerous examples are easily identified. For instance, if we hear a car approaching before it appears on screen, and then as it passes off screen we gradually hear it fade away, we have crossed this border. Similarly for music, we can see a piano player in a cocktail lounge, and then cut to a conversation at the table with the piano music continuing, crossing the border from onscreen diegetic sound to offscreen diegetic sound, frequently sold to us through a change in perspective from mixing techniques.

At the border between nondiegetic and offscreen diegetic sound, one example is a common treatment of music, as in the opening sequence of *The Shawshank Redemption* (Darabont, 1994). Here we are initially presented with a jazzy musical nondiegetic score, experienced as nondiegetic due to full volume and rich equalization, but as the camera slowly pushes into a stationary car with the protagonist sitting at the wheel the sonic treatment of the music gradually changes to that of an old mono radio, becoming diegetic/source music, onscreen when we see the radio, but offscreen when we don't. Thus, it crosses all three of Chion's borders. This is a very common technique with music, a corollary being when source music in one scene, like from a live band or a stereo, continues full and rich as musical score in the following scene.

However, there are areas of sound selection, editing and mixing that are not satisfied by this model, some notions that Chion misses. We will provide some examples below and others will emerge from our artistic research and case study of the *Koyaanisqatsi* sound design later in this chapter. These areas of departure from the model form a major motivation for the new paradigms of discourse that we seek in this thesis. But first, let us discuss Chion's own ruminations on the possibility of exceptions.

In 1994 when Chion first presents, in book form, his three-zoned diegetic construction

of the tricircle, he is already cognizant of criticisms but strongly defends the binary construct even while allowing room for creating new categories (Chion, 1994, pp. 74–75). He points to a few exceptions that seem to us to be rather trivial, like a character's voice when her back is to us – is this onscreen or offscreen? – or an electronic sound source – is the radio or telephone sound really onscreen just because we see the object that coveys it? He further cites an adult voice articulating the thoughts of an infant as in Amy Heckerling's *Look Who's Talking* (Heckerling, 1989), for can that possibly be the diegetic thoughts of the baby? Lastly, he cites the common practice of background environmental sounds which we never see. Here Chion notes that it "seems rather ridiculous to characterize them as offscreen, on the basis that we don't 'see' the little birds chirping or the wind blowing" (p. 75). This particular exception we find odd, as it seems to presume that we must at some point "see" the source of a sound in order to consider it *offscreen diegetic sound*, a contradiction with his own definition of *offscreen sound* just a page earlier, "in the narrow sense offscreen sound in film is sound that is acousmatic, relative to what is shown in the shot: sound whose source is invisible, whether temporarily or not." We find it curious that the examples of exceptions provided in his 1994 work are rather trivial relative to the kinds of exceptions that we will later discuss, which are quite common and more impactful. Furthermore, Chion defends, stating that "anyone who brings up such exceptions in order to claim the categories useless or trivial is throwing out the baby with the bathwater" (Chion, 1994, p. 75). Yet in 2003, when Chion re-presents his tricircle, he makes no mention of any critiques or ambiguities. Rather, his tricircle is presented as film sound theory de facto.[33] This harkens back to Castelvecchio's earlier-referenced thoughts on the need for theorists to scrutinize and refine terminologies and the "attendant conceptual apparatus" (Castelvecchi, 2020, p. 25). Castelvecchi might argue that Chion has pushed aside the hard work of furthering the conceptual apparatus. Let us now examine Chion's border zones, not discussed in his exceptions but elaborated upon in the final section of his chapter (Chion, 2009, pp. 259–260). Here we interrogate the validity of these borders and their crossing, an examination that will inform our artistic research and case study of *Koyaanisqatsi*.

In the border between onscreen and offscreen diegetic sound, Chion describes "the mentally visualizable trace" of an image with a concretized sound that is carried by the sound

[33] We ask ourselves if this was a reflection of how widely accepted his theory had become from 1994 to 2008, so much so that he no longer felt the need to address critique? Or was it more of an authorial strategy, that raising the question undermines the message? Of course we cannot know.

as the image moves offscreen, the sound being acousmatized. Moreover, movement in the opposite direction, sound becoming visualized or deacousmatized, is said to carry with it a "not-yet-revealed quality" of its source, before landing on the source. Though true in some cases, there are many in which an offscreen sound effect or music provides counterpoint that is not *of the source* to be later identified, until the moment of synchresis, of magnetization to the image. As an example, we address an often-studied early sequence of *Apocalypse Now* (Francis Ford Coppola, 1979), a sound effect of a helicopter is heard. The visual "trace" or "not-yet-revealed quality" is that of a helicopter, even if not seen. What one lands on visually are the rotating blades of a ceiling fan. Synchresis assigns the signifier to the ceiling fan, even though the sound is a helicopter. We then cut to a characters face, looking up at the ceiling fan. But before we land on the character's face, there is no sense that perhaps this is an internal sound of the character. On the shot of the fan, can this sound be diegetic if we know it is a mismatch between sound and image? As a form of narrative counterpoint that enters into discourse as will be addressed later, can this sound rather function as nondiegetic score, evoking a response that is outside the world of the character? The answer is unclear. Chion continues to describe the nondiegetic-offscreen border as the most mysterious, and here we must concur. He describes this border as being the least tangible and states that "opening this border, which amounts to its loss, is just about the most poetically fateful thing one can do in the cinema"[34] (p. 250). Later in this chapter, in our case study of our artistic version of *Koyaanisqatsi*, we will illustrate a number of examples that highlight and support Chion's assertion. But meanwhile, as a sound designer and mixer, let me provide a theoretical case that illustrates this poetic effect with a concrete example. For instance, when offscreen nighttime crickets and frogs are manipulated in the mix via volume and equalization changes to function as tension-inducing score, following, if not actively promoting, the emotional arc of a scene, this sound is dancing between diegetic and nondiegetic space. It is of the story world defined by the screen, but not mixed realistically and therefore functions more like musical score. In fact, especially with immersive surround sound, this dynamic change in the sound can rupture the territory of the screen, obliterate it, and therefore contribute to the phenomenological and embodied experience of the listener-viewer. We previously cited Mark Kerins, in Taylor (Taylor et al., 2007), noting the

[34] Michel Fano refers to congruence and non-congruence between what is heard and what is seen as the least understood relationship between image and sound and, presumably the most powerful (Fano, 1987, p. 6). Chion's nondiegetic-offscreen border seems to embody the same space as Fano's.

curious way in which immersion, or sound coming from multiple sound sources in a theater, defies traditional notions of diegesis. One must ask, therefore, if this border really exists and if Chion seems to point out a weakness in his own theory. As previously mentioned, after criticism among academic circles for the over-simplicity of the original model, Chion revises it in 1994 (Chion, 1994, p. 78) albeit defensively. In this new model, shown below in Figure 4, he removes the visualized and acousmatic labels, adds an "on-the-air" arc to encompass things like telephones and radios, includes a space near the intersection of all three zones for "internal" sound, and places ambient sound along the full border length of "on-the-air" and nondiegetic.

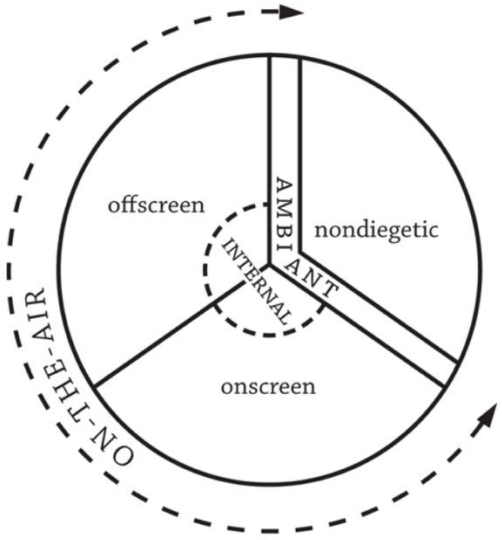


Figure 4: Chion's updated Tri-Circle (Chion, 1994, p. 78)

The revision does satisfy some anomalies of the first model, even the exceptions that he mentioned and we earlier noted. For instance, he defines "internal sound" to be sound related to both physical sounds like breaths or heartbeats – *objective-internal sounds* – and mental sounds like memories or mental voices – *subjective-internal sounds* – of a character. Our previous example of the helicopter SFX in *Apocalypse Now* would, once the image lands on the character, be considered a subjective-internal sound, a memory, triggered in our character by the visual of the ceiling fan. But before we see his face, when we have only seen the ceiling fan blades, we are in a liminal space where the sound functions entirely differently. What mixers know, our praxical knowledge applied, is that this type of treatment defies traditional diegetic space or, as cited earlier by mixer Randy Thom, the diegetic placement is "irrelevant to the effect the story has on its audience" (Taylor et al., 2007). The effect is altered by the positioning of sound; if the transition were silent, we would still understand the visual metaphor provided

by the edits or by a dissolve from helicopter to ceiling fan to a person staring at the fan, in increments. But with sound, we are taken beyond metaphor to something more experiential, phenomenological, enabled by the liminality offered by gradually-transitioning sound for which, in the middle shot of the fan, we are yet to have a concrete reason for the mismatch in sound.[35] Thus, where Chion points to crossing a border that separates two discrete places – from diegetic onscreen sound, like the helicopters we see and hear, to subjective-internal sound, like the helicopters we hear over an indoor close-up of the protagonist – he ignores the fact that along the way, we are in neither of those positions. This is but one example where function defies diegetic construction.

Additionally, in both the original and updated versions of his diegetic theory, Chion avoids mention of mixing techniques and how they may impact the diegetic analysis, including reverb, equalization, dramatic or irregular level adjustments, and the paradigmatic choice of excluding a sound for which there is a visual signifier. For instance, when reverb is added to a diegetic sound the reverb becomes quite literally part of the new sound. Yet, if the visual space does not warrant reverb, is the reverb perhaps a nondiegetic component of the sound signifier, even if the raw sound itself is diegetic? There are many such instances where a sound can be simultaneously diegetic and nondiegetic due to mixing techniques, quite powerfully, and even the above-mentioned helicopter on the ceiling fan is one such moment. This simultaneity is not addressed in Chion's tricircle and, in fact, it may challenge the construct because nowhere does he assert that the construct allows for simultaneous positioning of a sound in two or more areas of the tricircle.

Another example with respect to mixing is evident when we examine volume. How do we regard the omission of a sound within Chion's construct? If for instance we don't hear a massive explosion of a car that we do see, when we know our protagonist's child is in the trunk and she watches from afar, arguably the silence itself is nondiegetic since in the world of the film we should hear it. It certainly has a nondiegetic sensibility, like musical score, impacting or highlighting emotion. Or, if such silence is considered to be an absence of a sonic signifier, then nothing is being denoted by that silence and, therefore, it has no placement in the tricircle nor in Metz's diegesis. And yet, is not something connoted? One can argue that there is a represented instance by the removal of that explosion, one that functions symbolically because,

[35] We will introduce phenomenology in greater detail later in this chapter.

as an accepted part of film language, as part of the code, it communicates deep trauma felt by the horrified mother, and subjectivizes it vis-à-vis the listener-viewer. From that perspective, it must live somewhere in the construct, but where? The answer is unclear, highlighting another failing of the model with respect to mixing techniques. Aristotle might argue differently, if only we could invite him to a screening. For him, it seems the scene would be diegetic mimesis, and the silence from the removal of sound *is* the storytelling or narrating act and therefore a diegetic act. And for Genette, as we will see in Chapter IV, the diegesis of the silence would be irrelevant but would rather fall under his umbrella of narrative.[36] Suffice to say, Chion's tricircle takes film sound and shoehorns it into an agreed-upon notion of diegesis that is problematic across film and literary scholarship. It is a formative structure that certainly helps with sonic analysis as it relates to image in film, however it seems incomplete and, at times, feels like an attempt to frame sound theory *inside* image theory, particularly due to the ways diegesis is tied to the screen, as previously examined by us. This penchant to frame sound predominantly in a subordinate relationship to image can be found elsewhere in Chion, although through some prisms he offers greatly expanded insights as we shall explore in our examination of rhetoric to follow.

We have, thus far, defined sound design and positioned Chion in the canon of film theory, laying out his theories of sound diegesis, but first engaging in a discussion of the evolution of the term and some of the ambiguities it embodies, particularly the penchant with which film theory ties diegesis to story and the screen. We've identified a few conceptual examples that Chion does not address but are relevant for our research in that it is precisely these types of instances, which are quite common, that form the motivation for our artistic exploratory research. We now move on to a discussion of sound rhetoric, examining Chion's rhetorical theory, another important analytical framing device he developed. Rhetoric articulates with diegesis in its further attachment to the screen, and yet it provides some insights into functions of sound in film that reach beyond diegesis. For us, and somewhat ironically, it is the one structure Chion provides that strongly hints at a richer and deeper discursive role for sound in film, a fact that deeply informs our eventual adoption in Chapter IV of Genette's *Narrative Discourse* as a potential model for film sound analysis.

[36] As we will discuss later, Deleuze might suggest that the silence helps us abandon self into new abstraction of sensation devoid of memory while Merleau-Ponty might suggest the silence positions us to experience, but still grounded as a spectator.

1.3.2 SOUND RHETORIC

> I would classify any attempt on the part of a filmmaker to influence our thoughts and feelings as an "argument" (Behrens, 1979, p. 3).
>
> Laurence Behrens, "The Argument in Film: Applying Rhetorical Theory to Film Criticism"

Here, we enter first into a discussion of the notion of rhetoric itself, precisely due to Chion's use of the term in his chapter "The Five Powers", where he posits these powers via what he calls the "five rhetorical figures of the said/shown relationship" (Chion, 2009, pp. 385–404). We will then examine Chion's theory in detail as he presents it, coming to an understanding of both strengths and weaknesses. This further prepares us with the theoretical tools that we will apply to artistic research later in this chapter, to interrogate the prowess of existing theories, and we will refer to these concepts throughout the thesis.

Where does rhetoric fit within film analysis? As defined by Aristotle, rhetoric is the art of "discovering…the available means of persuasion" (Cooper, 1988, p. 3). This notion of persuasion persists as the most common understanding of the term today, as illustrated in the Cambridge Dictionary's primary definition: "speech or writing intended to be effective and influence people" (Cambridge Dictionary, n.d.). Persuasion is not often thought of as the goal of most fictional films and, within film criticism, Lawrence Behrens, cited above and in his article "The Argument in Film: Applying Rhetorical Theory to Film Criticism", notes in 1979 that rhetorical theory had rarely been applied systematically to film. He hints that some might find its use appropriate only in examining didactic films, non-fiction films, or films that are clearly tools of propaganda or persuasion, citing as an example *Reefer Madness* (Gasnier, 1936), the 1936 anti-cannabis propaganda fictional film (Behrens, 1979, p. 3). However, Behrens argues that all fictional films are not only "dramatic constructs" but "rhetorical as well" and, as cited earlier, finds that the filmmaker's job is one of argument, inasmuch as they "influence our thoughts and feelings" (p. 3). Behrens goes on to apply the classical rhetorical categories of logos, pathos and ethos to film criticism, mostly aligning them to verisimilitude, emotional play, and auteurship respectively (pp. 4-8). He further provisions us with an organizing device used in modern rhetorical and discourse theory, the *communication*

triangle,[37] built on definable rhetorical categories, but further notes that "it's less important to be able to categorize a film than to be able to view it as a function of all of these dynamics" (p. 9). It is in this play of dynamics that he conjures the *rhetorical stance*, citing Wayne Booth (Booth, 1963, p. 141). Booth depicts the *rhetorical stance* as the balance among three elements of communication: the "available arguments" of a subject; the assumed "interests and peculiarities" of the audience; and "voice", defined as the implied character of the speaker, further noting that this balance is difficult to describe. What we see in both Booth and Behrens is an acknowledgement that classical rhetoric, adapted, can play a role in analysis, yet it is not categorization where they find insights, but rather in the "stance" among the categories.[38] Moreover, Behrens specifically finds application to film: equating Booth's "available arguments" to logos, defined in film by those dramatic choices that render the narrative convincing and absorbing; equating Booth's assumed "interests of the audience" to the function of pathos in film, elaborating with film's ability to bring emotion to an audience; and equating Booth's "voice" to ethos, the implied character of the filmmaker. Behrens cites Booth once again, stating that "'the balance is harder to describe than the clumsy poses that result when it is destroyed'" (Behrens, 1979, p. 10).

To the extent that, with sound design, we intend to impart meaning through rhetorical structures, we are forced first "to reconstruct rhetoric as a cognitive theory" which, according to Dascal and Gross's "The Marriage of Pragmatics and Rhetoric", requires us to "place inference at its center" (Dascal & Gross, 1999, p. 108). If the choice of available arguments implies an act of communication, Grice's pragmatics assert that the interaction between speaker and hearer, the *talk exchange*, is scaffolded by the notion that "understanding is a matter of inference from what is said to what is meant" (p. 109). We add that, even though we tend to shy away from thinking of film rhetoric as persuasive and have already made an argument for the filmmakers' *argument*, it is worth noting that persuasion, or rhetorical exchange, is considered by Dascal and Gross to be a special case of Grice's talk exchange, in which there exists an "initial agreement" of the recipient to be open to persuasion, which should constitute

[37] This triangle is comprised of the expressive (encoder; artist-oriented), Pragmatic/Persuasive (decoder; audience-oriented) and referential/mimetic (reality; universe-oriented). Embodied within the triangle is the Formal/Literary/Objective (signal; artifact-oriented).

[38] It is not a surprise that this position on rhetorical stance may be fruitful, resembling in some ways Chion's tricircle, and it will therefore resurface in our case study work.

the pragmatic *cooperative principle* for Grice.[39] The speaker's introductory statements, "designed to obtain the attention, interest, and trust of the audience" (p. 109), as for classical persuasive rhetorical argument, expose an interesting analogue for film reflected in studio distribution logos where the sound often starts before image to obtain attention, which move on to opening sequences that are created for maximum interest, and ultimately the establishment of a reliable protagonist. Moreover, our audiences are either in an information exchange, where cooperation is presupposed, or in a persuasive exchange, where cooperation needs to be established and maintained. We argue that audiences can be in either of these positions, which can shift over the course of a scene or film. But overwhelmingly, film's organizing principles of continuity and fidelity, the erasure of the apparatus of classical film and even most documentary, seems to play a large role in *maintaining* the cooperation and therefore can be seen entirely as rhetorical. In addition, meaning-making with rhetoric includes the "possibility of misdirection and deception" (p. 110), often found in dramatic film. Dascal and Gross note that Searle addresses this deception with respect to utterances that have the intention of inducing a false belief, or that the recipient unknowingly infers a lie. We find that with every sound design choice and execution, from selection to editing to mixing, an utterance that has inference at its heart has been made. However, such inference can go beyond explicit meaning and lie in pathos and ethos, as "inferences to particular beliefs are a necessary condition of particular emotional states" with a clear "relationship of emotion to reasoned argumentation" (p. 120). Moreover, style can equally be a part of cognitive and rhetorical communication, when it is "a set of semantic and syntactic variants within a register" that "must depend for its effect on inference" (p. 122). Thus, sound tropes and immersive mixing techniques that are often dismissed as style but rather encase inference, are shown here to be part of a cognitive rhetoric. Dascal and Gross conclude asserting that rather than a hodgepodge of techniques, rhetoric can be rethought as "a coherent theory in the cognitive class" and "genuine hermeneutics" (p. 129) whereby for us, all utterances of sound in film, even abstract ones, imply inference, argument, and meaning.

From this discourse, we come to accept Chion's use of the term rhetoric in defining

[39] The cooperative principle is the assumption that in a conversation each participant engages in "the accepted purpose or direction of the talk exchange" (Grice, 1975, p. 45).

his Five Powers, not from the sense of rhetoric as political persuasion, but more in line with the idea of argument, an argument that demands an ever-shifting balance between logos, pathos and ethos, or between verisimilitude, emotional play, and voice, with utterance, inference, and meaning. Where sound fits in this paradigm will emerge in part in our following discussion of Chion's rhetorical powers of sound and our artistic research, and will be much further elaborated in case studies that follow, content analysis, and our application of narrative discourse. But for now, let us elaborate on Chion's rhetorical figures.

In his chapter "The Five Powers" in *Film: A Sound Art*, (Chion, 2009, pp. 385–404), Michel Chion advances a theory in which sound in film has five powers that reside in five rhetorical figures. These figures lie in the relationship between what is "said" and what is "shown". As we will show, it is here where Chion seems to address some of the ambiguities and omissions of the tricircle and where, more than anywhere else in his writings, he reaches towards rhetoric and narrative discourse as analytical tools. Still, there are strong limitations to this theory and the way he discusses it, even while it opens a gate towards expanded thinking. To set up his discourse, Chion gives us the following definitions:

> I call the said in a film that which belongs to the verbal sphere (words read or heard, uttered by a diegetic or heterodiegetic voice); the shown consists of the seen and concretely heard elements and, beyond framing, coloring, or staging the said, is the very raw material of cinema (p. 386).

Moreover, in a footnote he clarifies that "the sound of the voice, the material of that voice, belongs to the shown—shown sound—no less than the other concrete objects of a film" (p. 404). To clarify, only words themselves are *said*, not their characteristic sound, which is *shown*. With this conceit, sound effects would also lie in the realm of *the shown* to the extent that they are "concretely heard". We note a few assumptions at the outset that are problematic. Chion grounds the *said* in words, and words are tied to the screen. Omitting the case of a nondiegetic narrator for the moment, dialogue always comes from the screen. Even when a character speaks offscreen, their presence with respect to the visual frame is extrapolated by the viewer, usually from a previous or later shot, and laid out according to the geometry of the space as defined by camera positioning, shot size, and montage. The construct is therefore somewhat visually determined on both sides of the rhetorical divide: *the said* is grounded to the screen while *the*

shown includes everything on screen visually but extends to what is concretely heard. What does he mean by concretely? Perhaps he means an identifiable sound, one that can trigger a phantom image. Or does he include, for instance, musical score or abstract acousmatic sound? This is unclear, but in a chapter filled with examples, he never includes unidentified sounds, and only discusses musical songs with lyrics, which therefore embody a kind of dialogue or *said*. Therefore we can assume that musical score and abstract sound do not accommodate Chion's rhetorical *shown*. With respect to offscreen sounds, Metz might take issue with the term *shown* to describe them for he notes, with some disdain, the persistent cultural purview of, "the conception of sound as an attribute, as a nonobject, and therefore the tendency to neglect its own characteristics in favor of those of its corresponding 'substance,' which in this case is the visible object, which has emitted the sound" (Metz, 1985, p. 158). We find here, even in Metz, a shared skepticism towards visually-determined sound theory.

Nevertheless, let us examine Chion's "five rhetorical figures of the said/shown relationship" with brief mentions of case studies, where some of the ambiguities and omissions of the tricircle find stronger resonance. Note that he often describes these figures rather than providing definitions.

Scansion – Here Chion puts forth a relation between dialogue (*the said*) and visuals+concrete sounds (*the shown*) in which the rhythm and punctuation of speech are "scanned" via an event. As he notes, pro-filmic events can include the sound of wind or a bird, or the image of a person passing by, or the image of the speaker taking a drag of a cigarette. Filmic discursive events might consist of a camera put into motion, or the change of a quality of sound through filtering/equalization. An interesting example Chion cites is the gradually increasing reverb as Laurence Olivier repeats the line "except my life" while walking away in Hamlet (Olivier, 1948). This example illustrates Chion's supposition that the dialogue *words* themselves fall under what is *said*, while the sonic treatment of the words constitutes what is *shown*. As he notes, scansion is very common, and through it we can begin to see that sound treatment (or mixing) and sound whose diegetic placement is difficult to position in the tricircle have a home in Chion's *shown*, as long as such sound is "concrete". Chion's term *concrete*, however, remains ambiguous. Does he mean identifiable in terms of a source - the source of the above-mentioned reverb perhaps being a distant sound-reflective space into which the character walks? Does he mean separable as a distinctly heard sound even if not identifiable - we hear reverb, what causes it is not important? Or does he mean structured as in composition

through montage of sound objects, like in musique concrète,[40] - the reverb is a sound object, or part of one anyway, and has been put there through mixing and montage. He notes that the echo is a "concrete… wordless signifier". It certainly can signify a retreat into distance, a tunnel, or a cathedral, but also a memory. Chion goes on to note that scansion is an effect of the signifier. It need not have any meaning at all. This seems to contrast the example he gives since indeed "concrete" implies some identification of a sound, at least as an icon like when we identify the sound as *reverb*, but even as symbol where we sense the sound tells us it is a *memory*, or even as an index where due to the sound we recognize a retreat into the distance. The lack of clarity here begs the reader to wonder what sounds does he believe are not concrete, and where do they fall in terms of scansion specifically and in terms of the rhetorical relation of "said/shown" more generally.

C/omission – Here Chion presents the situation in which "what is said causes what is not said to emerge in a particular manner, like the phenomenon of the elephant in the room that no one mentions" (p. 387). Let us revert here to Chion's original French term which, as the English translator of *Film: A Sound Art* discusses in a footnote, is *creusement*, a word that gives a sense of "hollowing-out, creating a negative space or a void" (p. 404). In creusement, characters do not utter or refer to something that is clearly happening. The *said* is a void, an omission, an unsaid, against the *shown*. Chion presents as an example a moment in a protracted sequence of *A Men Escaped* (Bresson, 1956) in which the protagonist has been naming sounds through voice-over as they happen, sounds that might relate to getting caught during an escape from prison. This naming sets up the expectation of continued naming, but when one sound, a metallic vibration of a wire that resonates throughout the sound field, rings loudly and nearly causes him to be caught, the sound is not named, not mentioned, and therefore is given great dramatic effect (p. 389). Chion considers this to be creusement, or c/omission. Voice-over is *unsaid* here, against the *shown* of a traitorous sound.

Contrast – Chion doesn't define contrast at all, rather he gives a few examples in which what is said in words contrasts with the visual experience. He cites a few Hitchcock instances of characters kissing or approaching each other romantically but talking about something banal or about murder, providing contrast between the words and the image. Another example refers

[40] An approach to composition pioneered by Pierre Schaeffer (Teruggi, 2007, p. 213) and a field in which Chion is well-versed.

to David Lynch's Blue Velvet (1986) where a character hears great threats of violence while only receiving a few blows. Note that in this rhetorical figure, he offers no mention of concrete sound, i.e. non-dialogue, as being what is *shown*, only visuals.

Counterpoint – In counterpoint, which Chion notes is the "situation we most frequently encounter…something is said while something else – neither the opposite nor in obvious contrast to what is spoken – occurs along an independent track" (pp. 391-392). The example he cites is from The Sacrifice (Tarkovsky, 1986b), where a conversation happens between a man and a postman who circles around him on a bicycle. The camera's tracking position, the circling of the bike, and the philosophical conversation "make for temporal vectors in counterpoint, sometimes intersecting and sometimes not, in any case neither opposing nor contrasting" (p. 392). The idea here is that something is offered by the relationship between said and shown, something new, not in contrast or contradiction, but like musical counterpoint, a complementary melody that resonates. Yet, with this rhetorical figure, Chion again makes no mention of non-dialogue sound acting as a *shown*. It is only in the discourse between the words spoken and the visuals that he finds counterpoint in his only example of the figure.

Contradiction – To explain contradiction, which Chion finds to be both rare and of limited use, he describes a character saying "I won't hurt you" while acting violently. He goes on to cite a number of examples in which a character offers voice-over narration that contradicts what we see on screen. Here too, Chion never offers non-dialogue sound as part of the *shown*, in contradiction with what is *said*, only visuals.

Below in Table 2 is a summary of the discursive mechanisms of the five rhetorical figures, indicating the types of examples given by Chion to describe the function of each of the figures.

	Said	Shown	
	voice-over or dialogue	Visual Examples	Non-dialogue sound examples
Scansion	✓	✓	✓
C/omission	✓	✓	✓ (a rare example)
Contrast	✓	✓	✗
Counterpoint	✓	✓	✓ (?)
Contradiction	✓	✓	✗

Table 2: Chion Rhetoric Examples

As noted, for all of these rhetorical devices, Chion gives excellent examples of when the voice-over or dialogue of a scene, in discourse with image, provides narrative tension or expanded meaning, or affects the listener-viewer in some way, although to be clear, he doesn't offer an explanation of just what the listener-viewer experiences or is meant to experience. According to Behrens cited at the top of this subchapter, if Chion calls these "rhetorical figures", they may provide "argument" of some kind, or perhaps they are a reflection of Booth's *rhetorical stance,* the balance among "available arguments", assumed "interests and peculiarities" of the audience, and "voice", as will be explored later. Nor does he approach a cognitive rhetoric, as expressed by Dascal and Gross, that would reach for intended meaning. We thus find that Chion's canon of examples does not provide enough fodder for rhetorical discourse. More specifically, he seems limited where non-dialogue sound takes up the position of the *shown* in his rhetorical structure, though he goes to great lengths to remind the reader throughout the chapter that "concrete sounds", a term as yet not well-defined by Chion, are indeed included in the *shown,* as we previously addressed. Only in scansion, which is indeed very common, and c/omission/*creusement* is a clear sonically *shown* example offered. For *creusement*, however, we are reminded that the example is a rather rare type in Bresson's uniquely structured film, *A Man Escaped* (1956), where the protagonist names sounds throughout, and where minimalistic but precise sound that is usually described by the protagonist is a conscious part of Bresson's aesthetic in his endeavor to create a phenomenological film experience (Belton, 2008). Therefore, it is only *that they are named* that the omission of naming a sound becomes notable in Chion's construct of *creusement*. Decidedly, Chion's rhetorical structures are heavily rooted in an image paradigm, for the large majority deal with dialogue against image and, moreover and as previously discussed, dialogue (except in voice-over) is always tied to the screen even if the speaking source is offscreen. The dominance of both image and dialogue is addressed by Beck in discussing Chion's take on the term "soundtrack" where Chion posits that "each audio element enters into a *simultaneous vertical relationship* with narrative elements contained in the image […] and visual elements of texture and setting" (Chion, 1994, p. 39). Beck (2008) notes that Chion's positioning here suffers from the industry's "biased approach" to the filmmaking process that "perpetuates the idea of the dominance of image over sound in film" (p 73). Moreover, Beck adds that the "rule of sonic signification", a penchant to make sure that narrative is served first and foremost by

ensuring all sounds relating to narrative are prioritized, is responsible for "elevating dialogue to the top of the [sound] hierarchy and shaping basic film sound-studies terminology." We find that again, this grounding of theory in image and dialogue hampers a richer discussion, in this case of possible rhetorical analysis, and later we will offer a modification to Chion's definitions that opens our discourse somewhat, but still leaves us searching for more.

It is important to note that Chion does indeed expand beyond dialogue, returning to The Sacrifice (Tarkovsky, 1986b) in a discussion on counterpoint. Here he alludes to sound effects as the source of counterpoint – offscreen diegetic sounds that are never visualized and whose signifiers, in discourse with the film's narrative, provide much more than the iconic signifier of the obvious sound object. A gust of wind in Tarkovsky can be more than just a gust of wind when it feels like an answer to a character's question from a metaphysical space. For Chion, the wind is the shown, which means it must be in counterpoint against dialogue to work in his construction. He seems to indicate that the wind is indifferent to the constant vocal chatter of the father in this scene.[41] However, if that dialogue were not present, we argue that the wind would still offer counterpoint. If the scene were rendered dialogue-silent, with the image of the father and son walking in the barren landscape, the sounds of wind, birds, lapping water could all contribute counterpoint, and all position Tarkovsky's argument, define his rhetorical stance within the diegesis. The wind might still offer a contrapuntal idea, perhaps in the very same way as offered by Chion, one of indifference. Or perhaps something else, depending on the narrative unfolding at the moment, or the paradigm. Thus, is dialogue, *the said*, even necessary to have such counterpoint, or do his "vectors" of meaning emerge from the combined discourse that produces narrative? With scansion as well, does a sound effect of an eagle screech placed in the right moment, marking time, scanning a narrative beat, require bookending with dialogue? It doesn't, but it must scan a narrative beat, or as is often the case, create one. Narrative, we argue, is a greater determinant of what is *said* than dialogue, and Chion's very insistence that sound effects fall under the domain of the *shown* constrains a discussion of his rhetorical figures and the rhetorical stance of the author.

Chion makes no mention of narrative or diegesis in these rhetorical constructions, except to describe nondiegetic songs with lyrics, and to provide a special place in the rhetoric

[41] We are not sure if the "independent vector" in counterpoint is the wind itself or the indifference that it connotes. Chion is not clear on this.

for this. Here, he asserts, lyrics – in this case the *said* – universalize a character's journey or state of mind, and provide an "acknowledgment of the hand of fate" (Chion, 2009, p. 404). But what if this music begins with a diegetic song, first "visualized" or "synchronized" or "diegeticized" by a character dropping a needle on an album, as with the opening scene of Hal Ashby's classic, *Harold and Maude* (Ashby, 1971)? As a diegetic song, it doesn't fit into Chion's discussion. And yet, what we have in this example are lyrics rendered both diegetic and nondiegetic or, once again, we have a simultaneity, reinforced by the mixing choice, a filmic discursive event, that makes the music sound as if *not* playing on a record player in a room, even though it is. Therefore with songs, the one place Chion addresses diegesis in his rhetorical figures, his discussion misses equally interesting ways lyrics can provide counterpoint and universalize a journey regardless of the diegetic construction. He again positions lyrics in an image paradigm unnecessarily because he seems to defend that since they are words, those words are *said* and not *shown*. But faced with a visual needle-drop evoking the words, are they not first *shown*?

From here, as a thought experiment, we ask what if we were to alter the definition of *said* and *shown* in Chion's rhetoric? What if we rendered these terms in alignment with our multisensory capacity, defining the *said* as what is *heard* and the *shown* as what is *seen*, in other words putting sound against image in a rhetorical positioning of equals. We can examine through this prism greater possibilities in Chion's five rhetorical figures.

Scansion, as earlier noted, scans or marks rhythms and punctuation of the unfolding of narrative, whether that narrative unfolds through visual or sonic information. Nondiegetic sound like musical score can scan, as well as diegetic sound like music from a scene, or a sound effect thought to be of the world. In many instances these can overlap diegetic categories as will be illustrated later. In *The Birds* (Hitchcock, 1963), Hitchcock uses a diegetic song in a scene often studied, that of a schoolhouse, where children sing inside while Melanie Daniels (Tippy Hedren) sits on a bench outside the school, the ominous birds gradually gathering on a jungle gym behind her. In his own analysis of the scene, Chion mentions six functions of the song, three of which involve time: temporal continuity, the exasperation of waiting, and the mathematical structure of time (Chion, 2009, pp. 168–170). He would not consider this to be scansion because in his construct, scansion only scans or marks dialogue. Were he to place rhetorical structure between the *heard* and *seen*, freed from dialogue, he would note that this music clearly provides scansion. It marks the rhythms and punctuations of the unfolding

narrative and growing danger, regardless of the fact that there is no dialogue. Additionally, when examining the function of the song, we find that diegesis is unclear and uninteresting. The song takes on many functions that can place it in any category of diegesis. Lastly, as rhetoric, it may infer many ideas in addition to marking time, carrying a cyclical temporal vector that elongates with each loop, like the diameter of a conch shell moving out from center, and makes the final shot of a jungle film covered in birds seem mysteriously impossible.

C/omission in the *heard/seen*, or *creusement*, that fateful "hollowing out", the creation of a "void" of sound, can be rendered in dramatic moments when sound effects are omitted, as noted in an earlier example of the omission of the sound of a car exploding, inhabited by a woman's child, as the mother watches from nearby. The void, the unheard, offers a counterpoint of unspeakable trauma in a powerful way (the painful loss of the child, the witnessing of horror) and even universalizes it, precisely due to the narrative context and the filmmakers' rhetorical stance. For another example, in the opening scene of *Living Out Loud* (LaGravenese, 1998), during a hushed confrontation between a husband and wife in a restaurant, as the realization that her husband is having an affair dramatically sets in, the background sounds of diners, silverware and a live piano slowly disappear, creating a sonic void that registers emotional pain while also scanning the protracted moment, a moment without dialogue. This too reflects a rhetorical stance in the weight put on the realization while imparting meaning. But is this silence an utterance? As mixers we say absolutely, though we know linguists might disagree. The silence is a statement, a stylistic move that contains inference and imparts meaning at Dascal and Gross's intersection of pragmatics and rhetoric.

Contrast between what is *seen* and what is *heard* can go beyond dialogue to illustrate even contrasting tones between music and image, for example hearing bucolic music against an impending dark scene, whether that darkness is visual or narrative, or something violent. In *Koyaanisqatsi* (Reggio, 1982) for example, Philip Glass' music is at times in contrast to the image. The music can be very violent with rapid arpeggios, over images of bucolic natural scenery, or to the contrary, quite beautiful over images of destruction, as in the gorgeous cello melody that plays over the abandoned, almost bombed-out images of the decayed, windowless and evacuated Pruitt-Igoe housing project memorialized in the film. Let us be reminded that we can also look at this in the reverse – the image is in contrast to the music.

Counterpoint is offered, as previously discussed, in the overall audiovisual discourse as exemplified by the winds in *The Sacrifice*, regardless of dialogue. Similarly, the sound of a

lawnmower offscreen in the distance, for instance, offers counterpoint through symbolic and indexical representation – consumerism and suburban home nearby, respectively – whether or not it is juxtaposed against dialogue.

Contradiction, still rather rare, can be found in our new construction in examples where we see one thing and hear its opposite. A baby can sleep soundly but we hear it screaming. In this example the figure of contradiction, or rhetorical stance, might ask us to remember an earlier scene, or anticipate another, or contemplate the terror we may know underlies the seeming serenity. An opposite example can be found in the film *Clean Shaven* (Kerrigan, 1993), where in the first shot sequence that can be identified as a scene, the protagonist seems to brutally beat up a teenage girl, off screen. We hear screams of the girl, punches, and a violent dog barking in response, but we visually cut to a German shepherd, calm and quiet, wagging its tail, watching, still hearing screams and punches but no dog. We cut away and the vicious dog bark sound continues. This forges a complete contradiction. We also may ask what is the real diegesis here? Is the visual world accurate or is the sonic world accurate? Or are there two, one for the character and one for his surroundings? It's an important question but the answer is irrelevant to the experience. Interestingly, over years of asking students after a single viewing to describe the dog, they unilaterally insist it is viscous and angry. The sonic dominates, contrary to even Chion's position of sound against image, as noted earlier in our discussion of his notion of the "soundtrack," supporting Beck's point that the theorists succumb to a "blind spot" perpetrated by an industry that favors image and dialogue.

We offer this thought experiment to illustrate that the rhetorical effects that Chion tries to capture seem better served when not tied to the screen or dialogue. They can find greater resonance in evoking argument, or a filmmaker's rhetorical stance, when expanded to define the rhetorical relationship between sound and image on equal ground, rather than dialogue and everything else. Moreover, expanded further to put sound in conversation with other sound and with image, we may indeed find greater rhetorical function. We will use these expanded definitions, where *said vs. shown* becomes *heard vs. seen* as a tool of vocabulary in case studies throughout this work. More important, this discourse that emerges here fuels the artistic output that we create as an object of examination and knowledge transfer.

Chion's rhetorical devices are useful and relevant in that they open a conversation about rhetoric within the audiovisual text. Though he constrains most of the discussion to dialogue, the theory does expand well when applied more generously to sound effects and

music, as illustrated in our modified construct above, and are a foundational contribution to the study of sound for film as rhetoric, something we will expand upon later in this work. However, in his focus on dialogue, Chion fails to take a deep look at how *presumably* diegetic sound effects and diegetic music work to provide counterpoint, contrast and contradiction in a nondiegetic way. To do so would challenge their diegetic placement in his tricircle as we earlier noted. In addition, where he beautifully addresses the rhetorical power of silence or sound omission as a signifier, he does so only with dialogue, ignoring the omission of sound effects, and not addressing how silence fits into diegesis.

In summary, when looking at diegetic and nondiegetic sound effects or music in their most creative use, that is, in Chion's "border" areas and through the prism of rhetoric, the pillars of a binary diegetic begin to wobble, rendering diegesis somewhat difficult to define, or perhaps a false construct with respect to sound in film. Moreover, Chion's rhetorical theories are very limited to one paradigm, that of juxtaposing dialogue against everything else, either visual or concretely heard. Yet, Chion offers a glimmer of something different when, in an earlier chapter of the same volume, *Film, A Sound Art* (2009), he queries if we will someday see "the development of an outright rhetoric of the audiovisual relation" (p. 231). He concludes that an "audiovisual language, if there is one, cannot be envisaged or encoded like a visual language." So while elsewhere in the volume he positions this language in a rhetorical framework, he admittedly challenges this approach and remains skeptical that the "true subversive force of the rhetoric of exceptions" will be illuminated. This skepticism is a fundamental reason for the research at hand, hoping to arrive a vastly different framework with which to examine sound's discursive role in film. In order to advance a path forward, we will first discuss additional theories that shed light on our continued journey.

1.4 EXPANDED THEORETICAL CONTEXTS

Here we wish to provide some additional theoretical contexts that become relevant in our analyses moving forward. In order to work towards our aspired goal of liberating sound design theory from image paradigms, we inevitably have to examine ways in which sound functions, and these will emerge naturally out of our artistic output and case study with sound design created for *Koyaanisqatsi*. The functions and responses that we seek to solicit involve

things like attention, meaning-making and interpretation, emotional response and engagement on the part of the audience. It is not possible to put these in discourse without bringing into the conversation certain theories and vocabulary that go beyond sound theory, and at times beyond film theory, but have been applied to film by other scholars, and even occasionally to film sound. This is where we will address, in abbreviated format, these concepts and their relevance to our work. To go into great depth in any of these domains would be beyond the scope of this work, however their omission here as part of our foundational theoretical vocabulary would certainly be an oversight.

1.4.1 ARTICULATING EXPERIENCE: EMBODIMENT, PHENOMENOLOGY AND SURRENDER

In her article "The Immersive Spectator: A Phenomenological Hybrid", Maria Walsh (2004) posits the term *immersive spectatorship* as a way to reconcile what have often been seen as the contrasting views of Merleau-Ponty's phenomenology, based on embodied experience and anchored subject, and Deleuze's claim that film audiences lose themselves via an immersion into a flow of images. We note, however, that Walsh's 15 pages refer almost entirely to the relationship between the viewer and the image, with only three mentions of sound, all of which are embedded in citations and frankly feel like asides (p. 171, 173, 179). Nonetheless we find that Walsh's articulations resonate with our discussion of sound and our analysis of audience responses by bringing into conversation a number of vectors that point to the impacts and experience of sound in cinema for audiences. Walsh notes, citing Laura Marks' "The Skin of the Film: Intercultural Cinema, Embodiment, and the Senses", that Deleuze finds that cinema "can give us 'the genesis of an unknown body, which we have in the back of our heads, like the unthought in thought'" (p. 173)[42] and further that for Deleuze this unknown body is completely abstracted from our sense of self, "removed from autobiographical knowledge and ownership" (p. 174). This seems to be the strength of the experience, to Deleuze, as this abstraction from self allows for "liberation and transformation rather than habitual repetition," as Walsh further notes. This is a compelling idea that, to us, postulates film's ability to bring forth personal transformation rooted in abandonment of self. In addition, Deleuze finds in this liberated immersion a haptic response tied to the optical experience. To him, sight and touch are

[42] Walsh cites (Marks, 2000, pp. 147–148) who cites (Deleuze, 1986, p. 210) in Cinema 2: Time-Image

intricately related, as are other senses, but there is a "derangement of the senses" (p. 175) such that the sensory apparatus triggered is not one that emerges from precise memory but is rather mediated, transformed, "not owned by a particular body" (p. 175). Walsh concludes that this derangement of the senses is fundamental to Deleuze in order to "avoid the reassertion of […] a body returned to its owner in the immersive spectatorial space." Extrapolating this to the audiovisual experience, particularly with immersive sound, to us Deleuze suggests that sound would as well trigger deranged senses, and that deranged sound would be triggered by other senses. This articulates with "materializing sound indices (MSI)", which Chion defines as "aspects of a sound that make palpable the materiality of its source and of the concrete conditions of its emission," like breathing, mouth pops, rasp or the sharpness of walking on gravel; these keep the audience tied to actions and character, and represent the "rendering of the real" (Chion, 2009, p. 480). But for Deleuze, this real render would be out of the spectatorial body. Additionally, the notion of phantom sound, a sense of a sound triggered by a visual-only stimulus, which will be assessed later in this research, holds strong for Deleuze. To Deleuze, haptic responses should indeed be triggered by sounds, and phantom sound would be triggered by visuals. However, these sensations would not come from memory, but rather something new, devoid of self-reference, in a virtual world. In contrast, Walsh notes that for Merleau-Ponty's phenomenology "communication or 'interimplication'[43] between body and world is always occurring" (p. 179). In other words, in the filmgoing encounter, the embodied sensorial experience is not devoid of self, according to phenomenology. The real, the body, implicates the virtual and vice-versa. As Walsh summarizes, "the main difference between them is that Merleau-Ponty locates the body in the intersection between the points of crossover of the virtual and the visible" (p. 179). We would postulate that this crossover contains not only the visible, but the hearable and, in fact, we lean into sound as being somewhat more grounded in the real, more able to potentiate the body. As Kittler (1999) presents to us, the very nature of sound frequency curves over time, sonic envelopes, spatial indicators, and manipulability of sound collectively position recorded sound in the real, rather than the symbolic or imaginary. Sound in the theater is not nearly as representational as image, for the very physics of sound waves reaching ears in a screening, especially with greater sonic immersion and modern fidelity, is no

[43] Here Walsh cites Elizabeth Grosz, in *Architecture from the Outside:Essays on Virtual and Real Space*, (p. 89) using her term "interimplication" to describe an intertwined relationship between the virtual and the real.

different than the physics of the soundwaves it purports to represent. For vision, engaging with a two-dimensional screen without depth, and mediated via progressive frames and discontinuous edits, image is fully representational, or imaginary. As Kittler notes "phonography and feature film correspond to one another as do the real and the imaginary" (p.199).

As Walsh continues with Merleau-Ponty's phenomenology, she begins to resonate more deeply with our query into sound's less didactic functions. She describes a "dynamic passivity […] whereby one is captivated by/in an image, yet dynamically moving within that incorporation" (p. 179). She further discusses Merleau-Ponty in his *The Visible and the Invisible* suggesting that:

> poles of passivity and activity are continually crossing over and being exchanged with one another due to the embeddedness of the subject in the world" (p. 180).

We further posit that the screening environment plays a strong determining role in the push and pull between Walsh's poles, following Hven who asserts that when we query the interaction between mediated cognition and emotion, the "environmental scaffolding" (Hven, 2019, p. 7) is as much an element of the audience's phenomenological apparatus as are the three structural tentpoles of film phenomenology (film, spectator and other spectators). Ultimately, Walsh reconciles the embodied spectator with the transported viewer by elaborating on her concept of the "immersive spectator", describing an "interval" that "marks the scene of immersive spectatorship" where "liberated sensations unleashed by immersion in the flow of images pass through the body as the body punctuates the anonymity of these intensities" and where "the spectator is a willing captive" (p. 183).

Sound design, we posit, is critical to Walsh's immersive spectator, though it is never mentioned in her formulation. When we look at sound design, we have a wide range of sound: from materializing sound indices that tie one to the real with tactile sensation (and most often with movement on screen) to abstract, amorphous, acousmatic sounds whose source is not-yet or never revealed and from which no virtual image nor recognizable touch can be generated, and all along a spectrum between these. We see sound design, therefore, as the music that drives a dance between the real and imaginary, enabling the embodiment of the imaginary of both image and abstract sound, carrying the willing spectator through the flow in and between phenomenological engagement and objective cognition. *Surrendering* to the dance, to the ebb

and flow, is what filmmakers and sound designers hope audiences will do, and sound's flexible position between real and abstract, we posit, seduces the *surrender*. This also articulates with hermeneutics, a subject we address in the next subsection, but suffice to note here that Ricœur addresses a type of surrender when he states that when a spectator inhabits a film world, they lose themself as a master to become a "disciple of the text" (Ricœur, 1981, p. 113). For this reason, we will later refer to phenomenology, immersion, hermeneutics and embodiment with this dance in mind, and perhaps lean into the term *surrender* to encapsulate them all from the position of the listener-viewer: *surrender* not to suggest Deleuze's disappearance of self, but to enunciate surrendering control of cognition to a flow of anonymous intensities and cognitive awareness. This will be particularly relevant in the construction of our artistic piece and in our case studies and content analysis.

1.4.2 Experienced Interpretation: Hermeneutics and Phenomenology

Later in this work, we will interrogate the interpretive prowess of audiences based on a questionnaire with open answers. To look closely at interpretation, we must confront hermeneutics in order to expose variables and processes that impact how an audience does make meaning given the symbols provided too them. Furthermore, with our qualitative analysis, we as researchers will be interpreting data, the written words of the audience members, categorizing them via interpretation of text, and analyzing them. Therefore, hermeneutics also comes into play in our evaluation of data. Thus we find a discussion of hermeneutics particularly relevant, especially where it articulates with phenomenology. While a deep background on hermeneutics is beyond the scope of this thesis, we shall explore this briefly here as a theoretical foundational component for our subsequent discussions of narrative and thematic interpretation for our cases studied, including our artistic output.

In his book, *Hermeneutics of The Film World: A Ricœurian Method for Film Interpretation*, Alberto Baracco (2017) posits a methodology for interpreting film based on Ricœurian hermeneutics with a grounding in phenomenology. Film hermeneutics is an examination of the interpretation of film, taking into consideration the challenge that different audience members can find different meanings, and even one person's interpretations can change over time. It considers movies to be worlds and listener-viewers to be interpreters,

forming meanings that are inherently "relative and contextualized" (p. 86) via the subjectivity of the filmgoer and the filmgoing experience. Phenomenology, Baracco argues, is fundamental to film hermeneutics due to the "encounter between screen projection and filmgoer" where

> the filmgoer enters into an intentional relationship with configurations of meaning that are not reducible to such empirical data. Immediately, in film experience, multiple horizons of sense open up to the filmgoer, and many perceptual possibilities are available through which that data can be experienced (pp. 43-44).

Moreover, the listener-viewer may, according to Baracco, self-mediate perception by magnifying or diminishing intentionality of perception (p. 48), which to us shows a control over the level of surrender to experience, on the one hand, and meaning-making on the other. Perception and signification, the concerns of phenomenology and hermeneutics, intersect in the film world with the presence of the viewer in that world.

> From the perceptual origin of sense, film experience leads us to an interpretative analysis of the symbolic and philosophical meanings of the film world. The film world is the place where the perceptions of the filmgoer become meanings (p. 58).

Combined, therefore, in "the phenomenological hermeneutic approach" (p. 76), Baracco honors the interweaving and inseparability of experience and interpretation. This purview "aims to elucidate phenomenologically the lived experience of the film world and reveal hermeneutically its meanings through a process of explanation and understanding" (p. 87). Baracco further defends the approach noting that Merleau-Ponty's ideas of embodiment and intentionality form a strong premise for grounding the listener-viewer's perception in the film world, while the Ricœurian hermeneutic position "emphasizes the sociocultural and historical contexts from which film interpretation emerges" (p. 88), collectively offering us a more holistic assessment of the interaction between audience and film, experience and cognition. Moreover, this approach is in opposition to any pre-determined definitive meaning embedded in text, insisting that the process of interpretation "remains always open-ended and in constant search of meanings" (p. 91). For us, the notion that phenomena and interpretation are not mutually exclusive and even implicate each other is indeed compelling when we look at sound, and it resonates well with our earlier discussion of phenomenology in the context of embodiment and surrender. We posit that sound is a key and critical factor, as earlier discussed, in the embodied experience of Walsh's "immersive spectator" as well as for Ricœur's "disciple of the text" and

for our concept of *surrender* that places the listener-viewer in a dance between cognition and experience that at times encompasses both simultaneously and to varying degrees. Yet, we must point out that, not surprisingly, Baracco's film hermeneutics hardly mentions sound. In 306 pages of text, we find the word "sound" only 6 times, and "images" or "screen" 93 and 56 times respectively. Baracco therefore either falls victim to the idea the meaning is generated overwhelmingly from images alone, or he shies away from bringing sound into his analysis. We see this as a great challenge for our work, for we are interested to see how these ideas articulate not only with our case studies but with our content analysis that interrogates meaning-making and experience in the sound-image discourse, with sound on equal grounds.

Our remaining challenge is reconciling rhetoric and pragmatics with hermeneutics, as the former is concerned with a speaker's perspective (filmmaker) while that latter is concerned with the receiver's (audience). We will indeed put these in conversation through our content analysis where we can weigh intention on our part against interpretation and experience of the listener-view when sound effects alter the audiovisual text that we create as the grounding and interative component of this artistic research.

Sound designers and mixers know, via praxical and tacit knowledge, that the choice and manipulation of sound effects produces results that alter, drive, or support narrative via signification, cognitive framing, and interpretation, and also contribute vastly to the non-narrative component of an audience's experience, in terms of the emotional impacts typically associated with music, embodiment, and the phenomenological aspect of the movie-going experience. However, researchers and critics specializing in film sound have not demonstrated these impacts on audiences, which leads us to ask just what these impacts are, can they be measured, and what interesting take-aways might result from an analytical study that seeks to determine if goals of sound design are measurable. This framing of this study and the design and results via content analysis are the content of Chapters II and III. But first, we shall take the theories discussed and the weaknesses they expose to the edit room and the mixing stage, transforming an object of study along theoretical grounds and via artistic research, in preparation for content analysis.[44] This represents, at this stage, a form of artistic research that

[44] Throughout this chapter, vocabulary originating with a variety of theorists has been cited repeatedly. Moving forward, while we will frequently remind the reader of the context of these terms and their founding theoreticians, we will not continue to cite specific readings and pages.

can be called research-based art in that the literary review provided up to this point informs the artistic output that we create for further examination.

1.5 ARTISTIC RESEARCH – DESIGNING NEW SOUND FOR *KOYAANISQATSI*

An examination of the relevant foundational theories of diegesis and rhetoric, and their theoretical underpinnings in sound as elaborated by Chion, and further explorations on semiotics, phenomenology, embodiment, and hermeneutics have resulted in many questions and examples that probe the robustness of dominant image-based theories for fully examining sound's relation to image and other sounds in the audiovisual construct of film. As practitioners, our artistic research seeks to create a sound design that illustrates some of the unresolved aspects we have explored in a single film excerpt that can be easily accessed by the readers of this work and further used in content analysis to test desired outcomes of sound design. As already mentioned, it is in the realm of sound effects where theories seem to fail: in the binary of the diegetic vs. non-diegetic with respect to the screen and in the limited discourse on a rhetoric that positions primarily dialogue in relation to other sounds and image. We argue that there exists a far-more-complex discourse between sound and image, and between sounds on their own, and that a more holistic set of analytical tools would better serve the audiovisual experience and put image and sound on equal ground in materializing rhetorical stance, articulated with pragmatics and meaning. To accomplish this, we turn to a unique example from the documentary film world and tailor a sound design that illustrates desired outcomes, or rhetorical stance.

1.5.1 *KOYAANISQATSI*, BACKGROUND AND CHOICE OF FILM

***KOYAANISQATSI* IMAGE 1: TITLE FRAME**[45]

Godfrey Reggio's *Koyaanisqatsi* (1982) was commercially distributed around the world, mostly in art-house theaters and college campuses where it was a surprise hit (Godfrey Reggio, Cinematic Seer, 2019), and was the first in a trilogy of films. So strong was its impact that it was released on DVD in several formats and ultimately in a restored digital version with 5.1 sound on Blu-ray by the prestigious Criterion Collection in 2012 (*Koyaanisqatsi*, n.d.), 30 years after its initial release. The film is comprised of images with music by Philip Glass. There is no dialogue, and very rarely are there sound effects (SFX), typically light wind. On the educational website around Reggio's work, it is stated that *Koyaanisqatsi* "is not so much about something, nor does it have a specific meaning or value," as it is "an object in moving time, the meaning of which is up to the viewer" (The Institute for Regional Education, n.d.). As such, one can argue that this is a poetic film, with image and ideas, text and subtext, conveyed through rhythms of motion, shots and editing, and strongly by the music as well. Philip Glass's musical score is a critical component of the film for a number of reasons: a) it plays from beginning to end with only a few short moments of break; b) its minimalist structure, embedded with complexity within repetition, offers its own counterpoint or commentary on the images as will be discussed; and c) music is virtually the only sound and therefore it is nearly the entirety of the audio portion of the audiovisual experience. Furthermore, many critics note that the score drives the film, as noted more than 20 years after the film's release by The New Yorker magazine's music critic Alex Ross, "there is no more potent example of a score dominating a film" (Ross, 2005). Ross goes on to describe a juxtaposition that evokes Michel Chion's

[45] All images from Koyaanisqatsi are screen grabs from this 2002 MGM DVD edition (Reggio, 2002).

"counterpoint" or "contrast", if liberated from dialogue as a central point of rhetoric, stating that "instead of trying to make image and music serve the same ends, [Glass] play[s] one against the other, letting the disparity become an emotional experience in itself" (Ross, 2005). Thus, Glass's score already functions in many of the ways sound designers use SFX. For our purposes, as an object with both image and music, we begin with a multisensorial experience even without SFX sound design, so that the addition of SFX does not add sound itself as an entirely new sensorial element, which further justifies the choice of film as an object of study. Thus, a poetic film of image and complex music allows for the addition of sound design to illustrate theoretical strengths and shortcomings in film sound theory, and later to test for SFX's impact on many things, including emotion, engagement, interpretation and meaning. In addition, we return to the new piece iteratively as a translational tool for transforming literary narrative theory to a theory of sound.

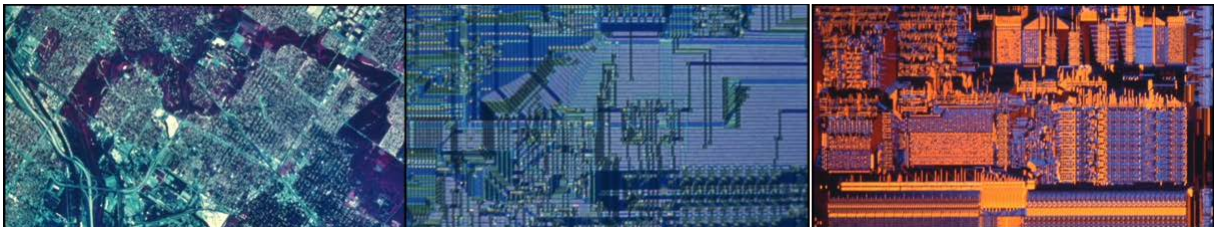
1.5.2 Artistic Process: Adding sound effects

In the 15-minute sequence used in this study, no sound effects exist in the original. As a sound designer, this author added sound effects to the piece, designed to tease out SFX impacts in a number of paradigms. The techniques include crossing diegetic borders, figurative and at times metaphorical use of sound, constructed counterpoint to drive meaning, pure diegetic mimesis, creating diegetic space that otherwise doesn't exist between shots (to examine if the connecting of such space leads to predictable constructed meaning), and silence as a nondiegetic element. Many of these specific design choices will be discussed below (section 1.5.3) where relevant theory from earlier in this chapter will be applied. In terms of the overall process, sounds were collected from the sound effects libraries of Sound Ideas (*Sound Ideas*, n.d.) and some created within Avid Pro Tools (Avid Technology, 2018 version 2018.7). After selection and editing in Pro Tools, many were transformed via various tools of sound design, and then they were mixed in both 5.1 and stereo speaker configurations in a film mix studio on a Pro Tools S6 console at Universidade Lusófona's teaching and research sound studio in Lisbon, Portugal. The two speaker configurations were made to accommodate different environments. The creative process of sound design, namely the selection, editing and mixing techniques, was targeted to demonstrate a much more complex deconstruction of sound, post-modern and rhizomatic, as sound directly contributes to narrative, expands discourse and breaks

diegetic conventions. Many of the techniques are very commonly used, although rarely in the aggregate in such a short piece, underlining the need for a new type of rhetorical analysis. For discussion purposes, we will use the terms diegetic and nondiegetic on a traditional basis in order to begin within accepted film sound constructs and to position discussions within Chion, and so that the weaknesses of that positioning can be exposed. Similarly for rhetoric, we will elaborate on both Chion's *said/shown* and our *heard/seen*. For reference, at this writing, the clip with sound design can be found here in the footnotes[46], and links to both versions can be found in Appendix A. Note that as this design was informed by weaknesses in accepted theories via literature review and years of teaching, the construction of this work can be considered as research-based or research-led artistic practice. We suggest that the reader watches and listens to the entire clip before returning to our discussion.

[46] https://vimeo.com/308269202

1.5.3 ARTISTIC PROCESS: APPLIED THEORY TO THE SOUND DESIGN



Koyaanisqatsi Image 2: Aerial Sequence
Timecode: 0:30 – 2:00

The clip opens with aerial shots of a city. We hear birds and an eagle screech, gently supported by air or wind. Yet, we are too high to hear such things, forging a "mismatching of sound and visual distance" (Fischer, 1985) that creates a contrast and adds weight to both sound and image by providing a subtle narrative tension – they don't quite match. The sound effects continue gently over alternating images of city, native American tapestries and computer chips. Thus we have, at first, a diegetic acousmatic and therefore off-screen sound, diegetic because we are in the sky, but with a mismatch of distance; a discursive mixing choice. This sound crosses a border, becoming nondiegetic with subsequent images of computer chips and tapestries, forming a connected metaphysical space between images not diegetically related to one another, but graphically related. This creates a metadiegetic space in the sense that a narrative within the narrative is evoked by tying together these elements, supporting the metaphor that the images are meant to create, however interpreted by the listener-viewer and augmented by the counterpoint of a rhythmic eagle screech whose return after return, ever-lengthening in time intervals between utterances, may alter temporality for the viewer. The completed sign, therefore, emerges from the discourse among and between individual visual and sonic signifiers. As noted in section 1.3.1 in our discussion of diegesis and the represented instance of Christian Metz, we created the term *thematic diegesis* to describe just this type of phenomenon, as a way to capture a unified created space of interconnected meaning, as opposed to a physical one or physical diegesis, and also articulated this concept with respect to hermeneutic interpretation. Our term *thematic continuity* speaks to the trail of meaning generated, and these terms are of course interrelated, emerging from the intersection of phenomenology and hermeneutics as expressed earlier in this chapter, for "in the film experience, two different plans of perception and meaning coexist for the filmgoer, that of the filmgoer as the subject of the real world and that of the filmgoer in the film world" (Baracco,

2017, p. 57). By entering the film world, diegesis includes the formulated thematic interpretation of the listener-view since, from this perspective, they are part of the world.

To restate, *thematic diegesis* is not of the physical world of the story nor of the storytelling, but it is of the world that the story communicates. It is neither an Aristotelian diegesis – the telling part of diegetic mimesis – nor that of the filmologists who swapped diegesis and mimesis, calling diegesis the story world. As we discussed, it may be closer to Metz's "*represented instance* […] – that is to say, the sum of a film's denotation" (Metz, 1974a, p. 98) if we are to consider the emerging theme to be denoted by the juxtaposition of sound and image, or the *sound-image*. This is, in a certain way, at the crux of our argument, for image provides signs and sound provides signs. But particularly if we are looking at indexical or symbolic signifiers, sound objects and their acoustic waveforms denote more than the identity of the sound object. The reason we embed the term diegesis in our *thematic diegesis* is because one can argue that the connected theme is pro-filmic, even when the instance is represented through filmic discourse, and therefore this expanded and qualified use of the term *diegesis* follows a certain linguistic and critical logic.

Koyaanisqatsi Image 3: Stadium to Office Towers
Timecode: 2:18 – 2:40

The next sound effect provides a similar example, with a cheering crowd at a full baseball stadium, sound arising, and rising in level, synchronously to an aerial shot passing over the stadium. The sound, lowered and with large reverberation via mixing, continues into the next shot, a nighttime cityscape with office buildings. As the distant crowd cheers and celebratory airhorns blow, we are provided with a strong *counterpoint* that asks a question; why do we hear the stadium cheering over images of workers in offices late at night? Diegetic cheers have crossed the border into nondiegetic space, yet assert a narrative element while acting as musical score. This is a simultaneity of diegetic placement and speaks again to our concept of forging a *thematic diegesis and continuity* via sound between otherwise disparate images. It

also relates to Chion's *counterpoint, contrast*, and/or *contradiction*: counterpoint in that the cheering forms a new vector of meaning against the buildings of workers (or vice versa); contrast with bright cheery sounds over a dark image of office towers; and/or contradiction in hearing sounds of pleasure or leisure with images of work. Note however that to call it counterpoint, contrast or contradiction we must use our altered definitions of Chion's *said/shown*, as discussed in section 1.3, namely *heard/seen*. We argue that, no doubt, it is the juxtaposition of sound and image, or the sound-image itself, even though the sound does not contain dialogue, that provides these rhetorical figures or, more precisely, Booth's *rhetorical stance*. On the receiving end of the communication, or pragmatically speaking, we are hoping for a cognitive response, implying inference and an approach of cognitive rhetoric.

Koyaanisqatsi Image 4: Factory Man
Timecode: 2:41 – 2:55

The next shot is of a man, sitting and smoking, in a control room of some sort, perhaps a factory or power plant. The sound effect that has been added is a warning alarm, along with a low frequency rumble. These are in direct contrast with his calmness. There is no visualized reason for there to be an alarm, and it is never grounded in an image, such as a flashing light. The man's lack of reaction defies simply an offscreen diegetic emission of the sound. Expanding on Chion, this is perhaps a *c/omission* where what is omitted is the visual rather than the sonic – we hear something but it is ignored. Let us recall that c/omission for Chion occurs when a void is created by not stating in words, in dialogue, something visually obvious. Image is in the foreground and dialogue ignores. In our expansion of his *said/shown* to the *heard/seen*, this can apply to a void of sound as in a silent explosion earlier discussed. There too, image is the foreground – the exploding car – and sound falls back into a void. But here we suggest that in the case of the factory man, perhaps we flip the script to illustrate that we hear the urgency

but see nothing that addresses it, for why must image be considered the foreground rather than sound? This highlights for us that the sense of void created in the rhetorical structure between image and sound can extend beyond the constraints of Chion's definitions. Moreover, the combination of the alarm and the visual seems to drive a different narrative than just the image and music, like a musical discord, and evoking an interpretation or tension determined or simply experienced by the observer. For instance, the man ignores the alarm of stress because he is bored, it happens every day, or he'd rather die anyway, as he puffs his cigarette. The possibilities are numerous, reflecting of course the variety of interpretations permissible and equally valued by hermeneutics, but they arise because of the rhetorical stance of the filmmakers, us, who have chosen this juxtaposition, which also comprises inference.

But why low frequency rumble? As mentioned earlier in 1.2, the advent of Dolby noise reduction and surround sound brought us low-frequency capabilities for film sound. As Dolby later advanced to 5.1 and beyond, the subwoofer was added to theaters and has now become ubiquitous in home-television screening environments as well. Subwoofers are able to reproduce very low frequencies, down to 20hz, the bottom of our hearing threshold for frequencies. As a discursive tool, these frequencies are often used for imparting emotional responses that are typically associated with things like anxiety. Many studies around music and low frequency noise (LFN) have been done in this area. For instance, a 2013 study in Korea exposed volunteers to sustained pure frequencies of 31.5Hz, 63Hz, 125Hz and 250Hz (Eom et al., 2014). Correlation was demonstrated between frequency and emotional responses such as unpleasantness, stress, anxiety, and annoyance as well as perceived physiological characteristics like ear pain, headache, nausea and dizziness. These responses were also greater with increased volume (sound pressure level). Importantly to our study, the idea of perceived physical vibration seemed to directly correlate with these negative impacts, with perceived vibration being far stronger with the lower frequencies, particularly 31.5Hz, which only a subwoofer would produce. This speaks to the embodied nature of sound and implies that via emotional and physiological changes in the audience, the experience of the audience can be altered. Furthermore, sensation and interpretation are interrelated as earlier discussed through phenomenological hermeneutics. As a rhetorical structure, as already noted with Dascal and Gross, if emotional impact or style embedded in the subwoofer design includes inference, than there is a pragmatic and rhetorical function at work as well.

Koyaanisqatsi Image 5: Elevator
Timecode 2:56 – 3:36

Leaving the factory man at 2:55, eight alarm buzzes continue, albeit diminishing in volume, over the next shot which has nothing to do with the factory shot but rather a very long single shot of crowds stuffing themselves into a lobby elevator. Visually, we do not interpret this as the same diegetic space as the factory, yet the factory sound continues. This "fade out" of a previous ambient sound over a new scene is a very common rhetorical devise that creates connectivity between locations narratively, if not spatially, as does the Phillip Glass score since it is a singular piece bridging these shots and sequences. How, then, do we define this delayed audio fade out? Like an edit or a visual dissolve, an audio dissolve *is not of the story* even if the isolated sound signifiers that make the dissolve are. We are reminded of Chion's assertion that a word is a word, but the sonic treatment of it, like the growing reverb on Hamlet walking away, is not part of his *said*, but it may or may not be part of the diegesis. In our case, an overlapping fade is clearly a nondiegetic part of the storytelling, discursive, rather than of the world of the story. And yet, we may have assigned diegetic function during the transition; are these spaces connected? This harkens again to the notion of a *denotated instance* in the very overlapping, or cross-referencing of two scenes, two locations, two parts of a story. This cross fade of a non-visualized sound effect dances within the tricircle, seemingly disinterested in diegesis but rather concerned with phenomenology and narrative or thematic interpretation.

Concurrently, in the elevator shot, there is no sound for the people, only a rumble, and from where we don't know. What is the silence? A c/omission? A counterpoint? Does it highlight the loneliness of the crowd? Is it diegetic in that the people do not speak with one another so it works as a metaphor for their world? If silence is a signifier of a psychological state of a character, is it not *internal diegetic*? What is more important is to recognize that the diegetic question plays a subordinate role to the phenomenological one; we can better feel or

experience the loneliness or the absurdity of the content because of the silence.

Koyaanisqatsi Image 6: Shots from People Sequence
Timecode: 3:54 – 5:24

Loneliness is further explored in shots that soon follow (3:54 – 5:24) of individuals looking into or near the camera. Sonically, the idea is supported by the metaphorical use of a telephone ringing and a dial tone – never picked up, the human connection never being made. This phone ring is pure neutralization of sound as for Eisenstein or Vertov, set in motion by montage in a juxtaposition against the images and over time, between what came before and what is yet to come, in a network of rhizomatic signs. Metaphor and counterpoint through sound montage is visceral and serves to reinforce narrative clarity. It calls for a *represented instance*, or an inference that emerges out of a striking utterance. This film is, in part, about disconnectedness in a "life out of balance", a translation of the word "Koyaanisqatsi" from the Hopi native American language (The Institute for Regional Education, n.d.). While the images certainly convey this, the addition of a sound effect acting as score – nondiegetic, driving emotion – that concurrently and directly reflects the state of mind of the characters – like diegetic voice-over or musical lyrics – defies a diegetic placement.

Koyaanisqatsi Image 7: Cigarette Lady
Timecode: 5:25 – 5:53

A similar effect is seen in the heartbeat that soon follows over the image of an older woman, struggling to light her cigarette (5:25).

Koyaanisqatsi Image 8: Cigarette Lady Connected to Emergency Sequence
Timecode 5:43 – 7:03

The next sequence of images contains a diegetic trick created with sound from 5:23-7:03. In the original film, the woman is thought to be in a crowded street. A visual cut is made to a new location where an unconscious man is being helped onto a stretcher. The next cut takes us to a young woman sitting in a fancy car, closing the window. This is followed by a disheveled man sitting half naked in his window, looking out, and throwing his arms up in frustration. The next cut is of a crowd gathered in 1970s South Bronx rubble with fire trucks, apparently the aftermath of a fire in the crumbling decay of impoverished urbanism, a recurrent theme in the film. There is then a cut within the scene to a fireman walking through the crowd in slow motion, isolating himself, and eventually stopping to see his own reflection in the residual water from fire hoses. In the original, these shots feel like a montage of neutralized images, connected only by the Philip Glass score and by our interpretation of, perhaps, decay. However, in our sound-designed version, sounds are added that denote an emergency location with sirens, police

radios, and crowds. But these sounds are advanced forward of the emergency images, fading up with the previous shot of the older woman with the cigarette, just as she looks up and off frame somewhere.

Carrying these sound effects from her, then over the cut to the unconscious man, now gives us the impression she is drawn into him. What is inferred is that she sees him and everything that follows and, as such, she becomes our proxy, observing the chaos and despair. Continuing the sound, we feel the young woman in the car who, with sound continuity, now seems to also watch the man being put on the stretcher and it is too much for her to bear, as a driver starts the car, denoted only by a sound signifier, and she closes her window, accentuated by a sound signifier and therefore mimetic synchresis (mimetic because it is not the actual sound). A metaphor is set up: the moneyed class, who can afford to look away, does, an inference solidified by the sounds. The half-naked man in the window is frustrated and angered by her departure, which is now happening in his sight and below his window only because of the sound effects connecting space; we hear the woman's car driving away on his shot.[47] The background sounds of emergency and crowds continue throughout, unifying a space that was once a collection of neutralized images and creating a thematic narrative far more concrete, yet still poetic, than what had been presented in the original. We have, in essence, joined these in both physical diegesis and thematic diegesis, as well as continuity.

Koyaanisqatsi Image 9: Fireman in Emergency Sequence
Timecode: 7:06 – 7:45

To cap this off, as the fireman walks away from onlookers to find a place to be alone, his footsteps become more and more isolated from the background crowd, eventually completely, and drowned in reverb as he steps into the puddle and sees his reflection at 7:45.

[47] The careful viewer will note that we even see just the top half inch of the roof of a car driving in his shot, a fortuitous accident.

While the footsteps, synchronized, are construed as diegetic, the reverb treatment is not – it acts as nondiegetic musical score in that it conveys an interior emotional quality of the character. Moreover, the manipulations of the mixing faders that gently reduce the background sounds to silence while accentuating the footsteps, that discursive fader move, it can be argued, is clearly part of storytelling or "style," but, then again, as a reflection of the psychological state of the fireman in the moment, it can be regarded as internal diegetic as well, like how Chion regards musical lyrics. It is certainly representational of his state, and therefore may be characterized as a represented instance within the diegesis, structured to relate inference, argument and meaning. On another note, given the embedded inference, the mixing move here is indeed a Searle utterance which speaks to the performative nature of mixing, something we are interested in exploring in future work.

The first example of the birds and then the stadium crowd provide what we have previously defined as thematic diegesis, that is to provide a sense of connected space that is beyond physical. There is no appropriate place for these in Chion's tricircle as they can exist concurrently everywhere in the circle, highlighting the visually-determined nature of Chion's construct. The last example of emergency sounds, as a background sound effect providing diegetic space for a sequence that is otherwise disconnected, is much more than simple sound continuity. In sound continuity, a background effect is used to reinforce the visual existence of a unified space, masking the apparatus of production and editing. Here, however, we are creating an interconnected physical space that never existed in the first place. This is not traditional diegesis and, perhaps, something like *sonic spatial construction* would better describe the phenomenon. As we move through this analysis, we seem to be asking not simply if sound has trouble fitting into Chion's diegetic construct, or anyone else's for that matter, but rather if it undermines it altogether and in this subversion finds its poetic, imaginary power, harkening back to Chion's own previously-mentioned skepticism about the ability for rhetorical analysis, including his own, to disclose the subversion (Chion, 2009, p. 231). The nondiegetic sound world works because the audience does not notice it. Typically music, these sounds secretly drive emotion, support tension and release, and trigger immense pleasure (Blood & Zatorre, 2001) as well as stress (Gerra et al., 1998) through neurological and endocrinological activity from the interplay of consonance and dissonance, rhythm and punctuation, counterpoint and instrumentation. When sound effects do this, particularly sound effects that can be thought of as diegetic, they subvert the structure of diegesis, moving an audience beyond the comfort

zone of what can be defined as "of the story," and reaching out to the universalities of phenomena. To us, this carries over into the previously described dance within the space of phenomenological spectatorship, or Walsh's "immersive spectator" as earlier discussed, engaged in "dynamic passivity" via both a phenomenological "interimplication between body and world" and a Deleuzian abstraction from self (Walsh, 2004, p. 179). This idea permeates throughout this discussion and the thesis.

Koyaanisqatsi Image 10: Man with Change
Timecode: 8:02 – 8:23

In one shot of the film beginning at 8:02, there is a man looking rather downtrodden, possibly homeless, shuffling unbalanced up the street towards the camera, his palm outstretched. There is small change in the palm of his hand. Here, we added the sound of children playing in a playground. One might believe this to be offscreen diegetic sound in Chion's construction, like children in a playground nearby. However, it is not meant to function that way, but rather as a nondiegetic figurative component that may add ideas along a spectrum of responses, from innocence or happiness to a darker contrast between hope and despair. Thus, again, if redefining Chion's rhetoric to exist in the juxtaposition of the *heard/seen*, it can be thought of as the counterpoint of a new vector or even a contrast. Diegesis of the children, by film school or Chion's definition, is irrelevant with respect to function, which lies more in a rhetorical posture. Moreover, the utterance of this sound pragmatically infers, implies rhetorical argument, and contains proposition by referring to happy children, illocution by being stated as a suggestion[48] and perlocution via its intent to generate a thought process or experience in the listener-viewer.[49]

[48] We base this on the naturalistic mix rather than something overtly loud.
[49] Many students over the years have expressed a changed response to this image, as well, from one of despair to one of either hope or childish joy on the part of the man, due to this semantic change from sound.

Koyaanisqatsi Image 11: Stock Market
Timecode: 9:17 – 9:46

Later, we are given an image of a stock market trading floor from above. There is a ghosting visual effect, most likely created as a superimposition of various segments of the same shot over time, with no movement of camera. To this shot we have added sound effects of crowds and machines from a casino, hoping to combine two ideas. On the one hand, through synchresis[50] of busy human and machine sounds, we mix in a sound object whose sound feels diegetic to the space presented, where noisy humans and machines are spinning. On the other hand, the choice of sound, the diagrammatic dimension of the sonic signifier, is not what it seems, providing some degree of figurative counterpoint.

There is much more to the sound design in this clip, however we shall skip to the last sequence. At 9:45, a rocket ship fires and takes off into the sky in a few shots followed by a single shot that lasts 3 ½ minutes. Sound design as added by us was meant to complement, not alter, the narrative in order to provide counterpoint and to extend the audience's ability to remain engaged. The rocket is given life with the sounds of an engine, with a sound as it breaks away from the launching pad, and with audible soaring through air. It is given the sound of an explosion and some synchresis of fire flares as a piece tumbles back to earth. The sounds used to create this, however, are not so straight forward. They defy the simple mimesis of diegetic audiovisual assignment through synchresis.

[50] Perception of a visual event and sound event as happening concurrently, where the visual action appears to cause the sound, due to placing them synchronously (Chion, 2009, p. 492).

Koyaanisqatsi Image 12: Rocket Subsequence A
Timecode: 9:47 – 10:43

To begin with, crowd sounds from the previous scene, comprised of cheers, are extended well into the beginning of the sequence, a shot that shows debris falling from the launch pad but at a framing that is unclear, suggesting some recent explosion. To these cheers are added subtle screams of horror, increasing the tension and placing into discourse sound with image and with other sound. Do we take these cheers and screams to be diegetic? Do we think there is a crowd watching some horrible destruction, like 911 or the Challenger disaster?[51] Or again, does it act as score, setting in dissonance two sounds and an image? The rumble of the engines is subdued – even for a closeup fiery engine, the actual sound used is merely a boiler room rumble with some low hollow tones added, not at all diegetically composed, but providing enough of a base to believe there is diegetic sound and to set up certain expectations, in tandem with the emotive and physiological impacts of low frequencies on the audience. As the ship breaks away from the bright red pad, a monstrous thunder with lightning-crash sound punctuates the break, a sonic blow accentuated by the subdued treatment of the rumble before it, highlighting syntagmatically-organized mixing signifiers – a quiet and low frequency followed by a bright thunder crash. This surprises and shocks, illocutionary acts of utterance. It is a combination of three rhetorical devices that Lucy Fischer (1985) calls "an abrupt tonal contrast," an "abrupt sound break," albeit in reverse, and a "metaphorical use of sound" (pp. 254-255), the metaphor itself offering a range of interpretations, perhaps a lightning flash of brilliance or an impending storm. Note, here we point out the semiotics and speech-act implications of mixing, beyond the choice and placement of sound effects. While this has and will continue to be mentioned, it is something to be explored in greater depth in later research on the rhetoric of mixing, as it is through mixing that all these nuances are achieved.

At 10:45, the long single shot begins as the Phillip Glass score returns to the theme from the start of the film, a meditational pattern based on the fundamental musical intervals of

[51] In our qualitative analysis, quite a number of tested viewers indicated that it was footage of the challenger or related to 911, even though both events postdate the release of this film.

5th and octave, in a minor tonality. This shot is stunning, impossibly-long, documentary footage of a rocket taking off and soaring through the sky, and more. It is 3 minutes 20 seconds long, and then jump cuts in, somewhat continuously, to an even slower motion of the shot.

Koyaanisqatsi Image 13: Rocket Subsequence B
Timecode: 10:58 – 11:41

As the shot begins, we feel some sense of concern but also triumph, giving us time to meditate on humankind's greatest technological achievements. Here, sound design adds a different rumble, a kind of soaring one, that has some sonic unsteadiness built into it, even though the image is rather stable. As the rocket moves away from us, there is a gradual, imperceptible shift via a long and slow discursive audio dissolve to an artic wind. We have taken flight. It is not at all a realistic sound, but lulls us into a sense that the rocket heads towards the heavens.

Koyaanisqatsi Image 14: Rocket Explodes
Timecode: 11:42 – 11:59

Suddenly at 11:42, without warning, the rocket explodes, already two minutes into the shot. Synchresis provides a sound for the explosion, but it is made from sonic manipulations of a screeching animal, feeling more like a sound that one imagines would be made by a dinosaur in agonizing pain. In this case, an imaginary is rendered real through synchresis. Metaphor,

abrupt tonal contrast, counterpoint; the surprise of the explosion has been given much more to experience and to think about, if one is inclined. The sound designer's hope is not that the audience will consciously draw a connection, but will experientially intuit something that might haunt them as they reflect on the experience of the film. The dinosaurs are extinct despite having ruled the earth. Or perhaps it's a simple anthropomorphizing of the rocket ship – allowing it to explode with pain. The signified is not important, only that a signifier has materialized a phenomenon and sparked the human imagination.

Koyaanisqatsi Image 15: Rocket Subsequence C
Timecode: 11:59 – 15:22

The remaining part of the shot shows the rocket in pieces, one large piece in particular, slowly tumbling back down to earth, beginning at 11:59. As it falls, the musical motif is complemented by a simple, repeated melody. Occasionally we hear a flare up of a flame, synchronized with the visual, supporting the fact that this piece of rocket feels almost human, still fighting to survive. Cold wind is gradually subsumed by the sounds of wind chimes, which feel like part of the score, but they are not. They create a mystic, magical sense that, to many listener-viewers interviewed, adds a kind of optimism: the human-looking piece of metal flares up, struggles to come back, and if returning to a spiritual orientation, it just might succeed despite its failures. Note, this last shot ultimately dissolves visually to ancient native-American cliff art, which in the full version of the film, is also shown near the beginning.

None of these sound effects are actually recordings of what we see. They come not from the footage nor from a library of sound effects of a rocket taking off and exploding. A camera placed on the ground with a long lens would not *hear* the rocket through the air, nor the flare ups, nor even a proper explosion. Recording the propulsion engines while the rocket is on the ground would result in nothing more than noise and distortion. And yet, most audiences attribute the sound effects to the image without question, thanks to the power of synchresis.

In summary, many of these design choices built in our artistic research are pro-filmic,

being diegetic in nature, and there are many examples of cinematic discourse as well, mostly evident from mixing techniques. Some work as metaphor, while others provide c/omission, scansion, contrast or counterpoint. Many use synchresis to reinforce an image, or cross diegetic borders or occupy more than one part of the diegetic circle simultaneously. Others bring attention to the absence of dialogue, function non-diegetically as musical score, or support narrative. Some establish the POV of a character, provide abrupt tonal contrasts, or create physical or thematic diegesis. For the entire 15 minutes of the sound design added to this *Koyaanisqatsi* excerpt, SFX seem to make a strong contribution to narrative while actively evading a concrete definition of diegesis, as if diegesis of sound is over-determined structural formulism. Sound design contributes deeply to the *represented instance* of Metz's more thoughtful sense of diegesis, even while being nondiegetic by traditional definitions. This artistic research shows us that when Chion first addressed the notion of a rhetoric of sound for film, he touched upon a major expansion of how we think about film sound, and yet he barely moved this idea forward, still locked to the screen, to old notions of diegesis and to dialogue.

Thus we have addressed our first of four areas of inquiry: are current ubiquitous structural techniques of discussing film sound adequate? We believe that we have demonstrated in our discursive reading of the literature along with our exploration via artistic research that, indeed, existing models are not adequate. This leads us to ask where, in rhetorical analysis, we might find a more holistic approach. Beyond screen-based mechanisms our artistic research reflected a *rhetorical stance*, or argument, through the use of sound by the creators, in its relation to both image and other sounds. The rhetorical functions are rooted in components of language theories previously discussed including pragmatics, utterance, and inference, and in the more experiential realm as articulated in our discourse on the immersive spectator in our section 1.4.1, embodiment, phenomenology and surrender. But first, if we are to build a discussion around rhetorical analysis, we must discuss the filmmakers' rhetorical stance and the speech act. Is there actually a cooperative principle in place that would support the notion of a cognitive rhetoric, for to apply sound with a rhetorical stance requires that a filmmaker or sound designer has intended specific outcomes with the sound work accomplished. Certainly, in our artistic research, there were intended outcomes. But pragmatics requires that we, as sound designers, are in contract with a recipient who can form both cognitive and emotive responses from the stimulus we provide. Content analysis will allow us to test if indeed these desired outcomes happen, if the communication sought-after is achieved. We expect that it will

demonstrate that many of the impacts of sound design, namely sound effects, that researchers and practitioners claim to be achieved are, in fact, defendable via content analysis.

CHAPTER II. SOUND DESIGN: ADDRESSING ASSUMPTIONS OF DESIRED OUTCOMES

In Chapter I, we discussed definitions of sound design and a number of theories as they apply to film in general and sound specifically, focusing mostly on diegesis and rhetoric. We examined the ways in which theories, particular those of Chion that are ubiquitously taught in film schools and sound programs, illustrate some of the functions of sound design but also fall short, in part because they are tied to the screen; theories of sound positioned in a hierarchy in which image is dominant. We further discussed other theories of language and philosophy that offer more intricate ways to look at sound's communicative and experiential functions in film. We illuminated some of the shortcomings of existing film sound theories through artistic research in which we created a sound design with sound effects, seeking to tease out a number of ways in which common use of sound design defies simple analysis via the existing theories earlier presented, illustrated by applying a theoretical analysis to the resulting audiovisual piece. Through this discourse, we were able to address our first area of inquiry to show that existing models are simply not adequate while also leaning towards expanded possibilities elsewhere.

Along the way, we have made or shared certain assumptions that we wish to address in both this chapter via literature review, and in Chapter III through content analysis applied to our artistic output. These assumptions form the basis of the research inquiry we here address, our second inquiry of four, which examines whether or not assumptions in the literature regarding intentional *desired outcomes* of filmmakers via sound are demonstratable, and therefore defendable. Here we will first discuss assumptions related to the functions of sound design, then move on to discuss sound design as a technology, defined by *desired outcomes* (or intentional functions). We will finish the chapter with a discussion of specific, hypothesized *desired outcomes* of sound design which will then be queried through content analysis in Chapter III.

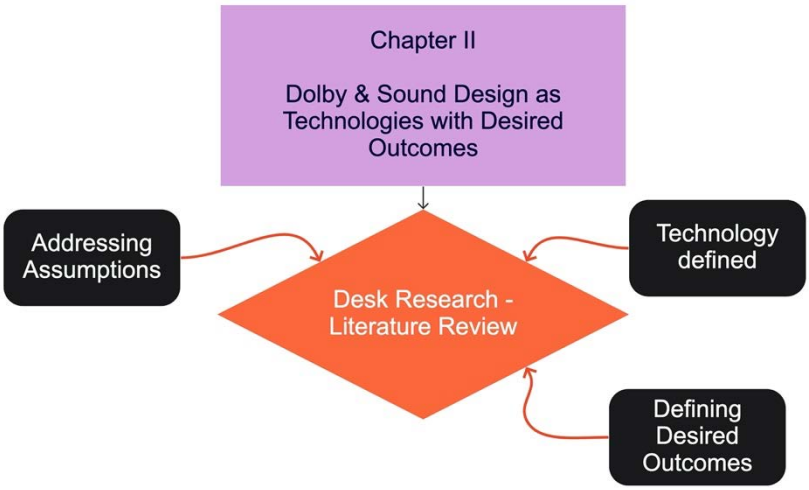


Figure 5: Chapter II roadmap

2.1 Addressing Assumptions

As we shall now briefly revisit, the assumptions that we wish to address are rooted in what practitioners and theoreticians call "functions" or "impacts", or any other such word that implies intentional results from sound design, an idea which is echoed in our underpinnings of the existence of rhetorical structure in the audio and audiovisual text. Some of the assumptions expressed in this work are our own based on my years of experience in constructing and executing sound design for film over the years, working in mixing studios with directors and editors, often making choices based on expectations of results. In that regard, the research is practice-led, iteratively tapping into tacit or praxical knowledge. Other assumptions that we address are either stated or implied in literature on sound design, as have been previously discussed, and will be revisited via diegesis and rhetoric.

2.1.1 Diegesis: Impact over Structure / Function over Form

With respect to conversations on diegesis, particularly on the side of practitioners, let us briefly revisit Taylor (Taylor et al., 2007) and their presentation of various positions on diegesis, where we heard from a number of practitioners who refer to what filmmakers can do with sound. Here, Nordan acknowledged that sound operates at the will of filmmakers when he stated that "assumptions about [diegetic and nondiegetic sounds] sometimes interfere with our understanding of what *filmmakers can actually do* with them" (n.p.). Sound Designer/

Rerecording Mixer Randy Thom added that diegesis is "often irrelevant to the *effect* the story has on its audience" (Taylor et al., 2007). Both these practitioners leaned into the idea of the *effect* as the dominating reason for making choices, and they seemed to assume that the choice and manipulation of sound effects produces results that alter, drive or support narrative and contribute vastly to the non-narrative component of an audience's experience, both in terms of the emotional impacts typically associated with music and the phenomenological aspect of the movie-going experience. They have highlighted the assumptions held by practitioners and many theorists that these effects, impacts or functions are real and are intentional, an assumption with which we shall engage.

2.1.2 Rhetorical Stace and Channels of Communication

With respect to rhetoric, we are reminded that Behrens (1979, p. 3) argues that all fictional films are not only "dramatic constructs" but "rhetorical as well" and, as cited earlier, finds that the filmmakers' job is one of argument, inasmuch as they "influence our thoughts and feelings" (p. 3), and he conjures the *rhetorical stance*, citing Wayne Booth (1963, p. 141) who depicts the *rhetorical stance* as the balance among three elements of communication: the "available arguments" of a subject; the assumed "interests and peculiarities" of the audience; and "voice", defined as the implied character of the speaker. This triangle is seen as encapsulating creative decisions, audiences, and the voice of the filmmakers/sound designers. Our entire discussion of rhetoric, therefore, both theoretical and applied to our artistic research in the creation of our *Koyaanisqatsi* sound-designed object of study, assumes that this communication channel indeed exists, which further implies authorial intentionality. Communication technologies enable open channels to send and receive and we posit that sound film serves as a talk-exchange according to Grice, implying a cooperative principal between speaker and hearer, which we extrapolate in the film context to exist between filmmaker and audience. We have further discussed that this exchange implies inference, argument and meaning and therefore suggests that both pragmatics and rhetoric are at play. We therefore ask ourselves if sound design, on whose functions we have already elaborated with respect to rhetorical stance, is indeed a communication channel of its own, and thereby a communication technology, rather than simply a supportive tool in visual communication. This chapter elaborates this inquiry.

In combining these underlying assumptions that we and others put forth, we uncover a foundational position that sound design (a) is a ubiquitous and diffused technology within the art and science of film, and (b) articulates well with the idea of a rhetorical stance. Through theories of technology to be explored next, therefore, we will examine if sound design is in fact a technology. If it is, it would have to embody and permit desired outcomes from those who use the technology. This chapter directly examines the elements required of a technology and seeks to define some desired outcomes of sound design, as illustrated in our research-led artistic output. Those outcomes will be examined in Chapter III to see if content analysis serves as a platform for illustrating that desired outcomes are met, illuminating rhetorical stance.

2.2 Is Sound Design a Technology?

Technologies are, according to Winston, utterances of a scientific language along the lines of Saussure, that is, "surface expressions of a deep-seated mental competence" (1998, p. 3). Winston elaborates:

> …just as in language a formal understanding of the deep structure of linguistic competence is not a prerequisite of utterance, so too a lack of formal scientific competence is no bar to technological performance. But the technologist will, at some level, have absorbed the science; just as a speaker, at some level, has absorbed grammar (p. 5).

Therefore, as we think of the sound designer or mixer at work, we are aware that they have absorbed the science – the electronic, acoustic, physical and psychoacoustic theory, the gadgetry of analog and digital recording and processing, and most importantly the very grammar of sound in the audiovisual environment—so that they are capable of making utterances. Therefore 'praxical knowledge' (Bolt 2007:34) of the artist enables an utterance of communication via a mastered toolset. This leads us to believe that indeed, sound design may be a technology, but in order to support our position that it is an element of *argument*, framed in rhetorical structures, we need to illustrate what desired outcomes sound design addresses: *what do these utterances do?* This question is fundamental because it encapsulates the idea of a diffused system, both creative and technical, that increases the likelihood of success for desired outcomes.

If sound design is indeed a technology, it is subject to certain definitions and parameters of a technology, the most relevant of which is its ability to predictably generate certain outcomes through the use of the technology. This will be discussed shortly, but first, let us revisit and expand upon some technologies associated with sound design, technologies that are used in our artistic creation and whose predicted impacts are subjects of either direct query or discussion in our content analysis.

2.2.1 TECHNOLOGICAL CONTEXT

Throughout the history of film sound, the technologies associated with it received numerous technical academy awards and were often marketed to audiences in a competitive play between studios who would form exclusive licensing arrangements with technology companies. For instance, Warner Bros. partnered with Vitaphone to develop, market and deploy the most successful sound-on-disc system for film prior to optical track technology, from 1926-1931, premiering The Jazz Singer (Crosland, 1927) on this system, said to be the first talkie (Frayne, 1985, p. 264). Even today, Dolby, Sony, and Xperi compete with their sound storage technologies for 35mm printed film – Dolby Digital, SDDS, DTS respectively – each with their own trailers/logos at the head of films meant to impress with their sound.[52] [53]

Going back through film sound technology via an examination of the Academy of Motion Picture Arts and Sciences' Oscar Awards for technical achievements through the years, we can extract a brief history of major touchpoints (*The Official Academy Awards® Database*, n.d.). The 1930s see a host of awards for optical track recording, microphone developments, and metering, supporting the first decade of commercial sound film. There are very few awards given for sound in the 1940s.[54] 1954 sees a number of awards for the development of magnetic sound recording, magnetic editing machines, and multiband equalizers reflecting the jump to magnetic sound in the production and post-production processes. There is another Oscar in

[52] One trailer from Dolby uses a train in an dramatic graphic to evoke the earlier revolution brought about by motion picture with *L'arrivée d'un train en gare de La Ciotat* (Lumière & Lumière, 1896).
[53] THX, created by Lucasflms and now owned by Razer, is a playback quality control system that certifies theaters for best practices in providing sound true to the original mix.
[54] Of note is an award to RKO for developing and building a reverb chamber, an early indication of sound processing (acoustic rather than electronic) becoming more ubiquitous.

1962 for a highly directional microphone.[55] The 1970s brings new speaker systems and in 1978 Ray Dolby wins an Oscar for Dolby's surround sound and noise reduction system.[56]

To underscore the impact that Dolby Studios has made, we further see that in 1989 Ray Dolby receives an Award of Merit from the Academy for continued contributions to the field. Dolby receives an additional Oscar Award in 1995 for the development of SRD, the Dolby digital sound system with 5.1, which first introduces the subwoofer, stereo surround channels, and a digital format compatible with 35mm prints. In 2003 the Academy bestows a lifetime achievement award to Ray Dolby, and in 2015, after his death, a star in his name is installed on Hollywood Boulevard's Walk of Fame, literally cementing the contribution of Dolby to the motion picture industry.

As noted in Chapter I, a detailed history of sound technology in general and specifically for film would be beyond the scope of this work, but we will briefly highlight what was previously discussed in terms of major steps in the modern era that are most relevant to our study:

i. Dolby Surround – creating a more embodied experience by providing sound sourced from around the theater;
ii. Dolby Spectral Recording noise reduction – enabling a much greater dynamic range from near silence to full loudness along with expanded frequency range across the hearing spectrum;
iii. 5.1 digital and the introduction of the subwoofer – splitting the surrounds into left and right while adding a subwoofer for very low embodied frequencies.[57]

We also again make note of the technical introduction of the subwoofer, producing low frequencies that heighten embodiment in several ways as earlier discussed. This remains relevant due to the ability for these to "whip up emotions", as previously mentioned, a phrase

[55] Interestingly this would be very beneficial for location recording, during a time in the early 1960s when Hollywood was hitting a crisis period, independent films in the US were developing quickly, and influences from European cinema location filming, from Italian neo-realists (who dubbed dialog (*Neorealism*, n.d.)) to French New Wave were being felt.
[56] In 1983 Jack Cashin of Ultra-Stereo Labs, received one for a similar device that is compatible with Dolby in theaters and reconciles some unwanted artifacts in the Dolby system, artifacts which by the end of the 1980s Dolby had engineered away.
[57] As mentioned, Dolby 7.1, 13.1 and ultimate Atmos expand on the ability to pinpoint source direction of sound and immerse audiences.

cited by Beck from film reviewer Ursula K. Le Guin's review of a 70mm print of *Close Encounters of the Third Kind* (Spielberg, 1977) (Beck, 2008, p. 70).

2.2.2 Pathway to Invention, Evolution and Suppression of Radical Potential

The very idea that engineers like Dolby receive Oscar awards for sound engineering is an indication of the intentional outputs of sound design. Just as a cinematographer chooses a differently-engineered grain of film negative to achieve specific results, the sound-for-cinema engineering world has filled gaps and advanced possibilities for creators in the sonic arts. We earlier discussed Dolby and its technical innovations that were accompanied by an explosion of creative output in sound that had previously been constrained by technical limitations: the cumulative noise floor, a limited frequency response, and a single-point source of mono sound emission. Along with the innovations, the concept of a sound design was born. While the term is perhaps overdetermined, since sound design naturally evolved out of existing sound editing and mixing practices, Dolby and sound design were marketed within the industry and to consumers as innovative revolutions. Even today, prominent film sound theorists write of this period as The Dolby Revolution. Michel Chion for instance, referring to the dynamic range provisioned by noise reduction, defines this period as a "quiet revolution"(Chion, 1991, p. 72). Theorist Charles Schreger titles his essay about surround sound "Altman, Dolby, and the Second Sound Revolution" (1985, p. 348).

Here we shall examine this so-called revolution through the prism of Brian Winston's theories on technology, which give us a view into Dolby and sound design as objects of technological innovation. If indeed Dolby and sound design are technologies, having emerged through innovation, then certain parameters of a Winston technology should be supportable, namely that they are *evolutions*, not revolutions, and that there exists a *supervening social necessity*, even in an artistic endeavor, like film (Winston, 1998, p. 12). Additionally, we shall incorporate a discourse between Winston's work and Everett Roger's *Diffusion of Innovation* (1983, p. 15) to force an examination of the social system under which sound innovation in theaters was implemented.

Evolution of Dolby Technology: In the introduction to Brian Winston's book *Media Technology and Society* (1998, p. 1), he asserts that what is "hyperbolized as a revolutionary

train of events" is ultimately a more gradual, evolving process. We can see from the vast body of writings on Dolby's wide acceptance that such hyperbole has been frequently used. Even in 2021, with the introduction of Dolby's new Atmos technology in the music industry, Dolby itself quotes a music producer tweeting, "This is a sonic revolution" (@Dolby, 2021). This self-referential form of hyperbole is not new to the film sound world. Since the advent of sound on film, patent-holders and their studio partners have aggressively marketed a sound *revolution*.

Figure 6: Warner/Vitaphone from Motion Picture Magazine, December 1928

For instance, a print ad for the 1928 Warner Bros. release *Noah's Ark* claims that "Swiftly–surely–dramatically–Warner Bros. VITAPHONE has revolutionized the whole world of entertainment" ("Vitaphone: True to Its Mark," 1928, p. 4). One must view with skepticism such a claim, that each new iteration of sound technology is a revolution. Even for Noah's Ark, there were only two short scenes with poorly recorded dialogue, hardly a revolution against the already well-developed sound narratives of radio. Rather, the use of the term revolution looks more like Winston's description of "a rhetorical gambit and an expression of technological

ignorance" (Winston, 1998, p. 2). In reality, it has been a slow and steady course, with limitations in the recording and playback media themselves putting constraints on progress, constraints around which Dolby (and many others) engineered their way to success, as Ioan Allen of Dolby himself notes:

> Historically the theatre industry buys a 35mm projector, they expect thirty or forty years of life from it, and it will work for fifty years probably…the equipment in the booth is expected to go on forever (Sergi, 2004, p. 104).

With this in mind, Dolby's surround matrix was engineered to accommodate the social or systemic-economic constraint that projectors in theaters globally were provisioned to read a maximum of two tracks of optical audio on a print (p. 105). Even in 1985, Rick Altman queries if theaters would fully adopt Dolby sound systems and "use these innovations to their fullest," highlighting that it remained "a separate – and economically problematic–question" (Altman, 1985, p. 50). These constraints can be regarded as abiding Winston's *law of suppression of radical potential* (Winston, 1998, p. 12), a requirement for an evolving technology. The Dolby matrix allowed four channels (left, center, right and surround) to be encoded, and therefore printed, onto just two optical tracks, so that any standard projector (the constraint) could play it and feed the signal to an external Dolby-patented decoder, with only a very modest upgrade to the projector. Without the need to accommodate such a constraint, "radical potential" might have led to new modes of multichannel sound in the 1970s and 1980s[58] [59]. It was not until digital audio technology became available 10-15 years later with digital sound printed on film, that new potentials like 5.1 would be met.[60] Even here, the old Dolby SR analog track was preserved for compatibility, which allowed for early adopters with digital readers to take advantage of the technology ahead of mass diffusion, aligning with Rogers' innovation adoption curve (Rogers, 1983, p. 11) which begins slowly before accelerating to mass use. Since diffusion is "a social process by which an innovation is communicated over time […] within a social sector" (Dearing & Meyer, 2006, p. 33), for Dolby technology we consider the social sector to be a combination

[58] Others were tried, using magnetically coated multitrack film fed to surround or quad speakers. These did not disseminate into the industry.
[59] Even Vitaphone could not last because it was a separate disk and could not accommodate inevitable changes of sync when distribution prints ripped in projectors and frames were lost.
[60] Later in 2010, Dolby Atmos introduced matrix-based rather than channel-based mixing, with virtually no limit to the number of discrete speakers, which was only made possible by the introduction of file-based prints for distribution rather than physical ones.

of film creators, producers, technologists, and distributors. It seems quite straightforward, therefore to assess Dolby technology as a technological evolution with constraints, but to apply such an analysis to sound design charges us first with asking if a creative process is an evolved technology.

Evolution of Sound Design as a technology: Sound Design has evolved over the years, in tandem with the evolution of the engineering that supports it, as previously laid out. It is relevant to stress that the engineering has provisioned, each step of the way, more latitude for sound design via ever-increasing fidelity, immersion, and manipulability, not just in terms of recording and playing back sound, but in terms of editing and mixing as well. Digital editing and mixing, in particular, have greatly expanded the creative possibilities of sound design by allowing more work to be done in less time, thereby decreasing an economic barrier, and by allowing for virtually an infinite number of layers with which to create new sounds and/or overlap previously recorded ones. Whittington (2013) notes that with digital editing and mixing,

> the nuance and detail of sound constructions [are] expanded, and, currently, it is not uncommon to see hundreds of sound elements aligned on a digital grid in a mixing program such as Pro Tools as they are being used to create one effects sequence (p. 68).

He continues, with respect to immersion, "within this new context, the importance of multichannel formats cannot be underestimated in establishing the new digital sound aesthetic" (p. 68). Thus the evolution of electronics and that of sound design's aesthetics have developed along parallel paths. As for any evolved art/technology, there have been early experimenters as previously discussed, with Chion making special note of Vertov, Hitchcock, Resnais, Lang, Bresson, Kubrick, Welles, Duras (Chion, 2009, p. vii) and there are undoubtedly many others.[61] While some, like Soviet filmmaker Dziga Vertov, took advantage of new technologies like the optical track in his first sound film, *Enthusiasm: Symphony of The Donbas* (Vertov, 1931), experimenting with sound as an asynchronous montage element, other worked outside of new technological developments. Orson Welles, for instance, pushed notions of diegesis and the acousmatic voice (Buckland, 2013, p. 243) in The Trial (Welles, 1962) at a time when nothing

[61] Perhaps the sound editors who worked with these directors should be equally credited and it is curious that Chion is auteur/director-centric

was particularly new in film sound engineering technology, for by then magnetic recording and magnetic-film editing were commonplace, and the film was released in mono optical track sound, a technology fully-adopted 30 years earlier. Welles highlights an evolution of sound design aesthetics that is tangential to an evolution in electronics, just as many trailblazing filmmakers utilize tools in all domains that are not new at the time. Thus sound design aesthetics are, at times, on their own evolutionary path.

Another requisite of a technology for Rogers is the technological duality of hardware and software; software being "the informational base for the tool" (Rogers, 1983, p. 12). Sound designers certainly make use of an information base in order to leverage the tools of modern theatrical sound, again echoing the notion of praxical knowledge. We have already demonstrated in Chapter I that this is done by employing myriad sound effects of many sub-types, music, dialogue and unlimited mixing techniques which form the underlying language and grammar of sonic signifiers and/or rhetoric, as we argue, within the audiovisual experience. This information base, which additionally includes knowledge of film sound theory and the knowhow for its implementation, constitutes software. As for hardware, sound designers, namely editors and mixers, employ complementary technologies in order to accomplish their goals; electronic processors, audio devices, complex sound editing software, etc. These tools have evolved, and so has sound design, traceable from live-sound-in-theater during the silent era and through the years, with many stages we have already discussed. This combination of hardware and software makes up the "absorbed" science of the technologist, as expressed by Winston at the opening of this chapter (1998, p. 5). He likens this to a linguistic utterance, provisioned by a speaker having absorbed grammar. The grammar of both the technology and the medium, therefore, have evolved, and from that base grammar the engineers and the sound artists perform their utterances.

Lastly, as previously mentioned with respect to Dolby, technology or innovation requires an evolution that has been constrained by "suppressors of radical innovation". We argue that a number of factors have served that role for sound design, including first and foremost the gradual evolution of the technical capacities of sound for film accompanied by the economic constraints in theaters as previously discussed. But we also argue that other factors like aesthetics and cultural biases come into play as suppressors of radical innovation. In an interview with Serbian-Portuguese film sound mixer Branko Neskov regarding shackles to the development of a culture of sound design in Portugal, he noted 12 barriers, some technical like

poor film stock, poorly-maintained and aged equipment, etc., but others not, including low funding to sound departments, short turnaround times for sound editing and mixing, funded scripts that offer limited aesthetic space for sound design, and creative resistance among directors[62] (B. Neskov, personal communication, March 5, 2021). The cultural aspects are also highlighted by Altman (1985) when he points to "historical and ontological fallacies" that have been perpetuated within film criticism (pp. 51-52). The "historical fallacy" leans into chronology, privileging image over sound because the poetic image film came first, ignoring the fact that sound films "are composed of two simultaneous and parallel phenomena, image and sound", and then, in turn, marginalizing sound. The "ontological fallacy" is the inherent claim that film is a visual medium. Altman argues that the perpetuation of the image-oriented stance has "failed to provide either the theory or the terminology" for sound cinema as it exists. From this we extrapolate that, to the extent criticism is part of a feedback loop with the industry, as argued by Peter Lehman in "Politics, Film Theory, And The Academy" (1988, p. 43), advances in sound design have been constrained by the image-based positioning of both academic and cultural writing and further constrained by creative resistance among directors and the screenwriting/greenlighting process.

With this we have established, for both sound engineering/technology and sound design, a technology with an evolution that is subject to a suppression of radical potential. However, we still need to understand what this technology accomplishes and for whom. According to Rogers (1983), a technology is "a design for instrumental action that reduces the uncertainty in the cause-effect relationships involved in achieving a desired outcome" (p. 12). How does sound design "reduce uncertainty […] in achieving desired outcomes," thereby abiding Rogers' definition of a technology? We believe that the desired outcomes lie in Winston's assertion that an innovation or prototype must have a "supervening social necessity" in order to undergo a transformation into a diffusible invention (1998, p. 8). First we will address the social necessity, for both Dolby and Sound Design, and how uncertainty for desired outcomes is reduced by them. We will then look at defining those desired outcomes.

[62] Neskov notes that some of these barriers still exist, particularly those around time, money, and directorial resistance, noting that only about 50% of Portuguese directors today are open to what sound can do for their films, pointing to an everlasting dominance of visual culture.

2.2.3 WINSTON MODEL – REDUCING UNCERTAINTY AND THE SOCIAL NECESSITY

We must remember that fundamentally, global *commercial* cinema is a storytelling medium and, as such, there is a contract between the audience and the storytellers; the writers, directors, editors, cinematographers, sound designers and others. That contract is an interactive one, one in which audiences take part in a feedback loop of decision-making at many stages of the process:[63] scripts are approved based on targeted audience assumptions; actors are selected based on potential box office, likeability, physical characteristics or fan-base, in an explicit play to audience expectations and desires; shots are often framed with the gaze of the audience in mind; music, especially popular music, is selected to engage a targeted audience; versions of a film are screened for test audiences who give feedback via surveys and/or focus groups and influence anything from the editing of a scene to a re-shoot of a new ending. As a commercial enterprise, the audience buys the tickets and they are therefore a conscious part of the consideration for commercial film decisions. For film creators and their creative staff who utilize technologies, *to reduce uncertainty in desired outcomes* is to address narrative and experiential elements of reception; for the goal of the storyteller is to provide both a *controlled narrative* understanding and an *experience* that leaves the listener-viewer entertained at worst, transformed at best, and recommending the film to friends and family. As earlier mentioned, we like to express this idea of piloting an unfolding engagement as *curating the flow of emotion and cognitive perception*, which, in rhetorical terms, might be the filmmakers' *argument*. Over the course of a single film, these desired outcomes may operate as a dance between cognition and experience, and will likely be a different dance for two listener-viewers sharing a popcorn while mesmerized by sight and sound. If we can demonstrate that the uncertainty of achieving communication goals – *curating the flow of emotion and cognitive perception*, conveying argument – is reduced by sound design, we can show that Rogers would define film sound as a *technology* and developments in sound design as *innovations*.[64] It is critical for sound design to assert that an innovation does not need to be engineering-based. It can be "an idea, knowledge, a belief or social norm, a product or service, a technology or process, even a culture, as long as it is perceived to be new" (Dearing & Meyer, 2006, p. 34). Sound design encompasses new

[63] Suggesting that even the "interactive revolution" is but an evolution of storytelling, not a revolution
[64] In a footnote, Rogers notes that they often interchange the words technology and innovation (p. 12)

ideas, new knowledge, new processes and possibly even new culture and therefore its evolutionary potential is met via gradual innovation.

Moreover, if a *supervening social necessity* can be defined for sound design, we can show that Winston's model of diffusion applies not only to Dolby's electronic advances, but to sound design as well. Winston includes the assertion that a supervening social necessity can vary in a spectrum from a true societal need to the "subjective whims of perceived needs" (1998, p. 6). By this definition, we ask if sound design, along with its typical use of immersive and emotive devices, arouses the emotive centers of memory and provides a richer, more memorable experience rooted in phenomenology and thematic interpretation, or Walsh's "immersive spectatorship", along with various other types of cognitive engagement. One can interpret that as fulfilling a need for social communion or entertainment. Equally, it can be regarded as satisfying the needs of the creators to tell their stories in the most compelling and curated ways. But are these truly needs? We take issue with Winston's term "need" since he goes so far as to permit a "subjective whim," rather we believe that a supervening "need" leans towards hyperbole. Rogers (1983) takes a different approach, describing four characteristics of innovations, one of which is "relative advantage" (p. 15), an approach that focuses on what an innovation *provides* the social sphere, rather than looking at what the social sphere *needs*. With Rogers, we need not argue that entertainment and storytelling engagement are *needs,* for his approach to innovations allows for perception and desire to play a role on purely Lacanian grounds. Rather, we must only explore if sound design offers a relative advantage in *providing* audiences with entertainment and engagement. We can therefore more simply question what sound design provides the social sphere, what relative advantage. These relative advantages will be our *desired outcomes*, which sound design reduces the uncertainty of achieving, within the social sphere of creator and audience.

2.3 SOUND DESIGN: HYPOTHESIZED ADVANTAGES / DESIRED OUTCOMES

Applying technology theory to Dolby is rather straight forward and, since it easily fits within the paradigms of technology and diffused innovation, there must be value in what it offers the social sphere: reduced noise, sonic immersion, low frequencies. Here we shall propose a number of hypothesized advantages not of Dolby but of sound design, as expressed

in our *Koyaanisqatsi* artistic output, in order to examine predicted outcomes to the social sphere through content analysis in Chapter III. These can be thought of as impacts, functions, or relative advantages, which collectively we call *desired outcomes* of sound design. We present them below, with a brief discussion of each. If we can show that the uncertainty of achieving these desired outcomes is reduced by sound design, then we will have illustrated the one remaining component to support that sound design is a technology, or innovation, that is diffused because of these outcomes, moving beyond the assumptions that we find in the literature and as stated by practitioners.

Based once again on praxical artistic knowledge, we interrogate if sound design provides creators and audiences – the social sphere – the following advantages:[65]

1. Increased attention and focus by audiences;
2. Support for narrative understanding and/or thematic interpretation;
3. Physical and temporal continuity (hiding of the apparatus);
4. Thematic continuity;
5. Imagined sound (engagement of mechanisms of phantom sound);
6. Experiential engagement or impression;
7. Emotional engagement.

Let us examine these in more detail.

(1) Increased focus and attention in audiences: Amy Herzog (2013) notes, in her sound analysis of *Sleep No More* (Punchdrunk, 2011), an immersive audiovisual live theater game, a hybrid intersection of real-world embodiment and representational narrative, that sound is an important force that attracts an audience's attention:

> The strategic use of sound was most apparent in the moments just before a scene was to commence. Elements in the recorded sound design would shift, sometimes with the introduction of a new musical element played at an increased volume or more often through atmospheric shifts induced by low-frequency chords. The performers' voices intruded in staccato bursts at these moments as well, as did the sounds of their movements, slamming objects or other characters with reverberant force. Depending

[65] As will be discussed in Chapter III, the null hypothesis is simply that sound design does not provide creators and audiences any of the mentioned advantages.

> on where one was within the installation, these sounds served as attractors, turning heads and luring curious wanderers into range" (pp. 203-204).

This example illustrates that the sound magnet works in both our living environment and mediated space, and sound in the cinema is often used for similar effect. Countless films, like *Velvet Goldmine* for instance (Haynes, 1998), introduce sound before the first diegetic visual frame appears, with music and/or sound effects drawing our focus to the screen-based world about to unfold. During the course of a sound film, sound adds weight to visual objects, grounding them in diegetic reality. Pre-lapping[66] sound into a change of scene draws our attention to a new diegetic space about to unfold. "Materializing Sound Indices (MSI)", as mentioned earlier, which Chion defines as "aspects of a sound that make palpable the materiality of its source and of the concrete conditions of its emission," like breathing, mouth pops, rasp or the sharpness of walking on gravel; these keep the audience tied to actions and character, and represent the "rendering of the real" (Chion, 2009, p. 480). For instance, in *A Man Escaped* (Bresson, 1956), the very textual MSI sounds of scraping during the protagonist's preparation to escape keep us glued to the danger highlighted in the action, as well as the danger the sound itself poses. Conventionally, over the years, sounds in films have been heightened through various means to seize the audience's attention, from harsh loudness to lowering nonessential sounds, or via stereophonic directioning techniques (Mancini, 1985, p. 363). For instance, a "c/omission" (Chion), as discussed earlier but between sound and image, can draw the audience's attention to a character's psychological condition even with a motionless shot on a steady face, as previously discussed citing *Living Out Loud* (LaGravenese, 1998), when background sounds of diners, silverware and a live piano slowly disappear, creating a void that registers the emotional pain of a character and also scans the moment rhythmically. We will use the phrase *focus and attention* in order to capture these varied ways in which sound steers the audience either to the screen, to the narrative, or to a particular object or person within the frame. Certainly, this is a *desired outcome* for filmmakers, as for any storyteller. If we are to *curate the flow of emotion and cognitive perception* in a controlled manner, we must demand both attention and focus, precisely when and where we want them.

(2) Support for narrative understanding and/or thematic interpretation: Again we put forward the idea of *curating the flow of emotion and cognitive perception* in a controlled

[66] Fading up on the sound of an incoming scene before the visual cut is made to the scene.

manner. In some way, this relates to attention, for as narrative unfolds in a constant flow, sound draws attention to objects and characters along the narrative trail, bringing clarity. Voice, music and sound effects have all been shown to support, contain or propel narrative. Dialogue, for instance, is obviously essential to narrative and sound design – including microphone choice and placement, dialogue editing, and mixing techniques guarantees its clarity when desired. However, we are including in this notion the idea of thematic clarity. Oftentimes films unfold more than simply a didactic plot, but rather engage the listener-viewer in broader themes. The digital blockbuster *Avatar* (Cameron, 2009), for instance, has a simple plot that can be recounted easily, but the themes it involves include messaging around ideas of extractive capitalism, love, interracial relations, environmental degradation/preservation, colonialization, and more. In most instances, we believe, filmmakers want narrative plot to be clear as it unfolds, and for themes to be either intrinsically or extrinsically accessible. We therefore put them into the same category of desired outcomes, particularly due to broader definitions of narrative that might include thematic representation, as will be discussed in Chapter IV.[67]

(3) Physical and Temporal Continuity, or hiding of the apparatus: Tellingly, in the glossary of Bordwell and Thompson's *Film Art: An Introduction* (2010), continuity editing is defined as "a system of cutting to maintain continuous and clear narrative action" (p. 490). It goes on to give a host of examples and related techniques, supplemented by an index entry (p. 501) with more than 24 subindices, but not one of these examples nor subindices mentions sound. Yet, continuity of sound is critical for audiences to believe that that the various shots in a scene unfold continuously in time and space, that is, within the diegesis of a scene. To have discontinuous sound is to break the illusion of continuity and undermine the best continuity picture editing. In a scene with dialogue, continuity begins on set with mic placement, but it is the work of the dialogue editor, polished by mixing, that creates continuity from the chaos of many microphones, ever-shifting ambiances from shot to shot and angle to angle, mismatched performances, and technical malfunctions. It is no surprise that Bordwell and Thompson forget the role that dialogue editors play in continuity, for as David Lewis Yewdall, sound editor and author of *Practical Art of Motion Picture Sound* (2012) notes, "Dialog editing is an invisible art form. If the work is good, the audience never notices" (p. 363). Of course, continuity through

[67] Some traditions of narratology would call plot narrative, others find narrative to be much more embracing, including themes. While narratology as a full subject is beyond the scope of this paper, we will address it more deeply in our discussion of Genette in Chapter IV.

sound is provided by more than dialogue. The placement and manipulation of continuous sounds that are meant to be present in a scene, which can include people, nature, environmental tones, traffic, source music or any number of continuous sounds like a distant police siren placed over an otherwise discontinuous edit, are but a few examples of sound's important contribution to physical and temporal diegesis. This harkens back to our discussion in Chapter I of the difficulties of a sound semiotics due to the temporal nature of the sonic signifier. Sound occurs over time, so if we hear a continuous sound, we are immediately more likely to believe that what we see during that sound is continuous time, and space, even with numerous visual edits. This is a critical part of the traditional aesthetic of hiding of the apparatus. Some less traditional forms, like the French New Wave, for instance, have intentionally subverted this process in both image and sound. Alan Williams notes that Godard extends this to include both dialogue and ambiance when stating that "far from concealing the sonic transitions to his location scenes, Godard emphasizes them" as opposed to the Hollywood tradition in which the dialogue and ambience editing go "unnoticed by seeming to answer to the requirement of the fiction" (Williams, 1985, p. 337). Goddard is put forth as an exception to the rule that overall, commercial films and most independent films follow rules of continuity the majority of the time. Therefore, we posit that sound design offers an advantage by helping to establish (if not establishing outright in many instances) physical and temporal continuity.

(4) Thematic Continuity: We coin this phrase as an attempt to capture the ways in which images can be brought into a thematic relationship by sound. We wish to differentiate this from thematic clarity as earlier defined, by stressing that here we are specifically addressing connecting images that might not otherwise be connected in time and space, that don't share a physical diegesis, but share more of a metaphysical space, metaphorical space or a space of theme. Music frequently serves this purpose. New Yorker music critic Alex Ross notes that "music can take control of the image" (Ross, 2005). He elaborates by citing, for instance, *Lost*, a television drama, crediting the menacing music of Michael Giacchino with conveying the theme that "surfaces are not what they seem." He further attributes a number of thematic ideas to Phillip Glass' music in *Koyaanisqatsi* (Reggio, 1982): loneliness, an angry God, decay, mania, beauty. As we have seen, the singular music cue in our sound-designed sample from the film connects vastly different shots, locations, and times. However, we propose that sound design exclusive of music can also serve this function by bridging disparate shots in a similar manner to music. We have previously discussed the intention of the metaphorical telephone

ringing in our research-led artistic sample, a sound effect that effectively creates connective tissue through individual shots of people along a theme – in this case disconnectedness, for example. We also cited from our artistic work the sound effects over aerial images of cities, tapestries, and computer chips noting that thematically these are further drawn together by ambiance and repetitive eagle screeches in order to establish not a physical connection but a thematic one. This can also occur over the course of a film, without the requirement that shots are contiguous. For instance, Hanlon notes that in Bresson's *Lancelot du Lac* (Bresson, 1974), a sound will "reappear throughout the film as a haunting reminder of earlier scenes and its symbolic function there" and gives examples including horses, armor, gagging on blood, ominous crows, and others (Hanlon, 1985, p. 331). Thus, in *thematic diegesis*, we use the term *thematic* to imply narrative beyond plot, and diegesis to evoke the sense of a *connectedness,* not of space-time but of idea. We suggest that there is a specific universe of thematic ideas embedded in a work, similar to the universe of the story, or diegesis, but along thematic lines rather than physical ones. We further maintain that sound design has the capacity to materialize these themes, or at least reduce the uncertainly that they will be related to audiences, by constructing or supporting thematic diegesis.

(5) Imagined sound/engagement of mechanisms of phantom sound: In examining the silent film period, Chion puts forth the notion of "deaf cinema" to suggest that "there were words and noises, but they could not be heard" (2009, p. 3). Beyond the pantomime of unheard speech, sound was at times "suggested" by shots of things that make sound, as in *Strike* (Eisenstein, 1925) with repeated closeups of a factory siren blowing (p. 5). Chion later cites Czech journalist Milena Jesenká describing a scene from the silent melodrama, *A Woman of Paris* (Chaplin, 1923) that poetically and emotionally moves her due to a shot of moving shadows over a character's face. Chion concludes that the interpretive impact of the shot comes from shadows that "evoke a train we do not see and the noise of train's movement we do not hear" (2009, pp. 180–181). We will call this phantom sound "imagined sound," defined as a phenomenon in which we expect to hear something with a visual stimulus and, even though there is no sound, we believe that we have heard it, or imagine it. MRI studies have shown that if sound is expected but not heard, along with a corresponding visual trigger, the secondary auditory cortex, a location of sound processing that engages with cognitive function, is triggered

as if the sound were actually present (Hughes, 2001)[68]. Our position here is that sound design, particularly the presence of sound effects generally provided for onscreen moving elements, sets up the expectation to hear, and therefore the omission of sound for a visual stimulus better triggers imagined sound.

(6) Experiential engagement or impression: We further explore that sound design contributes to the experience of audiences in some ways that are a bit difficult to describe. We use the phrase *experiential engagement or impression* to encompass an idea that leans more towards phenomenology as a result of creating memorable experiences, or moments, the kinds of filmgoing experiences that are carried by audience members and, hopefully, shared with friends and families. Memory certainly plays a function here, or rather, impressionability. In a paper on episodic memory and phenomenology, Perrin et al. (2020) argue that episodic recollection is a "metacognitive feeling-based account" of phenomenology (p.2). Thus for us, memory of an image, or a moment, or of a thought process like constructed meaning is an indication of impact and phenomenological engagement, and it can be tied to emotion, metacognition and haptic feeling, all of which are excited by immersive sound design. One study utilizing EEG's to measure the visual memorization process, broken down first into easy and hard tasks, and then into two components related to memorization, attention and sensory processing, showed that the sound of white noise increases visual memory equally across this matrix while Mozart music increases attention only during an easy task, and increases only sensory processing for a difficult task (Daud & Sudirman, 2016). It does not surprise us that noise, representing the full frequency spectrum of hearing and constant over time, might have broader implications than Mozart, particularly because different frequencies and volume have been shown to affect attention, cognition and emotion to varying degrees (Cuadrado et al., 2020). These results point to sound having a capacity to alter visual memory and memorability, concepts that we see as byproducts of phenomenological or experiential engagement.

(7) Emotional engagement: Film composer Mel Lewis notes that "music's primary function" is emotional. He posits that dialog, sound effects and music correspond to "Plato's three source of human behavior– knowledge, instinct and emotion" respectively and that as emotion, music "instills the spirit found in the heart" (Yewdall, 2012, pp. 474–475). We in turn

[68] In fact, visual stimuli feed into the secondary auditory cortex (Hughes, 2001) and auditory stimuli (whether real sounds or imagined sounds) feed in the lower (early) visual cortex (Petro et al., 2017).

posit that the emphasis found in film literature crediting music as the driver of emotion overlooks fundamental aspects of sound that should extend the emotional role to all sound. A 2020 study published in Frontiers in Psychology (Cuadrado et al., 2020) cites a number of studies demonstrating that emotional responses correlate to frequency, volume and clarity – the latter as determined by high-frequency content in the timbre of a sound – while other studies show that only 20-25% of emotional response is subject to these physical characteristics of sound, suggesting that sound content – rooted in the signified, interpretation, and meaning – plays a more significant role. This study also cites other studies that show, via electrodermal response, that emotional indicators like heart rate and temperature have a greater response with sound effects compared to silence. Moreover, another study cited indicates that levels of immersion, as determined by choice of headphones or speakers and their placement, correlates to emotional response. With all these empirical data related to sound and emotion, complemented by literature that assumes sound plays a role in emotional response in film, we maintain that indeed emotional engagement it is a desired outcome of sound design that serves the social sphere.

Having elaborated on the advantages, we would like to return to personal experience related to mixing sound for films, as it is relevant to our understanding of these desired outcomes. As a mixer, a film comes into the mix silent, that is, the visual element is separate from the sound elements. The typical approach is to first mix dialogue, put it at appropriate levels, equalization, consistent room reverberation, matching angles, etc., to make a unified sound that supports continuity and anticipates audibility based on expected other sounds. From this, we understand what a scene is supposed to accomplish narratively, structurally, and emotionally if, indeed, the scene has dialogue. Then we prefer to add background sound effects – the ambiences and tones that make interior or exterior spaces real. As these enter, even a mixer can start to feel the emotional and immersive play increase, along with the opening of offscreen space. It's a drastic change to the sound, and the faders never stay flat, they are massaged by some kind of artistic instinct, perhaps tied to praxical knowledge, to always be dynamic and highlight narrative, action, emotion, or all of these. Then come the foleys, which mostly serve as MSIs (Materializing Sound Indices (Chion, 2009, p. 480)) which render on-screen motion real, provide further continuity glue and create a kind of intimacy with the characters and their physical world. What is that intimacy brought about by the sound of

clothing rustle as a character turns? Is it narrative, or emotion, or punctuation, or a kind of metaphysical immersion? We are not sure we yet have the theoretical language to address that, but what we can express is that a mixer, and everyone else in the mix room, can feel with these foleys and background sounds a shift in the balance between representation and real. A typical response as this unfolds is for someone in the room to say aloud "wow, it is really coming to life". It is often at this point that directors, who for the first time "hear" their film, or more accurately "hear-see" it, experience either euphoria or panic as they literally realize they have a nearly completed movie that is soon to be shown to the public. Then other sound effects are added, imbuing many of the rhetoric figures we have discussed, like scansion, metaphor, counterpoint, contrast, as well as further materializing images with sound, expanding immersion, and evoking thematic continuity, particularly across the film. Finally, all the elements are remixed against music which, at a minimum, reinforces rhythm and emotional engagement. This, of course, is a simplified account. But the point here is to share that only in the mixing studio can one literally watch and hear these desired outcomes evolve before one's eyes, around one's ears, and as an embodied experience, sound by sound, layer by layer, dare we say utterance by utterance, even before music is added. Perhaps this underscores the overriding motivation embedded in this research and previously put forth in its introduction; by addressing desired outcomes of sound design and expanding a theoretical discourse to discuss them, we might bridge a gap between theory and practice that seems not to be captured by either in isolation, but that for an artist/researcher represents the gap in which we live.

With the hypothesized advantages upon which we have elaborated, we move into Chapter III where we will interrogate these advantages via content analysis using the *Koyaanisqatsi* sound design segment, continuing to address our second research inquiry. If indeed we can show that some or all of these advantages are met, or that the uncertainly of achieving them has been reduced, we will have addressed assumptions in the literature and will have shown sound design to be a technology that provides the social sphere with these advantages, further justifying the search for other theoretical paradigms that might better capture sound design's functions while spurring additional research in the area.

CHAPTER III. CONTENT ANALYSIS: ARE DESIRED OUTCOMES DEFENDABLE?

In the previous chapter, we looked at both Dolby and sound design as technologies. While examining Dolby via Winston and Rogers was a straightforward exercise, sound design appears more complex. For Dolby, the advantages, or desired outcomes, offered to the social sphere are fidelity, reduced noise, low frequencies, and immersion. For sound design, based on theory discussed in Chapter I and further articulated through our research-led artistic piece and personal experience, we have established seven desired outcomes that, if shown to be true, support further research in a number of domains: empirical neuro-cognitive studies, studies of rhetoric and underlying linguistic structures, a search for new paradigms to examine various physical and psychological effects from sound design, a phenomenological study of sound design, etc. This therefore continues our inquiry into our second research inquiry: are the assumptions defendable? For our research, it is critical to establish that the assumptions that sound design works in specific ways, the *functions* or *impacts* or *effects* that appear in literature and that are thrown around the mixing stage among peers, in fact exist. Only then can move forward with our research to examine new rhetorical structures that might frame them.

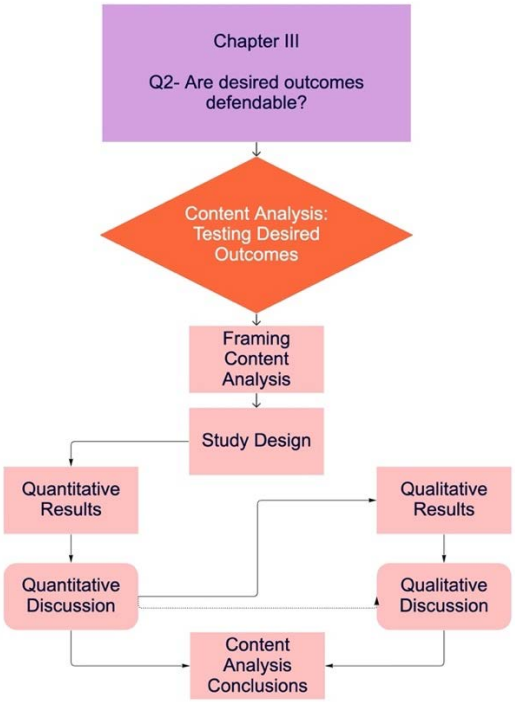


Figure 7: Chapter III roadmap

As evident in the above figure, after presenting the study methods, we will move on to quantitative results first and discuss them. We will then move on to qualitative results and discuss them, while oftentimes bringing them into conversation with the quantitative discussion along the way, as they become rather interrelated. From this combined discourse, conclusions and additional thoughts will be drawn. But first, we will frame our content analysis in its own theoretical underpinnings and justifications.

3.1 FRAMING CONTENT ANALYSIS

With our inquiry defined, that the uncertainly of success in achieving the aforementioned desired outcomes is reduced by sound design, we move forward with the design of the content analysis. But first, we wish to discuss, from a theoretical basis, why content analysis is appropriate and our overall philosophical positioning around it. To frame this discussion, it is sufficient at the moment to summarize our approach: we will show different audiences different versions of our artistic work, the *Koyaanisqatsi* excerpt. One set of audiences will screen the original with only music and image while the other set will screen the version with sound design as already detailed in Chapter I. They will complete identical surveys with both scaled and open-answer questions so that differences in outcomes, should they exist, might be measurable or become otherwise evident through data analysis. More details of the design will be provided in section 3.2, but first let us frame our approach to quantitative and qualitative measures.

3.1.1 CONTENT ANALYSIS GENERAL OVERVIEW

We approach content analysis with an acute awareness that some of the tested desired outcomes embedded in the sound design of this excerpt may not be measurable due to the very nature of measurability of responses to art. For instance, a deeper study would take into account any number of domains: audience studies, narratology, subjectivity, experiential memory and recall, and even the nature of surveys as a recall trigger, because all of these domains play a role in our endeavor to work with audiences, ask them to remember moments from a screening,

and complete a survey. There are bodies of study around these domains that introduce variables that are very difficult, if not impossible, to remove and beyond the scope of this study. Yet, regardless, this research has sought to structure a survey that might illuminate correlation between independent and dependent variables via data analysis while also offering insights into desired outcomes and audience response via qualitative discussion. Again, details of methods will follow, but first let us elaborate on the nature of content analysis, its application in the arts, and justification for its relevance in this study.

To begin with, pursuing a path of content analysis implies an alignment with positivism, for metaphysical assertions that are made about sound in film do not stand up as foundational pillars of future research in a domain that is increasingly populated by neuro-cognitive research. If we are to acquire knowledge in this sphere in the contemporary research framework, we believe the application of scientific method is warranted (Routledge (Firm), 2000, p. 696) and further positions this research as a springboard for additional work.

This content analysis is both appropriate and somewhat unusual. In artistic fields, this type of scientific approach as part of a larger theoretical exploration is uncommon. Yet in the context of artist research, Smith & Dean argue, “the unique combination of creative practice and research can sometimes result in distinctive methodological approaches, as well as exhilarating findings and artworks” (2009, p. 5). On the quantitative side, we are fundamentally exploring if responses to art are measurable and, to that end, we examine if our desired outcomes may be statistically measurable precisely because we can identify sound design as a variable for which we can create a control condition, namely no sound design, and an independent variable one, namely with design, and can apply a numerically scaled survey for some questions. A traditional scientific approach seems well positioned as a tool of such an inquiry. Qualitatively, for those outcomes that are based on open-ended questions, data analysis on text provides an additional means of measurement and, more importantly, informs insightful paradigms of discourse, which is the heart of the exploration. Such a hybrid approach at this juncture in the research, which involves a number of iterations in terms of revisiting our artistic output as well as putting our qualitative and quantitative discussions in an iterative cycle, sits well in our uniquely evolved path of artistic research exploration.

3.1.2 FRAMING QUANTITATIVE VS. QUALITATIVE RESEARCH

As previously stated, we are able to conduct quantitative research because we have the conditions for an independent variable with a control and will offer scaled questions to our audiences, therefore statistical methods will shed light on results by directly interrogating *correlation*. However, it is difficult to structure a numerically-scaled survey that sheds light on those desired outcomes that relate to interpretative and emotive impacts. For these, we utilize open-ended questions and qualitative analysis; but how shall we measure? Can we correlate? Can we infer?

As asserted in the introduction to the Sage Handbook of Qualitative Psychology, Willig (2017) cites Brickman[69] in noting that:

> [c]ontrary to what is sometimes asserted, science is a question of aim, not method. Science is an effort to make accurate observations and valid causal inferences, and to assemble these observations and inferences in a compact and coherent way (p. 6).

So if "science is a question of aim", how might we define our aim in non-scientific terms? The knowledge we seek to gain is rooted in the ways in which sound effects function as a communication technology between filmmakers and audiences, as discussed in Chapter I in our look at pragmatics. Some outcomes are measurable via self-reflection in a numerically scaled survey, but others are more problematic to measure as they relate to meaning-making, a deeply subjective process. We aim to better understand this relationship by making "accurate observation and causal inferences". Willig continues,

> from this perspective science does not have to be defined solely in terms of the hypothetico-deductive method. Other – qualitative – methods can be equally rigorous and valid (p.6).

Indeed, we have defined a hypothesis, or rather a position, as earlier stated, but it is in our qualitative discussion where we may depart from the strictly deductive method, aligning our inquiry more with Merriam (2009), who states that,

> All qualitative research is characterized by the search for meaning and understanding, the researcher as the primary instrument of data collection and

[69] Willig cites (Brickman, 1980, p. 10)

> analysis, an inductive investigative strategy, and a richly descriptive end product (p. 6).

It is from this viewpoint that we engage in qualitative inquiry as a strong contributor to artistic research, with a defined "aim" and an iterative process that hopes to enunciate results that increase knowledge, expand perspectives, and raise additional questions. It is an "inductive investigative strategy" and our analysis is a descriptive one that seeks to "describe poorly understood phenomena…understand differences between stated and implemented theories, and discover thus far unspecified contextual variables" (Marshall & Rossman, 2016, p. 161). Together, this informs our orientation towards the qualitative aspect of our study.

There are different types of qualitative research that researchers use when doing a study, and Denzin and Lincoln (2005) broke these down into eight research strategies: case study, ethnography, phenomenology, grounded theory, biographical, historical, participatory, and clinical (p. 6). According to them "qualitative research does not belong to a single discipline nor does qualitative research have a distinct set of methods that are entirely its own". Since there are several qualitative research strategies it is essential to stress that our research utilizes the case study strategy to arrive at its conclusions, the object of the case being our research-led artistic output. Furthermore, as researchers, our "search for meaning", again citing Merriam (p. 6), is a search to understand, on the surface, if desired outcomes are met by sound design but, more generally, to better understand the overall impacts of sound design on audiences. Since some of our tested outcomes interrogate interpretation by the audience of the audiovisual text, and furthermore require them to write about it, we are obliged to examine the meaning-making of our listener-viewers. Thus, we have a nested meaning-making at play: theirs as an object of study and ours as qualitative researchers engaged in Merriam's "inductive investigative strategy". With this in mind, we next will open a discussion of hermeneutics as it informs our approach, abilities and limitations in qualitative analysis.

3.1.3 Hermeneutics, Targeted Meaning, Memory and Phenomenology

As will be elaborated upon in the detailed study design to follow, we present a survey to audiences who have had some kind of audiovisual engagement, or experience, and then ask them to recollect moments of that experience, providing images in the survey to help them connect to moments or episodes of cognitive or emotional processing. We later code those

written responses in order to extrapolate themes that emerge from the audience's meaning-making, as expressed in their remembered moments. As earlier cited in our elaboration of desired outcome number 6, *experiential engagement or impression*, episodic memory, according to Perrin et. al. (2020), is "an epistemic feeling," and a "feeling-based metacognition"[70] (p. 6). When one is recollecting an episode from the past, one does not simply *know* that the referenced event – or in our case, a thought or feeling – happened; one rather returns to it with a degree of awareness. Elaborating and quoting Tulving, Perrin adds that "to episodically remember is 'to consciously re-experience past experiences' (Tulving, 2002, p. 6)". This can be very difficult for some subjects in the context of a screening room, peers finishing surveys before them, tiredness by the end of a survey, or myriad other reasons that might present obstacles to "consciously reliving experience". With the specificity of the environment, niche theory, which interrogates the relationship between environments and phenomenology, and therefore between environments and episodic memory, articulates that when we query the interaction between mediated cognition and emotion, the "environmental scaffolding" is as much an element of the audience's phenomenological apparatus "as are the three structural tentpoles of film phenomenology: film, spectator and other spectators" (Hven, 2019, p. 7). Indeed, this environmental scaffolding not only includes a physical space. One may respond differently when surrounded by peers, or in the presence of a professor, or alone in a living room. While a detailed discussion of niche theory and memory studies are beyond the scope of this work, they will be addressed briefly where relevant in our analysis of results.

Moreover, we are doing interpretive work, and here hermeneutics on our part comes into play. We seek, in part, to measure cognition or to expose interpretation by the listener-viewer of an audiovisual textual object, to see if *targeted meaning* conveyed via sound design is reached by the audience. Subsequently, we interpret written text, expressed by our subjects on a questionnaire, through a coding process with codes that we believe encapsulate targeted meaning. Furthermore, we attempt to interpret and defend the results that the coding provides. In her article on the hermeneutics in film phenomenology (Sera, n.d.), Mareike Sera cites Deleuze's thoughts on Italian neorealism where "The real was no longer represented or reproduced but 'aimed at'" (Deleuze, 1986). This is a prescient statement in light of our work

[70] Metacognition is the awareness or analysis of one's own learning or thinking processes (Merriam-Webster, n.d.-b).

and articulates with Brickman's earlier-cited statement on the "aim of science". We are seeking to elicit from audiences a connection to the real that they may have experienced as "immersive spectators*"* (Walsh, 2004), expressed through the symbolic representation of words on a survey, and further symbolized through our coding mechanism into themes that might represent the real. Moreover, as expressed on Stanford University's digital Encyclopedia of Philosophy project, when we undertake the task of analyzing experience through a phenomenological lens,

> we classify, describe, interpret, and analyze structures of experiences in ways that answer to our own experience. How I see or conceptualize or understand the object I am dealing with defines the meaning of that object in my current experience. Thus, phenomenology features a study of meaning, in a wide sense that includes more than what is expressed in language (D. Smith, 2018).

This is an important idea upon which to elaborate for it puts us, the researchers, at the center of the research, also a fundamental component of artistic research in general. This approach recognizes the role that we, as mixer/sound designers, play throughout this research, and justifies our occasional self-reflections. With respect to the qualitative analysis specifically, how we conceptualize sound design defines its meaning, or its function or role, in our experience of it and our conceptualization is, therefore, particularly prescient in any practice-based artistic research. This can be construed, in part, as a metabolizing or rendering of our personal experience into scientific research. Belton elaborates on Merleau-Ponty[71] by stating that phenomenology is an attempt to "put essences back into existence" characterized by "an approach to experience that seeks to describe rather than to explain or analyze" (Belton, 2008, p. 23). This may seem contrary to scientific method, however as we will find in our discussion of results, there are certainly findings that are *described* and, in an attempt to analyze, lead us to raise questions rather than draw conclusions. Thus, from Smith, we have inserted into our microscope a phenomenological lens that recognizes our own connection to the object of study, accepts description in the study of meaning above statistical analysis, and tries to touch upon invariable essences by moving beyond symbolic representation of a phenomenon. There is a certain parallelism here in that we have previously discussed inference as a fundamental aspect of both pragmatics and rhetoric as we looked towards linguistic properties of sound design, as

[71] Here Belton frames a discussion of Robert Bresson's *A Man Escaped* in phenomenology, and cites both Phenomenology of Perception (Merleau-Ponty, 1962) and Film And The New Psychology (Merleau-Ponty et al., 1964).

well as the balance between Deleuze's "abstraction from self" and Merleau-Ponty's "interimplication" in audience experience. For our research, Brickman allows, in a way, for inference to be a tangible part of our analytical output in the qualitative discussion while Smith allows us to leverage our own interimplication between body and world through our analysis, rather than running from it.

Perhaps together, therefore, these ideas of hermeneutics, targeted meaning, memory, environmental scaffolding, and phenomenology frame our conservation. Rather than offering criticisms of our scientific methodology, they offer insights into the admitted subjectivity of qualitative analysis as well as reasons for which some decisive conclusions may be elusive with audience studies, highlighting that we seek only to elaborate upon phenomena through scientific method in a critical stage of our exploratory journey. Fundamentally, we revisit Brickman, cited earlier through Willig, who highlights that science, or social science, is a question of goals, not orthodox methodology, and is "an effort to make accurate observations and valid causal inferences" (Willig, 2017, p. 10). Thus, we endeavor to explore and enunciate causal inferences that find expression in both our artistic work and our data.

3.2 Experimental Design & Methods

Hypothesis

Here, as we are entering into a statistical analysis, we shall position the experimental design in more traditional scientific terms. To recap from Chapter II, section 2.3, we hypothesize that sound design provides creators and audiences (the social sphere) some or all of the following advantages:

1. Increased attention and focus by audiences;
2. Support for narrative understanding and/or thematic interpretation;
3. Physical and temporal continuity – hiding of the apparatus;
4. Thematic continuity;
5. Imagined sound – engagement of mechanisms of phantom sound;
6. Experiential engagement or impression;
7. Emotional engagement.

The null hypothesis is that none of these advantages are met.

Summary Process and Audience Experience

In summary and as discussed in Chapter I, using Godfrey Reggio's *Koyaanisqatsi* (Reggio, 1982), we took the last 12 minutes of the film and added sound design elements, sound effects specifically – hard effects, tones and ambiances – and utilized various mixing and sound manipulation techniques. The full description of the creation of this object was given in Chapter I, section 1.5. This resulted in two versions: version A, the original without sound effects; and version B, the artistic research sound design with sound effects.

Four groups were each shown one of the two versions: 2 groups with version A; 2 groups with version B.[72] Participants watched and listened to their assigned excerpt either on their own (accessed through an outsourcing work platform) or in a theater on the campus of Lusófona University. Immediately after, all participants completed an identical survey digitally (a few in the theater without internet access did so by hand), which was comprised of some scaled numerical questions and some open-ended questions where participants entered text. Statistical analysis was conducted on quantitative and qualitative data.[73] Ultimately we use

[72] Full demographics and a discussion of the samples is provided below.
[73] Surveys are attached in Appendix B

independent samples statistical analysis for quantitative work and textual analysis for qualitative. Statistical processes were done in consultation with researchers at CICANT[74] for whom statistical analysis is a regular part of their work. Detail of methods follows.

CONTROL

The control in this experiment is the 15-minute excerpt of the original film, with the Philip Glass musical score. We must recognize that this choice of control affects outcomes. A different film, or a different score, or a different section of this film might produce different results and would certainly inspire a difference sound design. Additionally, it was important to keep the music in both versions because testing advantages between *no-sound-at-all* and *with-sound-effects* would also be testing a monosensorial experience on one hand and a multisensorial experience on the other which, in itself, would be a huge additional variable. Our study preserves multisensorial input in both the control and the independent variable. The control (Sample A Original Excerpt) can be found here: https://vimeo.com/308267921

INDEPENDENT VARIABLE

The independent variable in our independent samples study is simply sound effects, or to be precise the non-existence of SFX in Sample A vs. the existence of SFX in Sample B. Both samples retain the original music and SFX are added only to Sample B as the independent variable. It is further noteworthy to discuss certain specificities in the independent variable. For instance, at every moment in the creation of the independent variable, the author/sound designer has made a choice as to what SFX to use, as well as mixing levels, equalization and other audio signal processing. These can be thought of as both syntagmatic and paradigmatic decisions that have a direct impact on those outcomes being measured or discussed. Plus, as is always the case, impact and meaning are culturally determined at both the community and individual levels. Therefore it is possible, if not likely, that a different sound design would result in different outcomes, or the same study with a culturally different audience would result in different outcomes, important considerations. This speaks to the aforementioned discussion earlier in this chapter of the self-reflexive process of a phenomenological study and hermeneutics, as my own interpretive understanding of the *Koyaanisqatsi* piece, the utterances

[74] Center for Research in Applied Communications and New Technologies at Lusófona University, Lisbon, Portugal

of my sound design in the context of my embodied knowledge, and the ways in which I attempt to materialize abstract essences – or themes – into my design all play a role in the creation of this independent variable. This independent-samples study therefore has, as its independent variable, THIS SPECIFIC sound design. All subsequent references to the SFX used imply this design. The independent variable, Sample B - Excerpt with Novack Sound FX, can be found here: https://vimeo.com/308269202

DEPENDENT VARIABLES

The dependent variables of this study are those hypothesized advantages that were discussed in detail in Chapter II, section 2.3. Here, we will address each one in terms of the approach that the survey and analysis will take to interrogate them.

1. *Increased attention and focus by audiences*; the study asks a few quantitative number-scale questions that illustrate overall attention, attention in specific areas, as well as the consistency of attention over the sample. Quantitative analysis is applied.

2. *Support for narrative understanding and/or thematic interpretation*; The study seeks to see if, with SFX, more subjects take away story or meaning, and if there is a coalescing of interpretation. As earlier discussed, we place narrative and meaning, in other words story and interpretation, in a single category. For instance, a question like "what is the film about" can result in answers that relate narrative or answers that relate meaning. Either way, it is an effect of the conglomerate of audio and visual signifiers in their syntagmatic and paradigmatic contexts, and their interpretation. Written answers are coded for themes for qualitative analysis. Coding will be discussed further in methodologies.

3. *Physical and temporal continuity – hiding of the apparatus;* subjects are asked to rank, with a numerical scale, the extent to which they feel some series of shots happen at the same time and in the same physical location. This is designed to measure a sense of temporal/spatial proximity, or shared diegesis, between shots.

Often these visual shots have nothing to do with each other in terms of location or time. Quantitative analysis is applied.

4. *Thematic continuity*; as earlier defined, subjects are asked to rank, with a numerical scale, how connected certain series of shots are with one another in terms of *thematic meaning*, and to describe. Here, we ask if SFX increase the ability to ideate or infer meaning across disparate shots. As mentioned earlier, we have coined the phrase *thematic diegesis* to suggest that there is a universe of thematic ideas embedded in the work, similar to the diegetic universe of the story, but along thematic lines rather than physical ones. *Thematic continuity* interrogates if these themes are expressed across a sequences. Quantitative analysis is applied.

5. *Imagined sound – engagement of mechanisms of phantom sound*; As earlier discussed in Chapter II, section 2.3, imagined sound is a phenomenon in which we expect to hear something with a visual stimulus and, even though there is no sound, we believe that we have heard it. One question in particular addresses this phenomenon, with a numerical scale, and therefore quantitative analysis is applied.

6. *Experiential engagement or impression*; As earlier discussed in Chapter II, section 2.3, for us, memory of image and especially of thought process is an indication of the impression made on a listener-viewer or of phenomenological engagement. Several questions show listener-viewers an image and ask to what extent they remember that image, answered on a numerical scale. Quantitative analysis is applied, although qualitative results have a strong bearing on the discussion of results.

7. *Emotional engagement;* Emotion is measured via qualitative analysis to open-ended questions. Answer are flagged and coded for emotional words in the text. Details of the coding process will follow. This is therefore qualitatively measured via their responses.

POPULATION SAMPLING

While details of the sample populations are shown below in the Demographic Data section, it is relevant first to discuss our method of collecting these samples. The study had intended to pull in non-biased subjects that would represent the general population by offering rewards to randomly selected people in a mall for attending the screening and filling out the survey. However, just as this was being rolled out, COVID-19 hit and movie theaters were shuttered indefinitely. The decision was made to do two separate sample sets: one set for viewing versions A and B via Mechanical Turk, Amazon's digital work platform that has been used extensively for survey-based research of all kinds, called samples MTurkA and MTurkB; and a second set of A and B live in a theater at the University of Lusófona in Portugal where this researcher works and had access during COVID-19, bringing audiences local to that community, called samples PortA and PortB. These necessary decisions represented a switch from our targeted representational sample to more of a sample of convenience.

With respect to the MTurk online population, they were paid to conduct the study on a remote-work platform on which they are ranked, and this may have altered their approach to the task as a job with the possibility for bonus and additional work. Plus, the investigators cannot control for the screening/listening environment, and experiencing the sound mix on headphones or through laptop speakers would certainly be a different experience than in a theater with respect to immersion, noise, and low frequency content, all of which connect to our hypothesized desired outcomes. Some potential differences in results due to this are discussed. In addition, we eliminated eight MTurk respondents who did not provide demographic data. Having included these responses might have created odd biases if drilling down against gender or age. With respect to the Portuguese population sample, they were mostly students and they were in the screening room with their professors, which alters environmental scaffolding.

However, efforts were made to remove biases and approach more of a representational sample. For instance, MTurk populations were in many different English-speaking countries and were given an English survey. Meanwhile, Portuguese populations were given a Portuguese version of the survey. This created more diversity of language and culture overall. Portuguese students came from only two disciplines of study, law and communications, and therefore didn't represent a cross-section of the university. They are graduate and undergraduate students and therefore skew younger. However, all film/media arts students were excluded from the study.

By including graduate students we increased, to some extent, the age distribution and by excluding film/media arts students we removed some bias related to prior knowledge of the subject. To address remaining biases, the investigators have joined results so that combined independent samples are made up of: MTurkA+PortA and MTurkB+PortB. This configuration gives us the broadest samples of diversity possible under the circumstances and helps to correct for biases. Still we must recognize that the ideal scenario of more randomized solicitation in a theatrical space might have yielded different results. In addition, where separating the population-types yielded interesting results, we make note and discuss.

Demographic Data

Sample Size (n): Our samples sizes are well balanced in number. The two Portuguese samples have n=71 and n=70, while the two MTurk samples have n=94 and n=89. The combined A vs. B samples have n=165 and n=159. These are all balanced and high enough that statistical analysis should represent the population based on sample size.

Age: MTurk favors the midrange of age 36-50, but shows reasonable, even spread to the categories just above and below. On the contrary, Portugal is skewed heavily toward the 18-25 group, expected at a university setting. Both locations show a balance of age between their respective A and B groups, which establishes evenly-aged demographics across the control/independent variable divide, whether locations are combined or kept separate. However, the combined data shows 41% in the 18-27 year range, 28% in 36-50, and smaller numbers for other ranges. This gives an age skew towards young subjects and a bump at 36-50 that are not representative of the general population. Still, between A and B, or no-SFX vs. with SFX, demographics are consistent.

Gender: MTurk is balanced on gender. Portugal, however, shows a favoring of women. Combined samples yield 52.5% female and 43.5% male with over 4% of respondents identify as non-conforming or prefer not to answer. This is fairly well-balanced in terms of gender and approximates the population at large, which is closer to 51/49% female/male.

Survey for Data Acquisition

All data are acquired through a survey which can be found in Appendix B. All surveys are identical. The surveys were done digitally in Google Forms, however there were a handful

of subjects who did not have digital access post-screening and therefore completed the surveys by hand. Hand responses were entered by research assistants into Google Forms. We refer to the surveys and populations as SurveyA or SampleA (no SFX) and SurveyB or Sample B (with SFX). As previously explained, there is an English version (Mturk) and a Portuguese version (Port).

3.3 Quantitative Analysis – Data & Discussion

Data Processing

All quantitative analysis was conducted in SPSS (IBM Corp., 2020 version 27.0.1.0) using standard statistical measures. Since the study design compares two different sample populations in their responses to an independent variable, with or without sound design, T-Tests comparing the means is the best way to look for significance in an independent samples test.

In SPSS, single variable T-Tests were run to compare the means between samples A, no SFX, and B, with SFX, with a 95% confidence interval. In addition, certain types of questions that address categories of dependent variables were aggregated into "overall" data as shown in Table 3 below. All tests were run for the combined A vs. B groups as well as individually looking at the MturkA vs. B, PortA vs. B, and breaking out for gender in order to highlight any interesting results that may have surfaced, particularly given the differences in screening conditions and paid status. Full data of the combined group can be found in Appendix C. Below in Table 3 is a summary table of highlights where statistical significance was shown to exist. As is standard, a p-value of less than .05 is considered significant and the lower the value below that threshold, the stronger the significance.

Significant Data

Below is a summary table for all significant data in the combined samples, MTurkA+PortA vs. MTurkB + PortB, plus a demographic breakdown where appropriate to our discussion. Full data is supplied in Appendix C.

Question	Significance (p)	Direction change w/SFX
General Level of Attention	.610	increase
MTurk (online)	.287	decrease
Portugal (theater)	.050	increase
Memory of black man in close-up	.004	decrease
Memory of old man shaving	.021	decrease

Imagined Sound	.010	increase
Physical Diegesis Emergency	.024	increase
MTurk (online)	.019	increase
Men	.002	increase
Women	.703	increase
Portugal (theater)	.414	increase
Thematic Diegesis People	.006	increase
Overall Physical Diegesis	.024	increase
Overall Thematic Diegesis	.013	increase

Table 3: Significant Quantitative Results, Combined Samples, some breakdown

When combining our groups into our most diverse pool of demographics, therefore better approximating the general population, we see significant difference in a few questions plus in some aggregates, namely overall physical and thematic diegesis. We also see the cusp of significance for the Portuguese sample group on attention. In this analysis, we will interrogate each item with significance and elaborate on the technique used in sound design that may have contributed to the outcome.

Discussion of Results (Quantitative)

Desired Outcome: Attention and Focus

There was no significant different in measurements of attention and focus in the combined groups with the first question, "rate your general level of attention", giving a p=.610. This is one of the variables for which we had the greatest concern over bias. As previously mentioned, the MTurk population was paid for this work and therefore had motivations to pay close attention to the video regardless of which version. In fact, their self-reported attention ratings were quite high: on a scale of 5 the mean is 4.69 without SFX and 4.57 with SFX, with a standard error .061 and .094 respectively.

Meanwhile, the Portuguese audiences in the theater were in their school environment, pulled out of class, with professors present during the screening, establishing an environmental scaffolding where there are expectations put upon them. Interestingly their starting point for attention is 19% lower than for the MTurk population at 3.80 without SFX, standard error .095, and 10% lower with SFX, at 4.09, standard error .107. With a p=0.05 for an increase in attention for this group, the increase is on the cusp of significance. This increase in the mean on the cusp of statistical significance correlates with environmental scaffolding that includes the subwoofer and greater immersion in the theater environment. These mixed results speak to the need for an improved study design with a randomly selected audience in a theater as originally planned

before COVID.

In addition, and as previously discussed, the Philip Glass music is already a strong component in the audiovisual experience, a "potent example of a score dominating a film"(Ross, 2005). This can, as noted, play a strong role in the high attention numbers for both versions, for as elaborated upon in Chapter II, section 2.3, sound of any kind has had measurable results in altering or shifting visual attention.

Quantitative Analysis: Physical and Temporal Continuity (Emergency Sequence)

As specified earlier, in this sequence of people and actions, some of which showed emergency scenes, SFX were added with the intention of tying shots together in order to create the illusion that they are all happening in the same time and space.

Koyaanisqatsi Image 16: Emergency Sequence
Timecode: 7:06 – 7:45

There was an increase in connecting these shots in physical space and time with the addition of the SFX, showing p=.024. Interestingly, as shown in Table 3, this effect was only significant for the Mturk groups and showed very high significance among men in the MTurk sample, with a p=.002, and no significance among women. The hypothesized result that sound design conveys physical and temporal diegesis/continuity is clearly met by these results, but the differences found in different demographics is rather curious.

Quantitative Analysis: Thematic Continuity (People Sequence)

Subjects were asked how strongly they feel there is a connection in meaning between various images of people that are strung together in the film.

Koyaanisqatsi Image 17: Images From People Sequence
Timecode: 3:54 – 5:24

With a p=.006, there is a very strong increase in the connectedness of the people, which speaks to an expanded formation of narrative, or meaning, or some form of thematic connectivity, the materialization of the thematic diegesis, brought by SFX.

Quantitative Analysis: Overall Physical and Thematic Continuity

As there were many questions related to physical and thematic continuity, including some that did not, on their own, show significant change, consolidating those into aggregates serves useful to see if our significant ones are perhaps outliers. In fact, an increase in overall physical continuity with p=.024 and overall thematic continuity with p=.013 are both statistically significant. Physical diegesis or continuity has been brought about by the addition of background sound effects across shots, which unifies time and space because of the sonic continuity provided. Thematic diegesis, about the interconnectivity of shots in terms of narrative or meaning, is supported by a number of sound design and mixing devices used in different areas as discussed in Chapter I, section 1.5, artistic research. For instance, the aforementioned metaphor of a telephone SFX over shots of people, or elsewhere with cheering crowds from a baseball stadium that carry through the subsequent shot of nighttime buildings, are both sound design techniques that can draw meaning or interpretation across images.

Desired Outcome: Imagined Sound

Imagined sound shows a marked increase in activity with p=.010.

Koyaanisqatsi Image 18: Lighter, Imagined Sound
Timecode: 5:25 – 5:53

The specific question posed here is whether or not the listener-viewer heard the sound of a lighter being lit when, in fact, there is no sound. With the existence of SFX in the design, more subjects believed more strongly that they heard a sound. Put another way, there were significantly more phantom hearers of imagined sound with SFX in the design than without. By providing sound design in general, we have excited cognitive listening mechanisms even where sound is absent. This is entirely predictable from the previously-mentioned studies on imagined sound and activation of the secondary auditory cortex, yet it was quite pronounced to see the phenomenon appear as a statistically significant result. We would interpret this as a strong advantage, encouraging cognitive aural processing in tandem with visual stimuli, which is a form of interactive engagement.

Desired Outcome: Experiential Engagement or Impression

We remind the reader that we use memory as a measure for engagement or impression. There are two questions that show a significant *decrease* in the subjects' ability to remember a specific image when the SFX are present, and interestingly they are in the same sequence, and next to each other, contiguous.

Koyaanisqatsi Image 19: Black Man Close-up and Old Man Shaving
Timecode: 4:12 – 4:39

One is a close-up of a black man with p=.004 and the other is a close-up of an old man shaving with p=.021. This decrease was not an expected result, and that it happened for more than one image is even more surprising, plus these images are contiguous. The sequence of "people" in which these men are included has been altered with the addition of the sound of a phone ringing endlessly, not being picked up, going to a dial tone, and then starting again. Such a sound effect is meant to bring metaphorical meaning, that the people are disconnected in the world, or from one-another, or are hard to reach – many interpretations are possible, of course. The fact that the images are less remembered with the addition of the sonic metaphor supports the idea that the multisensorial experience of hearing the sonic metaphor alters cognitive processes, either bringing more attention to sound or more attention to meaning in favor of image, or having some other impact on memory we can only surmise. As will be discused in the *thematic diegesis* analysis below, there is a considerable uptake in meaning-making in this sequence with the SFX population.

In addition, for the Portuguese theatrical audience, there is a significant increase in memory of what we call Factory Man with p=.017, an image of a man, seated, smoking, looking depressed, at some technological workplace that feels like a factory.

Koyaanisqatsi Image 20: Factory Man
Timecode: 2:40 – 2:54

That this should be significant only for this viewing type is interesting, since the primary SFX of a low frequency rumble requires a sub-woofer for its full effect, which this audience had but for which we could not control in the MTurk population. Was this screening difference a source of divergent results for this particular moment? Unlike with the people images just discussed, this sound and that of a distant alarm were not metaphor, but rather meant to give a sonic value to the space and provide a contrast between sounds of anxiety and danger and the image of a rather still, calm, person ignoring their surroundings. The goal here is to generate some dissonance between sound and image that recalls Michel Chion's "contrast," one of his "five rhetorical figures of the said/shown relationship," were we to redefine "said/shown" to *heard/seen*. As earlier discussed, memory is tied to phenomenology and so we postulate that with such contrast provided on a long and steady shot, the audiovisual impression may be rendered stronger and more long-lasting to enable a positive shift in memory. This will be further discussed in qualitative results.

Quantitative Analysis: Additional Reflections

Here we must share that there are quite a few questions that did not show statistically significant change, more detail in Appendix C. Frankly some of them were surprising. Besides the results related to attention previously discussed, in terms of other non-significant changes, the overwhelming majority of examined responses did show some change, and in the expected direction. This, too, begs further discussion and experimentation around the demographics of the study and the screening environments, or environmental scaffolding. A more diverse sample population would be beneficial to tease out any contained bias. More importantly, a large study, conducted only in a commercial theater as intended before COVID-19 brought about a change

in execution, might have yielded different results. For many of these self-evaluative responses of an experience, the specificity of the environment comes into play as a critical part of niche theory. As mentioned, especially when we query the interaction between mediated cognition and emotion, the "environmental scaffolding" (Hven, 2019, p. 7) is as much an element of the audience's phenomenological apparatus as are the three structural tentpoles of film phenomenology; film, spectator and other spectators. Indeed, this environmental scaffolding not only includes a physical space. One may respond phenomenologically differently when surrounded by peers, or in the presence of a professor, or alone at a desk in a living room. Hence, niche theory applies equally to our theatrical university-setting population and our at-home MTurk setting. Furthermore, this shows that such a study might yield very different results if we are querying listener-viewers using mobile phones, a vastly different scaffolding, and warrants a separate study to that effect.

Quantitative Analysis: Summary of Results

Thus we have established correlation between sound design and desired outcomes or hypothesized advantages, notably in hypotheses numbers three, four and five, physical and temporal continuity, thematic continuity, and imagined sound, respectively. Hypothesis number one, attention, did not show correlation across the combined populations, although in the environment of the cinema theater, it was at the threshold of significance. Experiential engagement or impression, hypothesis number six, did not show correlation in an aggregated measurement of all possible sequence, but it did show correlation in two specific shots, in the opposite direction than expected. Correlation was also shown on a third shot, in the expected direction, for the theatrical population. These open an interesting dialogue as to possible reasons, a dialogue that we will continue in the qualitative results.[75]

3.4 Qualitative Analysis – Data & Discussion

Before meaningful data can be generated from textual surveys, the data must be coded. We used NVivo (QSR International Pty Ltd., 2021 version 1.4.1.4361) for qualitative textual

[75] We remind the reader that the other two hypotheses, numbers two and seven, are not examined through quantitative means.

analysis, imported all survey results, and began the coding process. Therefore, we begin by presenting our coding protocols in depth.

DATA PROCESSING - CODING PROTOCOLS

Coding is the heart of qualitative content analysis and it is riddled with minefields. To step around the mines as best as possible, we established coding protocols in conjunction with our research assistant. All responses of all audience members were coded into most categories. There were some codes created only for specific questions. A full list of codes and subcodes can be found in Appendix D.

To begin with, the research assistant and I separately reviewed a number of answers in order to establish the list of codes, or themes. Initially this list was longer, with many more sub-themes. We then separately coded a number of responses for these established themes and compared codes. Where we differed, we each defended our reasoning and came to a logical agreed-upon conclusion which further refined our coding protocol. We also ran a test analysis. After analysis we chose to simplify some of the themes by condensing their sub-themes when the aggregate of sub-themes became more interesting than dozens of sub-themes. Simply put, we had invested too much detail in nuances in thematic expression and chose to pull back from that level of nuance. With finalized codes and protocols, the research assistant then completed the coding of all responses. I separately coded some of the questions as a check and found no further discrepancies. What follows are specifics that illustrate the process with respect to question number six, "Please describe your overall understanding, if any (what the film is saying)". Below, words in quotes represent examples of words found in the textual data, i.e. the responses. Words in italics represent codes created in NVivo, also called nodes.

- *Narrative Clarity – Meaning:* The goal with question #6 is twofold: to capture overall expression of one hypothesized advantage of sound design, narrative or meaning, and to evaluate the strength of overall narrative clarity, something that will be further discussed in the analysis. Only those who clearly express no understanding, most commonly a blank answer or literally "no idea" are coded into *zero understanding*, meaning no narrative clarity. Anything that expresses a strong sense of meaning and/or narrative gets *high clarity*. Everything else gets *some clarity*. For *high* level of narrative clarity and/or meaning, we are looking for strong connectivity of ideas between

sequences, thus tying together an overall narrative or thematic understanding, and for every instance both evaluators must agree that the answer exhibits a high level of clarity.

- *Narrative Sentiment*: This category is part of capturing emotional response but we chose to attempt to separate the ideas of a positive/negative story from subjective emotion expressed. The responses are also coded into *negative, neutral* or *positive* if there is an understanding at all. Occasionally double coding is permitted here if there is a clear separation of positive and negative, as for "life can really be beautiful, but environmental devastation will destroy us". If the positive and negative words are tied into a singular idea, they are coded as *neutral*, as for "the rise and fall of humanity".
- *Emotion*: Emotion is a separate coding node that only gets a code if a sense of feeling is relayed. For instance, "the people in the image look very sad" does not indicate that the listener-viewer felt sad and does not get an emotional code, although it would be a negative narrative sentiment. "It made me sad" does get coded into emotion. These are either coded as a *negative* emotional response or a *positive* emotional response.
- *Themes:* This is a high-level code which contains the various themes that we saw evoked beta-testing, and it is here where we begin to examine narrative and meaning-making. They are:

Destruction	Environment	Money
Technology	Time	Humanity

Table 4: Themes for Coding

In addition, the theme *humanity* is further divided into subthemes:

General Humanity	Connectivity	Diversity
Life Struggles	Overpopulation	Routine
Urbanism	Work	

Table 5: Humanity Sub-themes

Figure 8: Word Cluster of Humanity Subthemes with relative weight

Some sample rules applied in thematic coding are as follows:

- Mentions of a "rocket" or "space" go into *technology*.
- References to "poverty" or "wealth" or "inequality" go into *money*.
- *Environment* is reserved for "nature" references, "planet," etc.
- *Time* gets a direct reference to "time" and/or "history", "progress", "evolution"
- References to "9/11" or "Challenger Explosion" go into *tragedy* which is a subcode of *destruction*.
- *Connectivity* gets all ideas of "selfishness" and "loneliness" (which also get a *negative sentiment),* and it also gets ideas of "connectedness" of people.
- *Humanity* deals with all answers about people or society. Society as a whole (unspecified) goes into *general humanity*, as do all answers of people or society that do not fit into the other *humanity* subcodes.

Moving beyond Question #6, the other qualitative questions ask for narrative or meaning-making around specific shots or sequences. These are only coded into *themes* and, in addition, every answer gets *zero understanding* if so demonstrated and a *positive, negative,* or *neutral sentiment*.

We also eliminated one qualitative question for all respondents, even though it showed interesting results. Upon coding the answers, it became clear that for Portuguese audiences, many believed the question was referring to a different sequence than it was. To avoid any mis-representation of data, the question was eliminated from the analysis.

With respect to coding for emotion and narrative sentiment, as noted earlier, we fully recognize that the state of mind of the individual during the screening and while completing the

survey is a key variable out of our control, as it is part of the environmental scaffolding (Hven, 2019, p. 7). Yet, we feel the population may be large enough to correct for that. We try this through two mechanisms. One is to code answers in *narrative sentiment*; is this a *positive*, *negative*, or *neutral* response? The other is to isolate the rare instances where a respondent specifically expresses their emotional state.

DATA PROCESSING – WEIGHTING AND SMALL NUMBERS

Because of the difference in the sample sizes (n) of each sample population, albeit small, we have weighted the frequencies of codes accordingly. All charts and percentages in the analysis refer to weighted numbers. Some of the frequencies of theme coding are large and others rather small. The goal of the filmmaker and sound designer is typically to use the tools available to engage a listener-viewer in story and experience, or as we are testing, "reducing the uncertainty" (Rogers, 1983) that desired outcomes will be met. Any measured increase in response towards narrative understanding means that even a few more audience members are coalescing on a particular theme, reducing uncertainty. Therefore, we highlight the large percent changes on thematic analysis regardless of the raw numbers, but always with transparency when numbers are small.

COMBINED DATA AND DISCUSSION OF RESULTS – QUALITATIVE

The detailed qualitative data results in chart form can be found in Appendix E. These charts contain the frequencies of different codes against specific questions. Below we offer individual graphs of notable data along with relevant explanations and a discussion of the results. In some cases the reader will find some cross-pollination with quantitative results already discuss earlier in this chapter. We therefore feel it clearer to combine data and their discussion into this hybrid format. After individually-coded results are discussed, we iteratively offer a re-synthesis of the results in section 3.5, organized around the hypothesized advantages that we interrogate.

Of the seven hypotheses originally stated, two are subject to qualitative analysis: (#2) support for narrative understanding and/or thematic interpretation; and (#6) emotional engagement. We will begin by looking at narrative sentiment and expressed emotion to capture audiences' emotional response. We will then move on to thematic analysis, one question at a time, interrogating meaning-making by the audience. Each item is coupled with a frequency

table to illustrate audience response and trends. Blue bars represent responses to the A version (no SFX), while orange bars represent responses to the B version (with SFX).

Narrative Sentiment

All answers to all questions were coded in one of the three categories, *negative, neutral*, or *positive*.

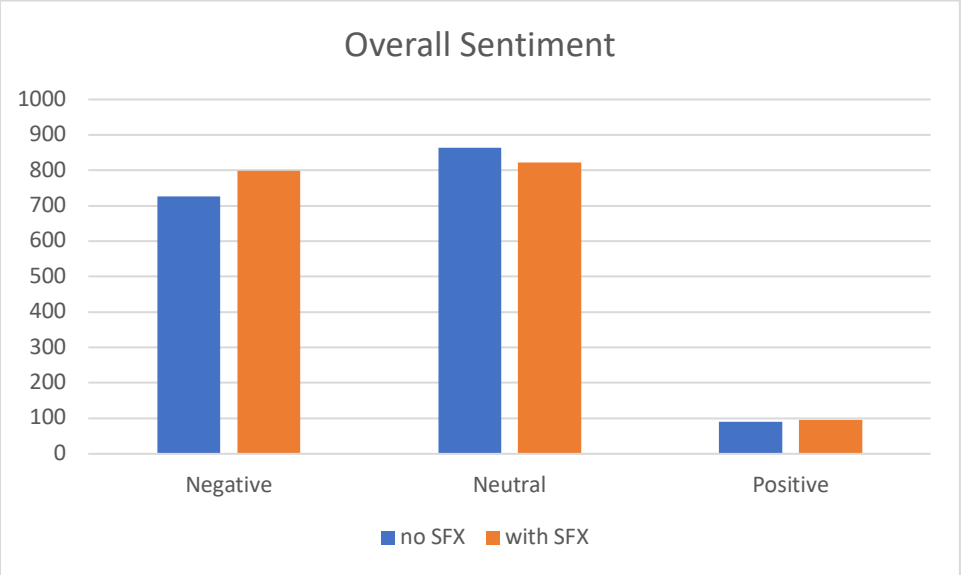


Figure 9: Narrative Sentiment Overall

From *no SFX* to *with SFX*, we see a 10% increase in negative codes and a 7% increase in positive codes, with a dropping of neutral codes by -5% percent. More people are expressing ideas with less neutrality and in more positive or negative tones. While perhaps not a dramatic shift, it is noteworthy.

Expressed Emotion

Expressed emotion was coded when a subject specifically expressed *their* emotions.

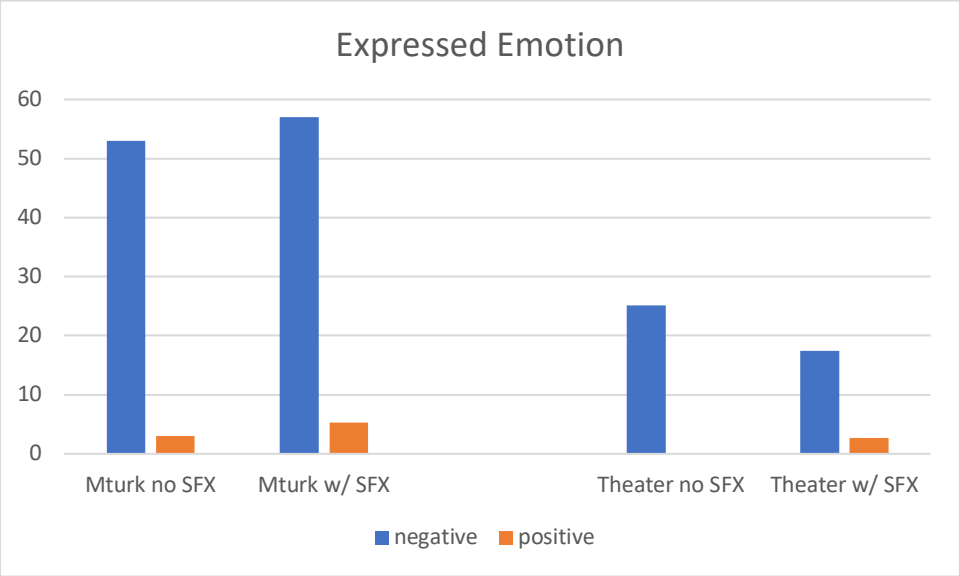


Figure 10: Expressed Emotion

There was a marked difference between the two viewing populations but the common element is an increase in positive emotion expressed. As earlier noted, this is one category that is very small in numbers so significance is hard to claim. However the addition of a few SFX specifically designed for that purpose, like children laughing, and magical chimes, might have had a small impact.

Combining narrative sentiment and expressed emotion, we do have a shift towards more emotion with sound design. It is not dramatic, but textual analysis of audience response does show it to be there.

Qualitative Analysis: Thematic Interpretation

The following analyses address specific questions of thematic interpretation one at a time. Note that for code charts, the code *humanity* is a bucket for sub-codes dealing with humanity, including *general humanity*, and therefore is an aggregate value. On the charts, the humanity sub-codes, *Ghosts* through *Connectivity*, are above *humanity,* the codes below *humanity* are other master categories. *General Humanity*, a subcode, is coded for responses that deal with people, society, or the world in more general, universal or abstract ways that cannot fit into the other more specific subcodes, thus every humanity response has a subcode. Where relevant, results from quantitative analysis are brought into the discussion.

Question 6: Please describe your overall understanding, if any (what the film is saying).

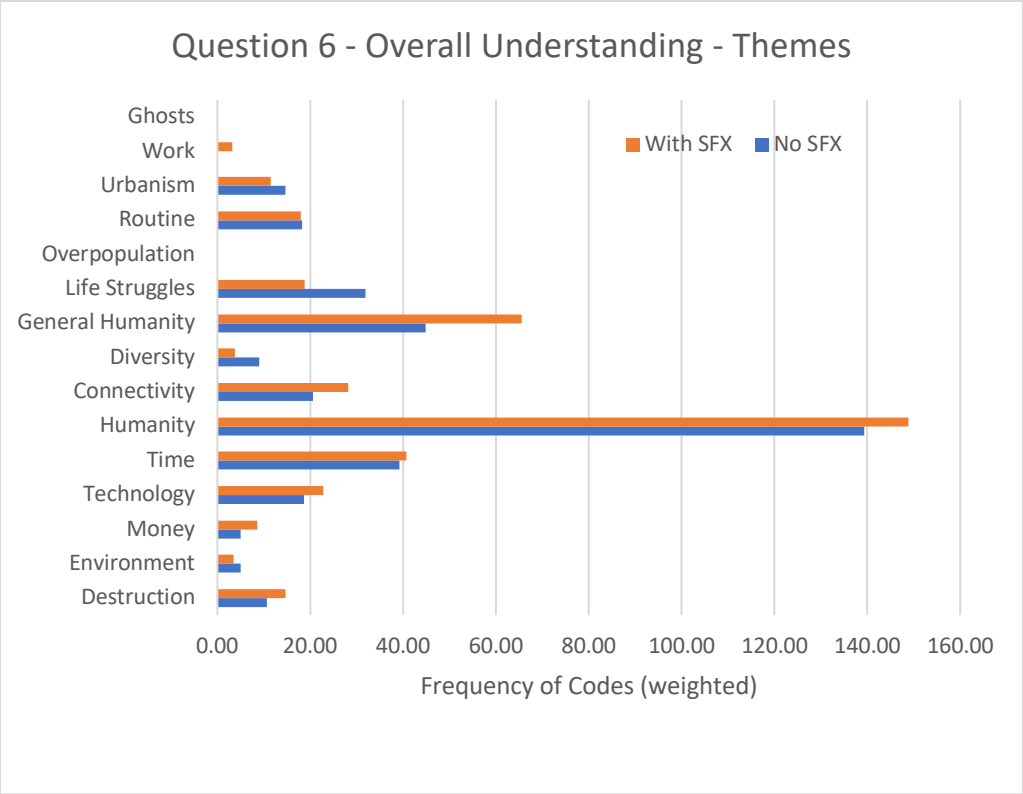
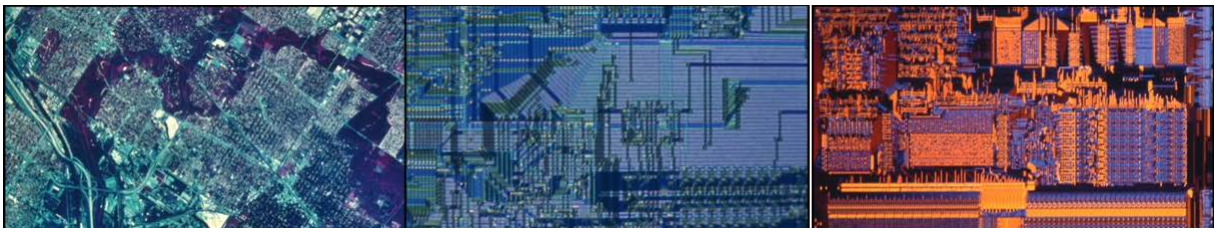


Figure 11: Question 6 - Overall Understanding [themes]

There is a 10% increase in codable themes. This points to a higher degree of interpretive work on the part of the respondents with SFX. While quantitative analysis (rate your level of understanding) did not show a significant difference, qualitative results illustrate richer interpretive responses. Respondents also show notable increases in the following themes: *destruction* (38%), *money* (72%), *technology* (23%), *connectivity* (37%), *general humanity* (46%). Respondents show notable decreases in *diversity* (-58%) and *life struggles* (-41%). *Humanity* as an aggregate increases as well (7%). Therefore overall, within humanity sub-categories, we see a consolidation of ideas towards general humanity and connectivity of people which can be expressed as how we are or are not connected. Considering the increased sense of *overall diegesis/continuity* with the sound design, both thematic and physical as registered significantly in quantitative data, this sense of connection of ideas and coalescing on certain human themes is a strong and expected result.

Question 9: Please describe your understanding of what this sequence might be trying to communicate.



Koyaanisqatsi Image 21: Aerial City, Chips, Tapestry Sequence
Timecode: 0:30 – 2:00

Question 9 refers to the long sequence of aerial shots of cities that are intercut with images of computer chips and native American tapestries. The added SFX of birds, winds, and a cyclical eagle screech whose cycle gradually gets longer are meant to encourage connectivity between these images, helping the listener-viewer to draw meaning from that connectivity.

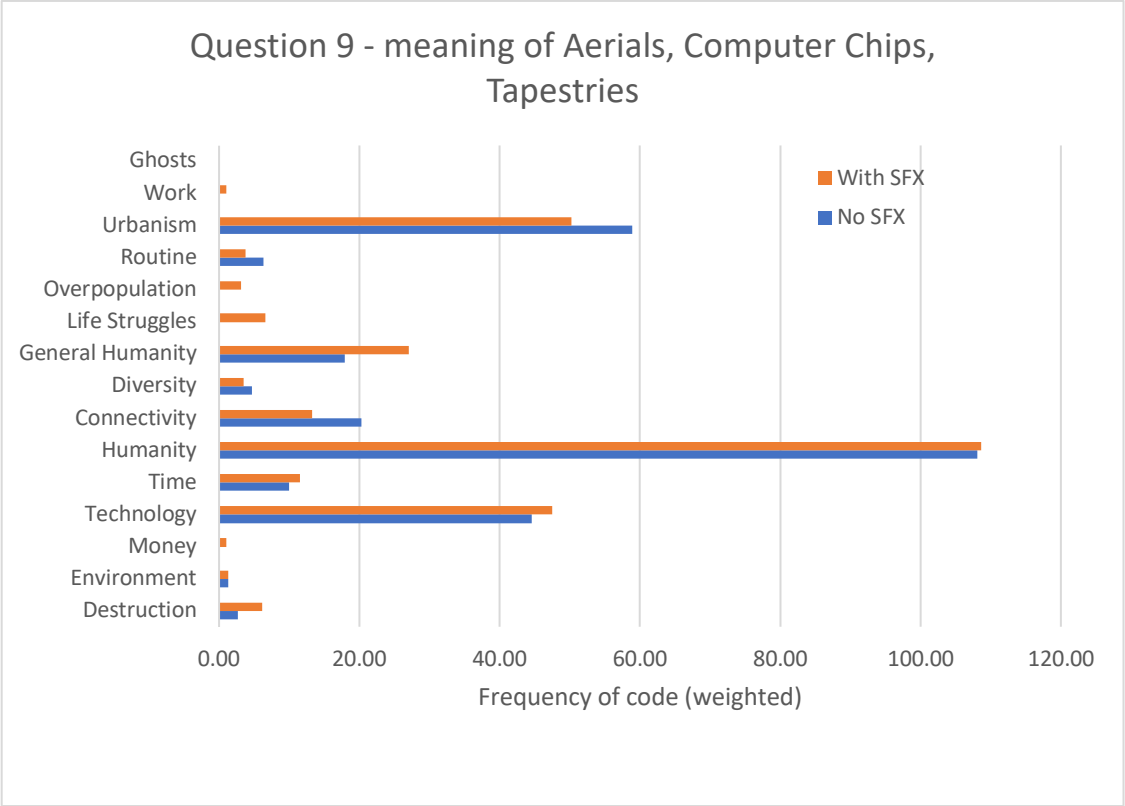


Figure 12: Question 9 - Aerials, Computer Chips, Tapestries

There is a small increase (6%) in codable themes. But what seems most striking here is the appearance of many more themes with SFX, four additional themes, which seem to come mostly from urbanism and connectivity, plus a few smaller categories. From a sound designer's

perspective, the intended impact of SFX here is to draw some parallel between the images through what we have previously termed *thematic continuity*, which has been shown in the quantitative analysis to have increased with SFX. Thus similarly to Question 6, we find that there is greater thematic takeaway for the respondents. The specific types of images – city, computer and tapestry – might offer clues to the analysis. Looking at *urbanism*, *technology*, and *general humanity* as relating to those images, we see a shift in the balance away from urbanism, noting that the city shots are most obviously cities, concrete images, and towards humanity and technology which are more abstract images and ideas. The addition of SFX therefore seems to help the audience extract connected meaning from the more abstract images by providing a continuity of sound throughout, specific nature sounds, and a rhythmic repeating sound, the eagle screech. In summary, more people felt a bit more strongly that images were thematically connected, while those who believed there to be a connection found richer discursive elements.

Question 12: Please describe your understanding, if any, of these shots.

Koyaanisqatsi Image 22: Stadium to Office Tower Shots
Timecode: 2:18 – 2:40

Question 12 deals with thematic continuity in the sequence of shots that are brought together by SFX. Images are of a long passing nighttime aerial shot over a filled baseball stadium that then cuts to static high-rise office towers at night with many lights on. Added SFX are a variety of cheering sounds which are attached to the stadium through synchresis, depicting the fans and then carrying over with other celebratory sounds to the steady image of the office towers.

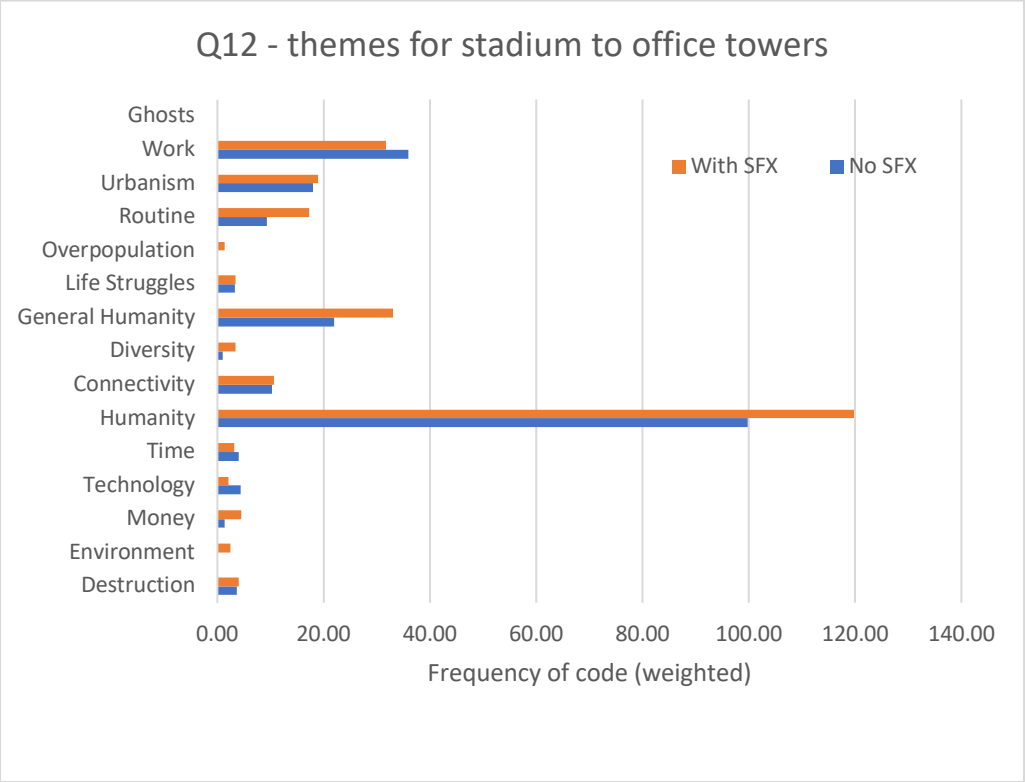


Figure 13: Question 12 - Stadium to Office Towers

There is a 20% increase in codable themes, more than for any other purely thematic question. This points to a higher degree of interpretive work on the part of the respondents with SFX. We are reminded that through quantitative analysis, listener-viewers showed a slight increase in connection that was not significant. Putting our quantitative and qualitative results together, we demonstrate that with SFX the answers illustrate richer interpretive responses, even if written by a similar number of respondents.

We additionally find a strong increase in coded values for *humanity* (20%) as an overall bucket, and specifically for *general humanity* (50%), and *routine* (86%). Together, these illustrate a coalescing of ideas towards part of the intended interpretation, that both images are about people and their lives, in a city, with everyday activities. Diversity and money both increased a great deal on a percentage basis, although they are low raw numbers.

While it is clear that people exist in the stadium image as colored dots seen from an aerial tracking camera, no visible people exist in the static office tower images. The addition of the crowd sound should draw attention to the crowd, and carrying that sound over into the office shot might establish the idea of people in those buildings. In addition, the choice of celebratory sounds was meant to instill the idea that people celebrate play/sports and possibly celebrate

work just as much, to a fault. We separately coded this question for *people* and *work-play* to explore if this connection emerges from the data. Audiences with SFX show a 94% increase in references to *people*. This demonstrates both that adding a sound to a shot through synchresis, like crowd cheers to a stadium, evokes a signified of that sound in the image, and that the *thematic diegesis or continuity* across these images works, drawing them together. As for the *work-play* code, there was little change, in fact a lowering by -4%. This implies that, at least for this example, the more direct sonic sign, 'sound of crowd cheering' signifying 'people', has more predictable results than a desired abstract idea or conclusion like 'sound of crowed cheering" signifying 'celebrating work to a fault'. This brings us back to Deleuze's targeted real that, along with hermeneutics, pragmatics and environmental scaffolding, frames our discussion, and to which we shall return in our final content analysis discussion in section 3.5.

Question 14: Please describe your thoughts, if any, during this shot?

Koyaanisqatsi Image 23: Factory Man
Timecode: 2:41 – 2:55

As mentioned earlier, for this shot of a man sitting in profile, smoking, looking somewhat depressed, with muted depressed colors, in a room that seems to be a technological control room for a factory or some other location, we added low frequency rumble and a distant alarm. With the addition of the SFX, we were interested to see what sense the audience gets of his location, so we coded separately for this. There is a large decrease in mentions of the location (-43%) in the SFX version. And yet, quantitative analysis showed that there was a significantly higher degree of memory of this shot, at least for the theatrical audience (p=.017), who also showed a decrease (-49%) in mentions of location. We consider this to be a very interesting result; stronger memory, less attention to location. Let's dig deeper into themes expressed by audiences for this shot and see what our results may articulate.

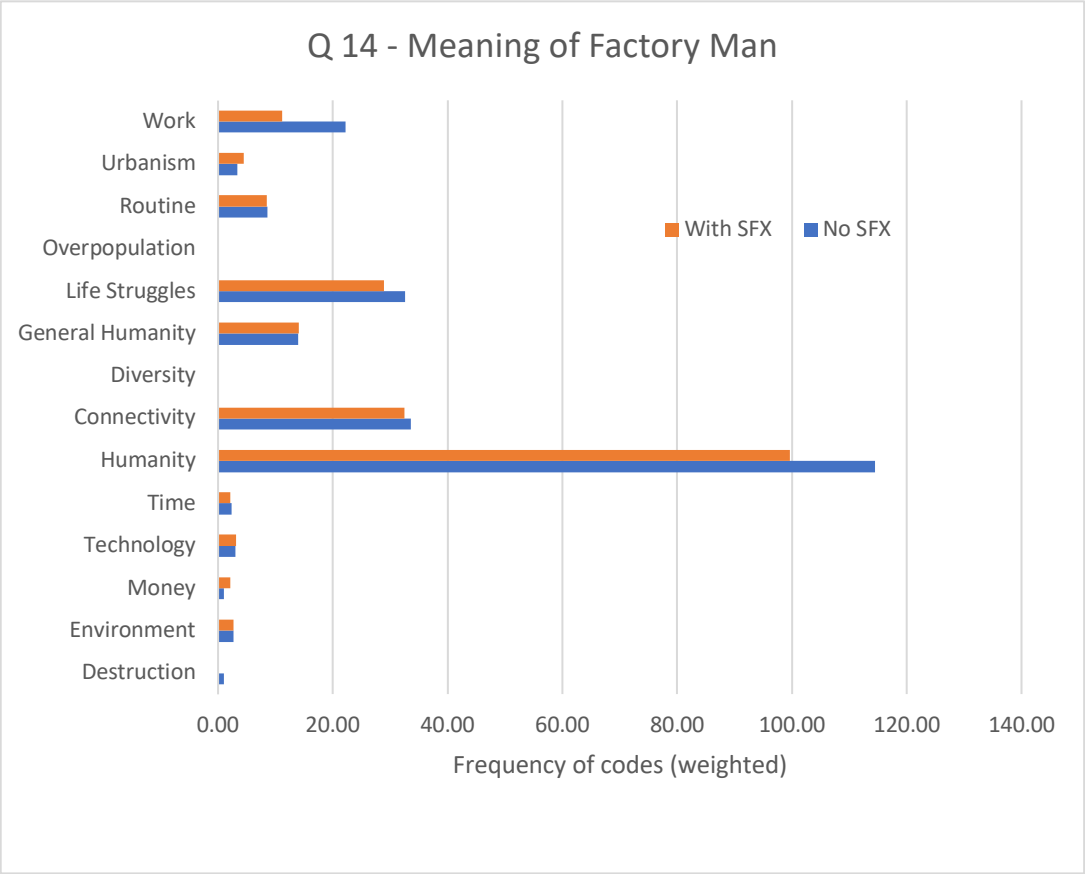


Figure 14:Question 14 - Factory Man

Major categories were dominated by *humanity* coding in both versions, although it decreased (-13%) with SFX. We can see that these decreases can be found in *life struggles*, *connectivity*, and *work*. This is quite curious, for here we present a shot of a man at work, sitting and smoking a cigarette, in a setting that seems to be a factory, with a very glum, stressed look. And yet, the sense of his location (down -43%), the idea of work, the notion of life struggles and references to humanity, or society or people in general are all decreased with sound design. What's more, codable themes overall are down (-12%).

Examining the sound design, the added low frequency rumble can be known to create anxiety as mentioned earlier. The addition of an alarm or hospital beeping sound is meant to create a sense of alarm, stress, or warning through phantom imaging offered by sound. Against the calm image, as noted earlier, this forges an uncomfortable "contrast" or "counterpoint" between sound and image, in the *heard/seen*: the soundscape contains information that is discordant with or counterpoint to the image.

Is it possible that this contrast or counterpoint makes the audiovisual moment more

difficult to process coherently, providing more descriptive or figurative elements that pull the audience away from simply thinking about the man and his location (information provided by the image), and either towards a non-interpretive stance, or towards other ideas that were either not memorable or couldn't be expressed?[76] This will be further discussed in section 3.5. Digging further, is this perhaps more exaggerated in the theatrical group who could hear both the subwoofer and the subtlety of the alarm better, since in the quantitative analysis it is they who remembered the shot more?

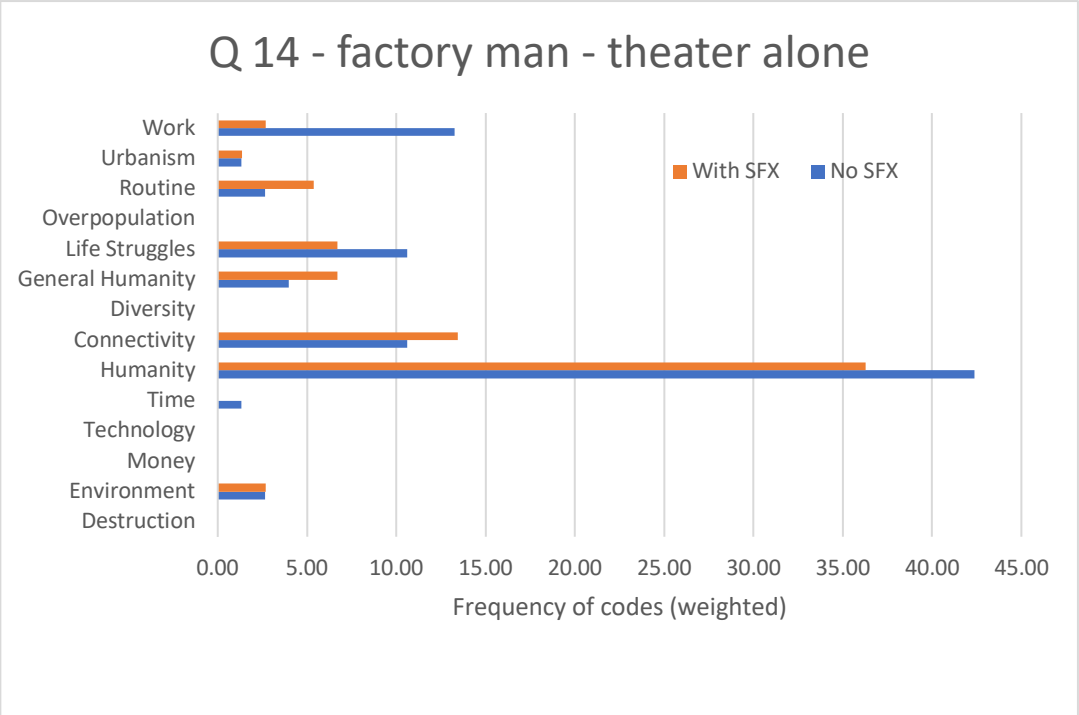


Figure 15: Question 14 - Factory Man [theater listener-viewers alone]

As the above chart shows, indeed, the effect is more exaggerated for the theatrical group and some changes are reversed for this group relative to the Mturk group. Within the *humanity* subthemes, there is consolidation heavily away from *work* and *life struggles* and towards *general humanity, connectivity* and *routine*. The overall picture shows that what was noted for the combined populations is stronger in this group, except that those themes that *are* expressed move away from the image of the man in his workplace, *work* and *life struggles*, plus the decrease in *location*, towards more universal ideas, with one possible exception: routine.

We cannot determine the extent to which the younger population of this group is a

[76] In our discussion of Genette and Proust that follows the empirical study, we will share assertions that the overwhelming descriptive nature of Proust brings a kind of phenomenological surrender, and then later discuss this aspect with respect to Velvet Goldmine.

determining factor beyond the screening environment. However, since both populations experienced a similar interpretive trend and that trend was exaggerated in the group with the environmental scaffolding that greater permits the reception of this particular sound design element, and coupled with the quantitative data that show our theatrical SFX population to have a significantly higher degree of memory of this moment, our data point in support of the screening environment with subwoofer as having a greater impact on this population. In turn, this speaks to the impact of technologies like immersion, subwoofers and noise reduction in supporting targeted outcomes. Additionally, with the higher memory for this group, coupled by a decrease in the expression of ideas and trends towards more universal ones, the phenomenological engagement from sound design, suppressing the analytical apparatus in favor of experience, seems more probable in the theatrical environment, with SFX. This phenomenological viewpoint will be further explored in the next section of this chapter, 3.5.

Question 16: Please describe your thoughts during the elevator shot

Koyaanisqatsi Image 24: Elevator Shot
Timecode 2:56 – 3:36

During this long shot of people crowding into a huge elevator, the beginning of the shot carries the low frequency rumble plus the alarm sound from the previous shot, the man in factory, which slowly fades out via filmic discourse, connecting the images sonically, and lastly leaving the rumble to carry through this long shot. As mentioned, when sustained this is a sound typically used to increase anxiety. Here, we were interested to see if this results in a more negative interpretation of the image. In fact, there was a +23% increase in negativity expressed, and fairly even across the two populations. Earlier in this thesis, we spoke of the silent, non-speaking nature of the people in the elevator. This, too, may play a role in negative emotion when we hear acousmatic SFX, the rumble, but do not hear what we see. As previously noted,

Chion might call this a c/omission, an instance in which what is said via dialog is either omitted or obliquely ignored (Chion, 2009). It can also work as figurative counterpoint through metaphor as the daily commuters are silenced as individuals. Perhaps then, the increase in negative responses reflects that the silence of people may not be as strongly felt in a viewing with neither dialogue nor SFX throughout, since nothing on-screen is heard, but *with* SFX, the absence of dialogue represents a formal change. The increase in negative responses was quite similar between the populations, +20% theater, +26% MTurk. However, for the theater group there was hardly any change for neutral or positive narrative sentiment while for the MTurk listener-viewers neutral responses were down (-57%) and positive responses were down (-47%).[77] Therefore, MTurk listener-viewers carried the overwhelming bulk of changes in narrative sentiment experienced by the combined group. This is counter to what would be expected, since there is less control of the production of low frequencies for this population, among whom we may assume very few are using a subwoofer.[78] Does this point to the connection of sound back to the preceding shot of the factory man shot and the cognitive dissonance produced there, or the c/omission effect between sound and image? Perhaps, but we cannot know. What we do know is there is a measurable change.

On themes, we see very little change with the additional of SFX. Therefore, the most powerful effect of the addition of this low-frequency sound effect is the expression of negativity and reduction of positivity and neutrality.

Question 32: Please describe your understanding of story or message, if any, from this sequence of people shots.

[77] Note that this is an instance in which raw numbers are small, and therefore small changes represent a large percentage change.

[78] While good headphones do produce these low frequencies, they are focused to the ear canal, which yields a very diminished impact. Much of the anxiety created by sustained low frequencies is thought to occur due to vibrational sensations in the lower torso, enabled by sound's ability to travel through us. Via headphones, this would not happen.

Koyaanisqatsi Image 25: Shots from People Sequence
Timecode: 3:54 – 5:24

This question, as mentioned earlier, pertains to a long series of shots of people, mostly medium shots, some close-ups, in the street. SFX added were meant to give continuity of space which we evaluated quantitatively, and thematic continuity simply by tying the shots together with a low frequency tone and the addition of a sonic metaphor.

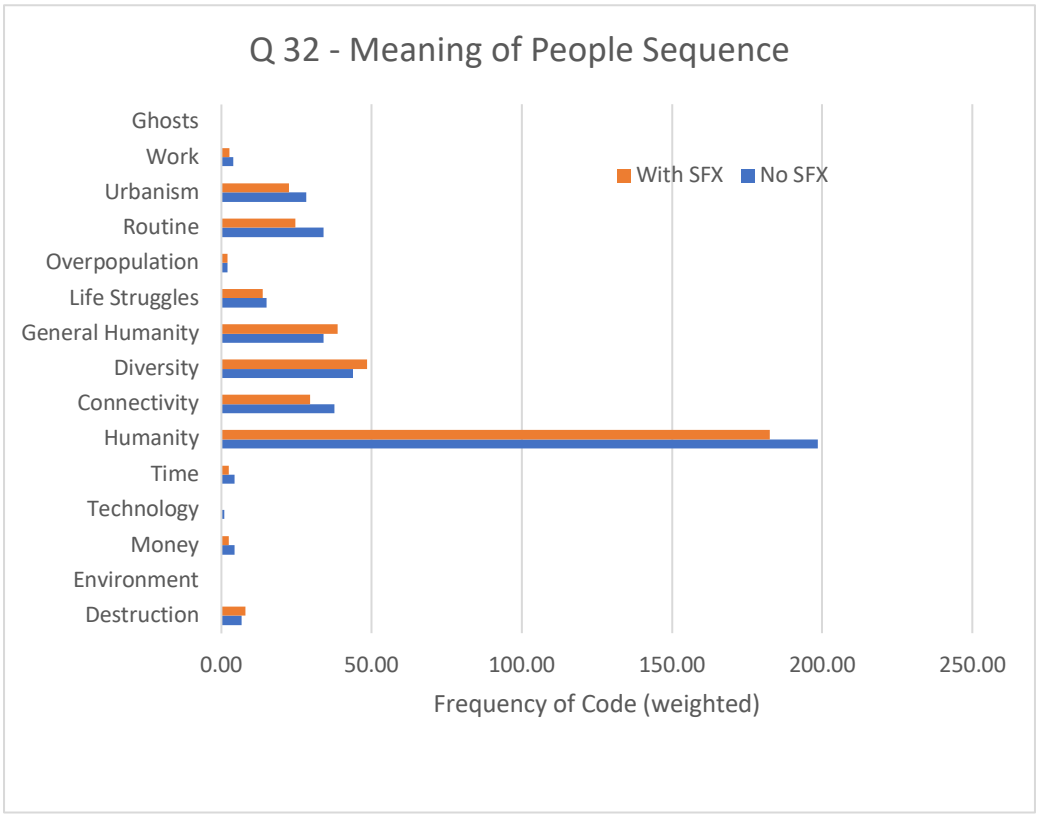


Figure 16: Question 32 - sequence of people

Once again, there was a decrease in the number of codes, but this time by only -9%,

with *humanity* as a bucket down -8%. With overall fewer ideas expressed, there is some consolidation towards diversity and general societal themes, albeit small.

However, the most interesting result here is, in fact, a surprise. The idea of the disconnection of the people, supported by the metaphor of a telephone SFX ringing and not being picked up, evidently did not work. Mentions of interconnectedness or disconnectedness of people were down (-22%). Did such an overt metaphor distract? Did it actually cause the reduced interpretive powers? Or perhaps it lead listener-viewers into a more phenomenological experience, disarming the analytical apparatus? Let us recall that quantitative analysis showed that the increase in a sense of connection between the shots was highly statistically significant: with the phone SFX listener-viewers feel the people shots are more connected, more of a uniform sequence, and yet the connected idea is not what we sought – loneliness, disconnectedness – or at least listener-viewers are not expressing that connection. This suggests, again, that the change brought about by the SFX here and as seen in quantitative data is more phenomenological or experiential than analytical on the part of the respondents, and highlights our inquiry with respect to hermeneutics and Walsh's dance of the "immersive spectator", as will be again addressed in section 3.5.

Question 36: Please describe your thoughts, if any, during the emergency sequence.

Koyaanisqatsi Image 26: Shots from Emergency Sequence
Timecode 5:43 – 7:03

In this sequence, a number of shots of unrelated people and places have been shown in the quantitative answers to be put together in physical space by the addition of SFX,

establishing physical diegesis in the minds of listener-viewers. Added SFX are various emergency-worker voices, an urban background, police/fire shortwave radio speech, and a siren. Here in qualitative analysis, we want to look at meaning.

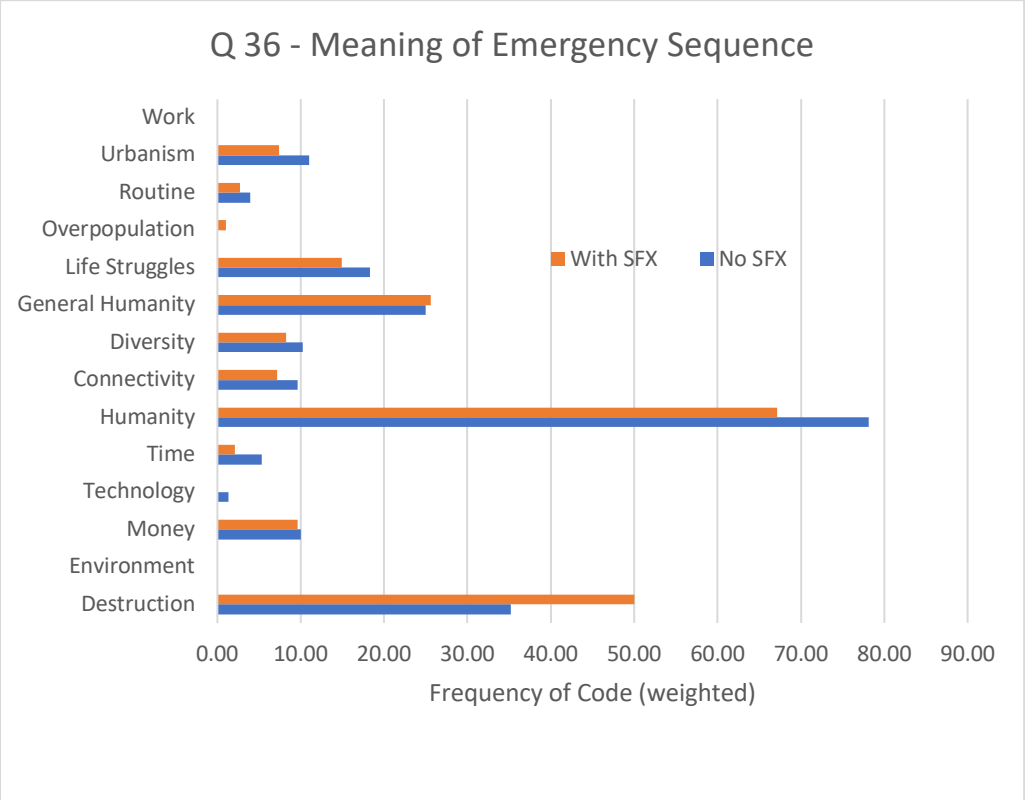


Figure 17: Question 36 - emergency sequence

Most striking is a +42% increase in responses coded into the major category *destruction* and out of the major category of *humanity* (-14%). Most of the images in this sequence show destruction and some kind of emergency, but some are also of a man sitting half-naked in a window, a blonde woman in a fancy car closing the window to which we added the SFX of the window closing, and other images that, when not tied to the emergency images sonically, can be perceived as discontinuous. With our sound design elements added, the overall messaging of destruction is clearly better registered by listener-viewers, while interpretations that connect to people and society are reduced. In other words, just music and image brought about more abstract thinking, while the literal nature of the added SFX brought audiences into more concrete grounding.

With statistical significance in the quantitative analysis showing physical diegetic continuity for this series of images, coupled with the thematic shift, we can conclude that the

connection of space puts all shots, in a sense, into the destructive emergency setting, which dominates the thematic response of the audience above the individual shots of people. We also consider that the pre-lapping audio in the sound design that was intended to create the impression that the woman with the cigarette was looking at the man being placed into a stretcher, has put her in the location and established her as our proxy witness. Whether that had a direct impact is difficult to measure without testing two versions with SFX, omitting that shot. There are many such desired effects that remain unmeasurable in this context due to the discursive nature of the overall piece.[79]

Question 41: Please describe your thoughts, if any, during the stock market shot.

Koyaanisqatsi Image 27: Stock Market Shot
Timecode: 9:17 – 9:46

The image of the stock market is a steady shot from above, with people moving around, and with layers of the shot shifted in time and superimposed. The sound design added is the sound of a busy Las Vegas casino, with the beeps and bells of slot machines.

We note a decrease (-12%) in codable responses, so we see a small diminishing in interpretive impression. However, there are some notable changes to coded themes.

[79] Much more on discourse will be discussed in the following chapters.

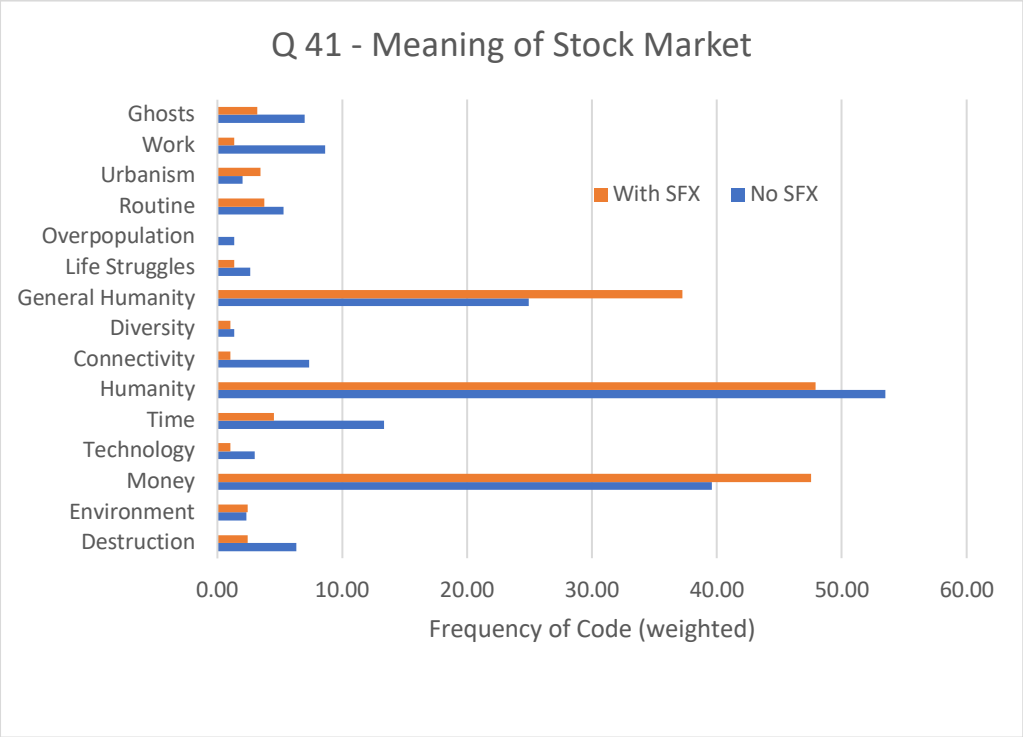


Figure 18: Question 41 - stock market

The most important increase with sound design is the coded theme for *money* (+20%), which is certainly a desired or targeted result of the SFX of a casino, reinforcing ideas such as gambling, handling money, wealth, financial risk, etc. Ideas that relate to general humanity have a larger increase (+49%) even while its bucket code of *humanity* shows a decrease (-10%). Where do these increases come from? We see decreases in *destruction, technology, time, connectivity, diversity, life struggles, overpopulation, routine, work and ghosts*. Thus we are consolidating strongly on two ideas: money and general humanity. Recall that *general humanity* refers to things like people, life, civilization, society without specifics that would put it into a different *humanity* subcode, in other words, more universal ideas. Therefore, the addition of SFX is both pushing the idea of money and drawing attention to the overarching idea of its relationship to society. The *ghosts* code was added specifically for this question. Without SFX, a small number of responses (7) contain references to ghosts, likely coming from the ghosting-effect in the superimposition within the image. With SFX, the number is reduced from 7 codes to 3 codes. It's a large percentage decrease (-55%) and suggests that the image is more powerful in making an impression of ghosting without SFX, albeit with small numbers. This concurs with general theories of multisensorial perception in film (Antunes, 2014). The SFX adds counterpoint that may, for some, draw away from the concrete, formal aspects of a

superimposed image or at least provide different thought patterns that are more memorable.

Questions 43, 45, 47, 49: Please describe your thoughts, if any [during four sections of the rocket sequence, asked individually].

Koyaanisqatsi Image 28: Selected Shots from Rocket Sequence
Timecode: 9:47 – 10:43

There were no interesting data points that emerged specifically from qualitative questions during the rocket sequence. Let us examine, in the interest of gaining knowledge and with full recognition that the researchers, as for any qualitative study, are implicated via the hermeneutics of analysis, some possible reasons that might shed light on this unexpected outcome, albeit anecdotally. We have already noted that we did not see any significant attention increase in the quantitative portion of the study, which may or may not be due in part to the paid nature of the screening for some of our audience, and the teacher-chaperoned nature for the other audience, in other words, the environmental scaffolding. In addition, these questions were at the very end of a rather involved survey, and survey fatigue may have been setting in for some subjects who provided answers in less detail, which gives us fewer coding nodes with which to work. Both of these may account for the lack of difference between populations. In addition, as discussed earlier around the phenomenology of episodic remembering, with these questions in particular coming at the end of the survey, the longest time gap between initial experience and recollection, and with possible survey fatigue, our subjects may be either less

capable or inclined to engage in the "feeling-based metacognition" necessary to generate a meaningful response to the questions. We will return to this in the next section (3.5) when we address memory, impression and its relationship to phenomenology.

3.5 Discussion of Results and Related Conclusions

Having examined and discussed the results of individual qualitative questions in the survey and significant data from quantitative results, we will now bring our discussion back to the hypothesized advantages, or desired outcomes, addressing each one relative to content-analysis results and underlying theory.

(1) Increased focus and attention in audiences: Quantitative analysis was applied here and no statistically significant difference was found in the data. As earlier noted, the presence of any sound at all, in this case the Philip Glass music, may already play a strong role in drawing attention to the screen just as Amy Herzog was drawn to an unfolding narrative in a live theater game, *Sleep No More* (Punchdrunk, 2011). We are reminded that Horstmann posits that the embodied experience of film spectating is defined by multi-channel reception of representation, image and sound but with sound separated into onscreen and acousmatic sound (Horstmann, 2018, p. s15). There is very little visualized sound added in our sound design and both versions therefore are dominated by acousmatic sound, from music and the sound design material. Embodiment, to Horstmann, is well satisfied by the original version and we propose that the embodied experience is not fundamentally different between the two versions, and that attention is greatly derived from this embodiment. This is supported by the fact that raw data on overall attention in the version with only music shows a mean of 4.31 out of 5 (standard error .064), quite a high level of attention to begin with, and an increase to 4.36 (standard error .073) with SFX. This supports the idea that attention is strong to begin with whether this is a result of music already existing in the original piece, or the environmental scaffolding as mentioned that includes being paid for the "work" in the case of half our population or seated in front of professors in the other, or some combination of the two. Moreover, from a phenomenological hermeneutical perspective, the listener-viewer may, according to Baracco, self-mediate perception by magnifying or diminishing intentionality of perception (Baracco, 2017, p. 48), so interpretation and attention are both tied to intentionality. We are also reminded that for the

theatrical audience, increased attention was on the cusp of significance (p=.050), where immersion and the subwoofer were controlled.

In terms of "Materializing Sound Indices (MSI)", that "render the real" (Chion, 2009, p. 480) by drawing attention to physical properties of things on screen, we did not add substantial MSI's except for very few instances in areas that we did not query, like the subway train on tracks, and floor fans. Therefore we would not expect to see increased attention from MSI's in our sample. We do use a "c/omission" (Chion)[80] between sound and image, in drawing the audience's attention to a character's psychological condition by lowering background sounds: with the fireman walking in the emergency setting to a moment of contemplation. In fact we registered a statistically significant increase of physical and temporal diegesis/continuity in this sequence with the added sound effects and a consolidation of themes towards destruction. Attention and focus may have played a role in this, but we cannot know. We don't have sufficient data to demonstrate a correlation between increased attention/focus and sound design with our study, and we believe the bias in the environmental scaffolding of the sample populations plays the greatest role.

(2) Support for narrative understanding and/or thematic interpretation: Again we put forward the idea of *curating the flow of emotion and cognitive perception* in a controlled manner, in this case not about emotion, but cognition. It is in this desired outcome where, more than anywhere else, we are interrogating rhetorical stance and argumentation as discussed in chapters I and II.

As earlier shown, the increase in thematic interpretation overall is a desired outcome, which is coupled with a statistically significant increase in thematic continuity. Therefore we have demonstrated correlation between sound design and our hybrid concept of narrative understanding and thematic interpretation. However, the notion of *rhetorical stance* as discussed in Chapters I and II leads us to further evaluate if the intended themes through authorship are met, the curating of interpretation. Table 6 shows specific scenes or shots, the intended thematic responses, and whether or not sound design reduced the uncertainty that these ideas would be met.

[80] Chion, as cited and discussed earlier, but our revised definition of the "said/shown" being the *heard/seen*.

Sequence or Shot	Intended Theme/Meaning	Was intention met via sound?	
		Yes	No
Aerial Sequence	city, humanity, technology	x	
Stadium/office	humanity, routine, people	x	
Office alone	celebration of work/work-play balance		x
Factory man	work, life-struggles, disconnectedness		x
People sequence	disconnectedness		x
Emergency sequence	destruction	x	
Stock market	money, society	x	
Rocket sequence	humanity, destruction, future		x

Table 6: Rhetorical Stance - desired outcomes

We can see from Table 6 that, beyond the overall consolidation and increase of interpretation already noted, some specifically-intended themes were met while others were not. This existence of even some movement in the intended direction supports our hypothesis that uncertainty has been reduced in meeting desired outcomes and, moreover, it supports the rhetorical stance of the artist being met, at least to some degree. However, and quite interestingly, the fact that themes at times consolidated on different meaning-making or didn't change at all also means that the interpretive responses of audiences cannot be well predicted through sound design, they can only be targeted.

For the aerial sequence, stadium/office sequence, emergency sequence and stock market shot, the frequencies of codes that correspond to intended meaning were increased with sound design. Let us see what is common to these four. All of these have a background sound that is consistent, possibly diegetic in traditional terms and, if the sequence is more than one shot, background sound establishes a type of continuity to connect the shots, in some cases physical and in some cases more thematic. Increased sound continuity seems to correlate with increased targeted meaning-making. Additionally, the sound design for the aerial sequence and the stadium-to-office sequence both have codable themes increased; more evident meaning-making. Plus, for the aerial sequence which contains abstract images in montage with concrete ones, and the abstract superimposition of the stock market shot, the implied content of those

abstract images is better reflected in the responses with sound design. This leads one to wonder if continuity with concrete sound over abstract ideas helps the targeted ideas of those abstractions emerge, for it seems so via our data. From a semiotics perspective, we are reminded from Chapter I that, as Francesco Bellucci shares in his discussion of "Peirce on Symbols", an index "has denotation but lacks an adequate connotation" (2021 np). Given this premise, we can postulate that with regard to meaning-making, these sounds are not functioning indexically. Rather, they function as recognizable sounds originating in the story-world, and there is a clearer "represented instance" (Metz) that may steer meaning-making in these four cases.

For those shots or sequences where targeted themes were not met, the office shot alone, the factory man shot, the people sequence, and the long rocket shot, we find a few things in common as well, the strongest of which is the existence of counterpoint or contrast between sound and image, as discussed in the artistic research case study in Chapter I, via a sound that is not possibly of the story-world by traditional definitions. For the office shot, no cheering crowd sound is actually possible at that location, and for factory man, the dangerous alarm sound is incongruous with the stillness of the man. For the people, it is the phone ringing that is clearly not of the literal story-world, nor are the vocal screeches and magic chimes that we hear over the rocket exploding and falling. In addition, the factory man and people both have low frequency rumble while the factory man and rocket are long shots, the rocket being interminable. Given this information, we posit that using *nondiegetic* SFX, as traditionally defined, correlates with reduced *targeted* meaning embedded in the sounds and images, supported further by the anxiety-provoking low frequency rumble and a long shot length.[81]

To explore how this may be rooted in our theoretical underpinnings, we are reminded first of our discussion of semiotic challenges with sound and Valle's assertion that as a result of temporality, "presence is turned into presentation" (2015, p. 76). In our design, we have provided sound signifiers that are presented over time, and we suggest that they each act via three mechanisms at once; iconicity, symbolism, and index. For instance, the phone ringing sound "possesses the quality signified" (Peirce et al., 1998, p. 307), that quality being the sound of a phone itself, and is thus an icon according to Peirce. It is a symbol in that the repetition of the iconic sound in the absence of the phone being answered is a culturally accepted symbol for "no one is picking up the phone". The potential metaphors are many, and it is here where the

[81] The long shot of the rocket will be discussed later in the section.

index emerges from discourse between sound and image: these people are disconnected, no one is listening, they are alone, their minds are vacant, etc. Thus the acousmatic and nondiegetic sound functions on many planes simultaneously and the presentation over time (Valle) becomes an essential quality of the signifier (Peirce). We query now if in this moment, the triadic nature of a "represented instance" (Metz), put in discourse with both image and haptic response from low-frequencies, renders sound's strongest role to one of presentation over time, rather than denotation or connotation, leaning away from meaning-making and toward phenomena.

For both sets of sequences/shots, we have sounds that are not easily placed in Chion's tricircle and can be experienced as diegetic or nondiegetic at different moments, by different people. Even those tied to the screen, like backgrounds of an emergency space, are immersive in a theater or with great headphones, and therefore not so tied. For all of these, there is a clear, albeit unpredictable, shift in registering meaning when the soundscape contains information, regardless of traditional notions of diegesis, as shown in our data. We believe our data show that sounds leaning towards screen-based take the listener-viewer towards more targeted meaning-making, while more abstract sounds or sounds that provides contrast or counterpoint, regardless of so-called diegesis, make the audiovisual moment more difficult to process cognitively, providing more descriptive or figurative elements that ultimately pull the audience away from an interpretive or analytical stance to one that is more phenomenological.

(3) Physical and Temporal Continuity – hiding of the apparatus: We have posited that sound design offers an advantage by helping to establish, if not establishing outright in many instances, physical and temporal continuity. Our data have clearly supported this position with statistically significant measurements that consolidate all questions in this area and specifically in the case of the emergency sequence, although as mentioned this was most evident among men in the MTurk sample, something for which we cannot account. Desired outcomes of continuity are met by our sound design, with sounds like background traffic, temporal sounds like sirens, voices on emergency radios, etc. joining visual shots in time and space. This satisfies, even in the documentary form of *Koyaanisqatsi*, the Hollywood tradition of editing becoming less visible, as "the requirement of the fiction" (Williams, 1985, p. 337).

(4) Thematic Continuity: To reiterate, we coin this phrase as an attempt to capture the ways in which images can be brought into a thematic relationship by sound. We previously noted that the music of *Koyaanisqatsi* already does this, citing the music critic Alex Ross (Ross, 2005), but propose that even with the strong contribution of the music, sound effects will still

connect shots thematically. Our data supported this desired outcome strongly in the people sequence with the metaphor of the telephone ringing, even though the targeted meaning of *disconnectedness* was not met. This highlights that our sense of what constitutes *thematic continuity* is not the same as *targeted meaning* or *rhetorical stance*. Is it enough, therefore, to use sound design to get the audience to think, to do the interpretive work, without the prescribed meaning being met? For us, the answer is yes, as this is a sign of overall engagement. Our *thematic diegesis* as constructed or supported by sound design correlates with helping audiences to draw from the specific universe of thematic ideas embedded in the work, and materialize these themes, or at least reduce the uncertainly that they would be materialized.

(5) Imagined sound/engagement of mechanisms of phantom sound: We have successfully evoked Chion's "deaf cinema" where sounds exist but cannot be heard (Chion, 2009, p. 3). Our phantom or imagined sound clearly showed a statistically significant increase in being experienced with sound design, as hypothesized. Specifically, the absent sound of the cigarette lighter was imagined to be heard by more people with sound design present in the mix, therefore the omission of sound for a visual stimulus better triggers imagined sound.

(6) Experiential engagement or impression: We further hypothesized that sound design contributes to the experience of audiences in some ways that are a bit difficult to describe and we used the phrase *experiential engagement or impression* to encompass an idea that leans more towards phenomenology as a result of creating memorable experiences, or moments, the kinds of filmgoing experiences that are carried by audience members and, hopefully, shared with friends and families. We previously shared the notion that episodic recollection is a metacognitive feeling-based account of phenomenology (Perrin et al., 2020). Thus, we used questions of recollection, valued on a scale from 1-5, in order to assess this memory or engagement. To our surprise, there were two moments that showed a statistically significant *decrease* in memory, the close-up of a black man and another of an old man shaving. As discussed, these occurred in an area where there was a significant uptake in thematic continuity by the audience. Yet, for another image, the factory man, there was a statistically significant *increase* in memory with sound design for the Portuguese theater population sample, a population that had a more controlled screening environment with subwoofer and immersive audio, and again we are reminded of Horstmann's multi-channel embodiment supported by the subwoofer and the haptic sensation it provides. We conclude that using memory alone as a gauge does not allow us to show correlation between engagement or impression and sound

design. However, there are other items that do point towards increased impression, depending on how we define it.

Throughout this study, we have repeatedly leaned into, and highlighted in our theoretical background in Chapter I, phenomenology and the balance between Deleuze's "derangement of the senses" (Walsh, 2004, p. 175) and Merleau-Ponty's "interimplication" and "dynamic passivity" (p. 179), concluding for us a sense of *surrender* made possible by sound. In the people sequence with two images of reduced memory, but an uptake in thematic continuity via a metaphorical counterpoint offered by the sound of the telephone ringing, we wonder if the rhetorical positioning between sound and image shifts the experience towards one of either conscious interpretation, which would be less phenomenological and more hermeneutic as it engages the analytic apparatus, or simply a more phenomenological, less analytical engagement, yielding to our *surrender* which articulates with Walsh's "immersive spectator". This would be supported by the fact that the expected theme of disconnectedness from the phone metaphor *decreased* with sound design. Audiences "got" the message without the sonic metaphor, so were they taken elsewhere with the metaphor – not to other themes, but to a place where they recalled less image and expressed fewer themes? The memory shift is statistically significant so there *is* an impact or change brought about by the sonic metaphor, well established by our study, but exactly what that change is remains elusive. Perhaps the stark contrast provided by a sound that is in no way screen-based in the traditional sense may tilt the audience into more of a phenomenological reception with less of a conscious analytical engagement. We also mentioned a possible shift for audiences on the factory man shot, where, particularly for the theatrical audience with more immersive and low frequency sound, the shot was more easily recollected and all themes easily associated with the shot were *decreased* with sound design, plus overall themes were down, too. Here we find an example that points to the analytical apparatus being diminished by the sound design in favor of experience, or *surrender*. These results point to sound having a capacity to alter visual memory and memorability, concepts that we see as byproducts of phenomenological or experiential engagement, but these notions seem intertwined rather than having a directly predictable relationship.

Before leaving our discussion of phenomenology, we'd like to address the rocket sequence, which showed no change in themes, and nothing statistically significant in terms of memory nor any other data. The intent of the sound design here is to anthropomorphize the rocket with light wind soaring, an animal roar for the explosion, the magic chimes during the

fall, with flame flares of optimism, and that these were among the final images and sounds experienced by audiences before beginning their survey. Therefore, its greatest impact may have been a phenomenological one, whereby the specifics of the moments did not register changes in the survey questions that addressed them directly, but were rather registered in the questions at the beginning of the survey, answered immediately upon finishing the film. These first questions ask for overall meaning-making for the entire film excerpt and are the ones that more likely embody changes experienced in these last moments of viewing. Those changes that registered significantly via the beginning of the survey include a greater expression of universal abstract ideas relating to humanity and the connectedness of people. Perhaps the sound design elements for the rocket segment had their due impact in a way that we did not directly capture in the last questions, but did in the first. A change of study design could attempt to capture this more directly.

As an additional point of theoretical interest, we shall address the fact that this last shot is interminably long, clocking in at 3 minutes and 20 seconds, and thus we revisit our discussion of temporality and phenomenology. As discussed in our theoretical backdrop in Chapter I, in her book *Temporality and Film Analysis* Mroz discusses the unique experience for listener-viewers when engaged with long takes, particularly with respect to Mroz's own analysis of Tarkovky's film *Mirror* (Mroz, 2013, p. 93). She reflects on Jon Beasley-Murray's argument that for Bazin and Deleuze, "long-takes are especially associated with an 'inhabitation' of the real duration of time, in which the bodily sensation of temporality is prioritised over its disjunctive narrative coding" (Beasley-Murray, 1997, p. 49). Moreover, she quotes Synessios' (2001) assertion regarding the impact of the long take on cognition,

> …the mesmerising camera movements, together with the unusual events taking place within the frame, confound all attempts to interpret the image. The emphasis is on directly experiencing it and allowing it to affect a deeper layer of our consciousness (pp. 50-51).

Though applied to Tarkovsky, these notions ring true for the end of *Koyaanisqatsi* and may even be applied to earlier sections where we added contrasting sonic information to long steady shots. Extrapolating from Mroz, we argue that registering specific cognition during this mesmerizing rocket shot is handicapped by the image itself, and possibly compounded by the equally mesmerizing and dominating Philip Glass score and SFX. The SFX serve as the

"unusual events" taking place within the frame – synchresis of an unrecognizable animal cry of pain, cool winds, magical chimes – seducing the listener-viewer experience to one of *surrender*.

(7) Emotional engagement: We use two coded measures to discuss emotion: narrative sentiment expressed and personal emotional states expressed by the audience. With narrative sentiment, sound design correlates with fewer neutral responses, and an increase in negative or positive. For expressed personal emotional state, sound design correlates with an increase in positive responses. We earlier cited a 2020 study (Cuadrado et al., 2020) showing that emotional responses to sound correlate to frequency, volume and clarity and also to sound content: signified, interpretation, and meaning. Our results support this, as our audiences with SFX experienced a screening with more low frequencies, more volume due to added SFX, and sonic signifiers that convey content. For instance, the magical chime sound on the falling rocket and the sound of children in a playground convey magic, innocence, lightness, fun, more positive ideas that may result in the increased positive response for both narrative sentiment and personal emotion. Likewise, the increase in negative coding may come from the combination of more low, threatening frequencies in many parts of the excerpt and signifiers like the screaming dinosaur, or unanswered empty phone ring. In fact, in the silent and long elevator shot where we only added low-frequency rumble, we had +23% increase in negativity expressed. Indeed, we find correlation between sound design and emotional engagement that serves the social sphere, even if defining specific causes of this shift is rather complex.

SUMMARY:

Through this quantitative and qualitative content analysis of our research-based artistic work, it has been shown that sound effects, or sound design, correlate to desired outcomes being met, or more accurately, that the uncertainty of those outcomes is mitigated by sound design. With sound design, the themes express by listener-viewers about the overall meaning of the film excerpt consolidated towards *general humanity* themes and the connectedness or disconnectedness of people. *General humanity* answers speak of society, people and the world in more abstract and universal tones, suggesting that essences of human existence are more brought to mind – or heart – with sound design, over the more specific nature of other themes. In addition, codable themes were increased by 10% with sound design, suggesting an overall

greater power of interpretation. While thematic consolidation or coalescing was not experienced universally across our sample, a number of sequences showed such consolidation, suggesting that sound design has the ability to create, steer or direct interpretation of the audiovisual in a particular direction. Concurrently, sound design has the ability to alter the listener-viewer's position between cognitive function of a spectator and *surrender* into an experience. Also noteworthy, low frequencies had the effect of increasing the negativity in responses.

In conclusion, this experiment suggests a host of impacts of sound design, some expected and some unexpected. Most importantly to the research design, we can conclude that desired outcomes of filmmakers and sound designers can be supported by sound design. As earlier stated, this analysis truly only applies to this sound design and this piece, a poetic documentary film. These effects may be more subtle or more exaggerated with other types of films or sound design choices and, most likely, won't be uniform from sequence to sequence within any film. Perhaps most importantly, this research is impressed by results that suggest a tendency for sound effects to move an audience towards a phenomenological experience, accessing greater human themes and abstractions, and drawing connectivity between shots and sequences. Christian Metz (1974a), in attempting to reconcile semiotics and phenomenology writes, "the cinema is the phenomenological art par excellence, the signifier is coextensive with the whole of the significate, the spectacle its own signification, thus short-circuiting the sign itself…"(Carbone, 2014, pp. 224–225) (p. 43). Our data have pointed in this direction. While they clearly show correlation between most of the hypothesized outcomes, all but attention/focus, we are stymied to see that targeted meaning-making is only sometimes met, and always by only a share of the audience. Metz asserts that the spectacle, the full sum of all the parts, is the signification. For us the spectacle is comprised of the multisensorial, the sound-image-haptic, even the smell of popcorn, and within sound it includes the simultaneity of diegesis, the triadic signification of sound effects as icon, symbol, and index, the interaction between layers within sound, the sound-pressure levels generated, the environmental scaffolding of sound, and more – conditions that defy diegetic and rhetorical structures tied to the screen but rather create experience. The changes that correlate with our sound design highlight the discursive relationships between sound and image, and sound and sound, that are at the heart of the audiovisual cinematic experience.

A close reading of Winston along with our content analysis of sound design

demonstrate that both immersive technology and sound design fall within the natural evolution of the technology of film. Moreover, sound design reduces uncertainty in achieving a number of desired results from the filmgoing experience. It can be regarded as an innovative technology that follows Winston's path of technology evolution, with elements that are complemented by Rogers' model of diffusion of innovation. Uncertainties remain, as with any innovation, and they reside in what Rogers calls "expected consequence" (1983, p. 13). We, the sound designers, who use sound design as a technology of communication will always face, according to Rogers, uncertainty in expected consequences, and we have demonstrated this to be true with our intended meaning. And yet, Rogers also notes that the corresponding beneficial reduction in the uncertainty in meeting desired outcomes, which comes from the information base of the innovation, represents the working strength of the innovation to solve a problem (p. 13). Thus, sound designers innovate and the collective information base of sound designers themselves and the tools they implement become the very technology that allows for further innovation.

Sound design provides a relative advantage to both creators and audiences through the formation of physical and thematic diegetic positioning, activation of cognitive auditory centers by auditory signifiers, increased tendencies towards phenomenological experience as exemplified in both memory and thematic properties. Additionally, the sound design study demonstrates some potentially powerful aspects of film sound that are to be further explored, to better isolate and define the myriad advantages provided by the methods or rhetorical devices of sound editing and mixing – work for further research.

As earlier stated, the primary goal of the content analysis has been to address our second research inquiry, whether assumptions in sound film theory and as stated by theorists and practitioners are defendable, or rather, to examine if claims of the power of sound design, specifically the use of SFX creatively, can be demonstrated rather than assumed. Statistics via quantitative research has shown that some desired outcomes, in fact, can be demonstrated. But it is the qualitative study that has, in many ways, yielded the most interesting and perplexing results, supporting desired outcomes often, but also raising additional questions while shedding light on audience experience.

Throughout the narrative of this research, from Chion's foundational theories to the transformation of the *Koyaanisqatsi* object of study via artistic work, and then through content analysis, we seek to generalize via epistemics the ways in which sound functions in film and yet we are stymied by existing theories which remain based on image paradigms, particularly

diegesis and Chion's rhetoric. The particular instances of usage and audience reception don't fit nicely in generalized patterns identified by film sound theory literature. Ultimately, we seek to draw connections that address a fundamental hypothesis: that the strengths of sound for film go beyond typical notions of discursive style, emotional signaling and diegetic positioning that are common to film sound analysis in the contemporary canon, and therefore warrant an exploration in another form of discourse. We have demonstrated many of those strengths, but are still in search of a discourse.

To reconcile this, we now consider looking at film sound not only as a critical element that is traditionally thought to derive its raison d'être via image, but as a fundamental part of the very discourse of film, approaching Genette's take on Proust and the complex network of narrative structure.

CHAPTER IV. ADAPTING GENETTE'S NARRATIVE DISCOURSE TO FILM SOUND

> The general is at the heart of the particular, and therefore…the knowable is at the heart of the mysterious (Genette, 1983, p. 23).

We have, thus far, concluded that sound design offers a host of benefits to the social sphere of filmmakers and audience. Some of these are our tested hypotheses from content analysis, for which can now go beyond assumption and state them to be *effects*, *impacts*, or *functions* of sound in film. There are many other aspects of sound referred to in the literature as discussed in Chapter I where we looked at the inadequacies of existing theories of sound for film, particularly diegesis and rhetoric. We also brought many other concepts into the paradigm, like semiotics, pragmatics, phenomenology, even cognitive neuroscience, and more. What has emerged from this intertwining of ideas, schools of thought, and ways to analyze the audiovisual experience is a sort of Proustian discourse, with twists and turns, a varied and complex tapestry, and remembrances of things past that print through layers of analysis like a palimpsest of film and linguistic theory. The common thread that we have shown, however, is this: existing theories are most often based on sound in relation to image.[82] Our analyses, first of the case study of our own research-led artistic work, and then of the content analysis, particularly the qualitative thematically-based analysis, may be somewhat Proustian but they demonstrate that there are myriad ways to look at sound on its own terms, and they emerge out of an attempt to reconcile what might be unreconcilable, that sound for film functions on many levels outside of its relationship with image, and many more levels within that relationship. The relationship itself remains elusive, mysterious, but as Genette states, "the knowable is at the heart of the mysterious" and therefore our search continues.

[82] Either directly or by positioning it via dialogue, as with Chion's five rhetorical figures.

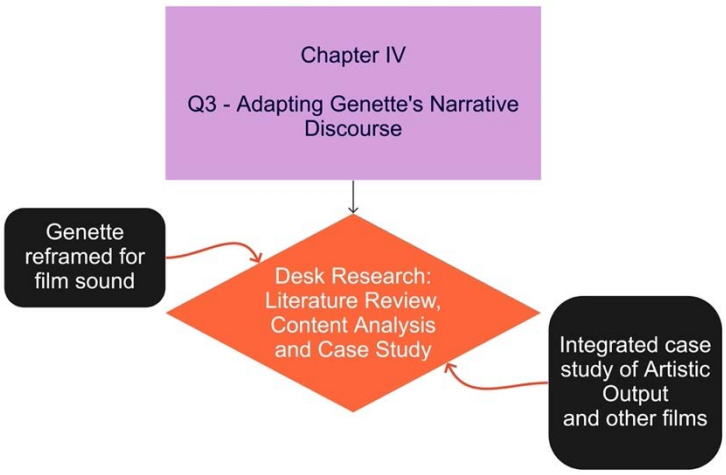


Figure 19: Chapter IV Roadmap

We remind the reader of our overall ambition with our research stated in our introduction: Can we theoretically liberate film sound – dialogue, sound effects, music and mixing – from strict image paradigms, and then bring it back into conversation with image on equal ground? This research continues by endeavoring to explore other theories outside of traditional film theory that might shed new light on how sound works. We are drawn in our research to look towards literary theory, particularly that of Genette, for two primary reasons: first, we see in the discursive role of sound with image, and sound with sound, a rhetorical stance that more mirrors the rhetorical figures of literary theory; and second, Genette roots his theory of narrative in a study of Proust, whose adjectival form, Proustian, captures, to us, the very essence of the complex tapestry of discourse in the image-sound relationship.

In discussing Genette, we will look at the overall construction of his theory of narrative and the rhetorical categories as he defines them. With each step, we will put his theories in conversation with film theory and sound theory and oftentimes use it as a case study on our *Koyaanisqatsi* sound design where relevant, demonstrating the strengths in Genette's work in analyzing film sound. Were we to leave the case study for after the illustration of Genette's theories, we would be doing a disservice, as the *Koyaanisqatsi* case is useful in that it illustrates Genette's concepts when they are applied to sound, so rather, we put them in dialogue as we move through the theory, iteratively returning to our artistic work as a means of illustration. Other films are cited as well as further examples.

In Gérard Genette's *Narrative Discourse: An Essay in Method* (1983), first published in 1980, he posits an unequivocal host of terms that might capture the "general…at the heart of the particular" to describe and analyze narrative, based on his analysis of Proust's *À la*

recherche du temps perdu (*In search of lost time*) (p. 21). If "the general is at the heart of the particular" (p. 23), Proust serves as Genette's "particular" case, one where the beating heart is challenging prose that is vastly descriptive, jumps in tense and time, and contains sentences that go on seemingly forever in a fashion that is so nonlinear that Walter Benjamin (1968) called them "the Nile of language, which here overflows and fructifies the regions of truth" (p. 201). Proust's writings address large philosophical themes like perception, time, memory, sexuality, and art, while pages upon pages of ink describe the tiniest details of quotidian life. Approaching 4000 pages, depending on the language/translation, there is a surrender one must make to get through it, a release "from the typical pressures of reading" as further elaborated by Oliver Munday, the design director at *The Atlantic* discussing his Covid-19 quarantine reading of Proust: "I stopped tracking my progress and let go of any expectation of what I'd read next." We are struck by Genette searching to reframe narrative discourse and his choice of Proust, who uses literary devices to structure and move within time, provide excruciatingly precise discursive detail, create a swirl of narrative via the obliteration of periods and commas, and demand surrendered reception by the reader that reverberates long after - so much so that Munday notes, when later affected by a glorious moment in nature, "I grasped for poetry to describe it, before letting go, remembering that experience precedes language. For the moment I just watched." Thus Munday is inspired by Proust to surrender to an experience and Proust, himself, embeds such a moment in his largely autobiographical central character, Marcel, who while reading is struck by words on a page, that:

> filled me with […] a joy I felt I was experiencing in a deeper, vaster, more unified region of myself, from which all obstacles and partitions seemed to have been removed (Proust, 2004, p. 96).

Munday reading Proust and Proust's own Marcel are both reminded that experience precedes language, brought into essences that expand the mind and resonate with abstract phenomena. For Marcel it comes from the words of his favorite philosopher-writer, Bergotte; for Munday it comes from Proust's nonlinear twists and turns, and complex rhetorical figures that call for surrender to the currents of his literary Nile.

The intersection of these notions parallels with film sound and its ability to manipulate time, provide discursive detail, obliterate periods and commas that mark the frame, the shot, the scene and the screen, and the reverberating impacts of filmgoing that can be so difficult to

describe and are so often shepherded by sound, as we have earlier discussed. In our effort to approach the mystery of sound so that we might know the mystery of the whole universe (Inayat Khan, 1996) sound researchers might find a strong parallel to Genette, whose discourse on narrative searches for an uncoverable "general" because "the knowable is at the heart of the mysterious" (Genette, 1983, p. 23).

In his quest, Genette is, first and foremost, concerned with the ambiguous nature of the term *narration* itself and how that ambiguity manifests in narratology. "Analysis of narrative discourse," Genette understands, involves an examination of the "relationship between a discourse and the events that it recounts" but also "between the discourse and the act that produces it" (pp. 26-27). These two sides put the discursive elements of storytelling into conversation with two different concepts of narrative. The former can be seen as narrative defined by a series of events that comprise a story, while the latter as the event of the storytelling itself; thought to be a much older concept of narrative. Aristotle, we are reminded from our discussion in Chapter I, section 1.3, used the term *diegetic mimesis*, placing the adjective "diegetic" in the position of storytelling and the noun "mimesis" in the story (Ricœur, 1990). Also, as earlier discussed, Genette, by the time of writing *Narrative Discourse: An Essay In Method*, had admitted to waffling on the use of the term diegesis, first basing it on the filmologists *story world* definition and later somewhat retracting that position. Thus Genette now seems to echo Aristotle in addressing both storytelling and story as part of narrative. Taken together, this expands narrative and narrative discourse to a more holistic concept of narrative, more along the lines of narrative psychology, a field that "values the stories *and storytelling* in giving meaning to individuals' experiences" (American Psychological Association, 2020). We find this a very fruitful place for positing film analysis, especially with sound because the stories and the storytelling taken together put sound and image on equal ground, particularly with respect to pro-filmic and filmic discursive elements. We have earlier mentioned mixing, for instance, as a filmic discursive and performative act. The resulting recording of, say, a receding echo becomes a sound object that is at once pro-filmic, since the character is retreating through a tunnel, and discursive to the trained ear, and renders the flat screen in three-dimensional sonic space. Extrapolating from the APA leads us again to phenomenology. Mroz cites del Río,[83] for instance, in describing the balance between storytelling and story in different terms, describing

[83] From (del Río, 2012, p. 15)

"'affective-performative moments' that seem to escape," where the "representational imperatives of narrative and the non-representational imperatives of the affective performative displace each other without ever completely canceling each other out" (Mroz, 2013, p. 34). To us, this captures the moments of experience when story and storytelling, when pro-filmic and filmic discursive events merge into embodied experience of the listener-viewer and surrender. Tomasulo (1988), in his article suggesting a return to phenomenology in critical film theory argues that it may offer "the potential for a non-naïve, post-critical, semi-autonomous, and active viewer who can go beyond the text through his/her experience"[84] (p. 29). Mroz shows us that with sound, Tarkovsky also leans towards phenomenology, when she cites him in *Sculpting in Time* (Tarkovsky, 1986a, p. 158) asserting that sound "must also be 'sculpted', isolated and manipulated rather than simply 'recorded' in order to form part of an aesthetic experience" (Mroz, 2013, p. 123). Let us allow Proust's Nile to now lead us back to Genette.

Genette defines for himself three aspects of text in order to navigate the choppy waters stirred up by ambiguous usage: *story, narrative, and narrating* (1983, p. 27). *Story* represents the "signified or narrative content," whether the story element is low or high in intensity or importance. In a footnote, the translator cites Genette clarifying that for story he will also use the term *diegesis*, "which comes to us from the theoreticians of cinematographic narrative".[85] *Narrative*, according to Genette, is reserved for "the signifier, statement, discourse or narrative text itself." *Narrating* describes "the producing narrative action and, by extension, the whole of the real or fictional situation in which that action takes place". Genette is concerned with the analysis of texts and Alber and Fludernick (2014) assert that "the narrating instance represents events and existents (story), and they are thereby mediated in a particular (verbal, visual, or audio-visual) sign system (narrative)" (p. 2). To the contrary Horstman finds that Alber and Fludernick "misconstrue Genette's understanding of the concept of narrative mediation" citing Genette's position that narrative is "'stricto sensu [...] a verbal transmission' (Genette, 1988, p. 16) and can therefore be oral or written, but not – in his definition of the term – visual or audio-visual" (Horstmann, 2018, p. s10). We take issue with Horstmann, who positions Genette's

[84] This author feels additional resonances here, as this whole journey of uncovering a more useful "general" to describe how sound design functions is processed phenomenologically through my personal experience as a mixer. This of course, takes us back to hermeneutics, but embraces the unavoidable subjectivity in the analysis.

[85] We remind the reader that, as earlier discussed in Chapter I, Genette at some point resolves to using diegesis in the standard film theory sense, although in his eighties he expresses regret.

narrative as exclusionary of film based on too narrow an interpretation of 'verbal', and we rather highlight Alber and Fludernick who lean into Genette's narrating instance and, for us, Metz's represented instance, as narrative, opening Genette's analysis to the audiovisual text and expanding it to include non-dialogue sound. Moreover, for Genette, between his three categories of story, narrative and narrating, only the *narrative* category is open to narrative discourse, since story and narrating are not served by textual analysis. Applied to sound, it appears that sounds in film, as sonic elements or textual objects, would fall under Genette's *narrative* category as they are a series of signifiers, or "aural objects" whose unique characteristics work in the same way as spoken language (Metz, 1985). When Chion refers to *scansion*, which he loosely defines as something that "has the effect of marking or punctuating the discourse and aids the spectator in taking it in," (Chion, 2009, p. 489) he notes that scansion is an effect of a signifier, regardless of whether there is or is not a signified (p. 387) for instance, a car horn over a pause in dialogue, or an eagle screech, or even a dramatic pause in dialogue. Scansion may be pro-filmic like a car we see, or of filmic discourse, like an audio dissolve or the addition of reverberation. Moreover and perhaps more interestingly, sounds and silences often enter into discourse with each other or with visuals, either by their content alone or by sound mixing techniques, as exemplified by and previously discussed with the *Koyaanisqatsi* sound design example, further positioning them into Genette's *narrative*. In particular, offscreen sonic signifiers that provide counterpoint enter into direct discourse. For example, the sound of children in a playground, heard offscreen over a dark scene of terror with a bloody victim in close-up, discursively works with the visuals and other sounds, evoking ideas or emotions such as innocence lost, vulnerability, magic, witchcraft. In this example there is a signified, or a rhizomatic network of possible signifieds, but what becomes irrelevant is the diegetic construct. Are the children in the world of the character? In the character's head? Does she even hear them? Or are they more part of a thematic or purely aesthetic dimension? What *is* relevant here is the use of a sonic signifier as narrative rhetoric – where *narrative* is not story, and *rhetoric* is in the literary sense of the word, a linguistic tool open to exposure through analysis. This is not to suggest that, if Margaryta Sound Designer is the one to add those children to the scene, the question of Genette's *narrating* category would be irrelevant. *Margaryta* would be placing the children into the textual discourse, who is she, and why did she put it there? These are questions of *narrating,* and narrative as discourse "lives by its relationship to the narrating that utters it" (Genette, 1983, p. 29). However, with the exception of obviously

reflexive filmmaking, the *narrating* is generally invisible to the audience, even if skilled eyes and ears can and do detect traces of the narrating in the audiovisual text. The signifiers, including all acts of mixing,[86] are what create the audience's sonic encounter and color their visual one into a singularly multisensorial experience…this *is* the text and it is here where we find the discourse. Therefore, through Genette's notions, all sound elements, once embodied in the audiovisual text, are narrative ones, even music. Moreover, diegesis, or rather story, is outside of discourse analysis. Applied to film theory this would certainly represent a huge departure from the hegemony of diegetic theory when discussing sound, for as Proust (2004) notes in his fiction,

> Perhaps the immobility of the things around us is imposed on them by our certainty that they are themselves and not anything else, by the immobility of our mind confronting them (p. 6).[87]

However, such a departure from hegemony would certainly be a hugely liberating one.

Further breaking down the "problems of *narrative*"(Genette, 1983, p. 23) into a division he initially credits to Tzvetan Todorov[88] in 1966 and then modifies, Genette cites verb classifications as a rhetorical mechanism for more detailed clarification of narrative discourse: *tense*, *mood* and *voice* (p. 31). Let us look at these one at a time.

Tense relates to the temporal "relationship between story time and discourse time." Again, citing Todorov, Genette notes that, besides the ways in which time is depicted and manipulated, Todorov's *tense* includes the "time of [narrative] enunciating" and the time of narrative "perception". While Genette finds these latter aspects to "exceed the limits" of Todorov's definition of tense, when we are examining the filmic text they seem entirely relevant and sound plays a critical role in this space. Simply stated, an acousmatic sound has a specific timing of entrance and exit, or enunciation, a volume curve, which impacts the timing and intensity of perception, and when shifted alters the time at which a discourse is expressed or entered. This is as fundamental to sound editing and mixing as it is to frame shot sizes, camera movement, and picture editing. For film sound, the timing relates directly to the syntagmatic relationship between all audio and visual signifiers, including mixing choices. Changing order,

[86] And therefore what is sadly, reductively, and too often referred to as "style".
[87] from Swann's Way, the first of three volumes of the series In Search of Lost Time.
[88]Tzvetan Todorov, "Les Categories du récit littéraire," *Communications*, 8 (Todorov, 1966)

in/out points, length or volume strength therefore changes *tense*, and meaning, or as we have been defining the receptive dimension in the theater, *the flow of emotion and cognitive perception*: the interplay between the enunciation within the audiovisual text and reception, conscious or otherwise. Therefore, with this point we take issue with Genette's dismissal of Todorov, arguing that, at least with film, the timing of enunciation and perception are critical aspects of *narrative*, and should fall into his subcategory of *tense*.

Genette goes further into narrative tense to examine other concepts relating to time in literature. First is *temporal anachrony* which he defines as the tension between the temporal orders of story and narrative, which include *prolepsis* and *analepsis* (p. 40). In film this most commonly deals with story time vs. screen time in terms of linearity: at various points we may anticipate the future, a *prolepsis*, or refer to the past, an *analepsis*, in a story (p. 40). Moreover, this anachrony is further characterized by *reach*, a measure of how far we reach forward or backwards in time. For example: *In written literature, as we neglected to say in the introduction to this thesis, this often happens within a sentence or even a chapter, which we will address shortly.* Note that within our italicized sentence there is a phrase that serves as a prolepsis with quite a far reach back to the introduction, and another as an analepsis with perhaps a short reach forward, but both working as qualifiers to the temporal level of this paper overall, which is referred to by Genette as a narrative level called the *first narrative*. Back to *Koyaanisqatsi*, as a poetic film that is arguably without predetermined story, the first narrative emerges from an interplay between the audiovisual content and the listener-viewer, with the listener-viewers' cultural predispositions of course at play. Perhaps it is a narrative about ecological destruction, or individual loneliness, or the greatness of human achievement against nature. Note that none of these are *story*, but are addressed in the themes we captured through coding. Within that first narrative, we have several anachronistic moments in our sound design example, like the prolepsis of the sound of people screaming while the rocket is still on the ground, long before the explosion, or the analepsis of carrying a cheering baseball crowd sound over office shots, no longer in the space of the baseball stadium. The repetition of the phone ringing drags or lengthens, in a way, the perceived time of the shots of individual people. Tense is utilized through these devices and thereby either alters or becomes part of narrative. This can happen simply across a cut or over an entire film, for instance the film opens visually with a rocket and returns to one at the end. Because sound and image can travel on their own planes of time, and sound is multilayered, the discourse available through *anachrony* specifically and *tense* in

general is exponentially larger than with visual alone.

Additionally, Genette introduces the concept of *duration*: how we compare the duration of story vs. the duration of narrative. To compare them would be to put these into discourse with respect to tense. He struggles with not being able to measure the duration of narrative in literature, noting that written narrative is "unlike in movies…[where a film] allows us to determine a 'normal' speed of execution" (p. 86). Because of this, he sticks to a more macroscopic view, for instance, a 300-page book can tell the story of someone's single day or their entire life. But in film, one of the richest tensions cinema has to offer is this temporal discourse, where duration of a story-scene or more importantly a story-moment can be shortened or lengthened within the narrative, practically without limitation. In particular Mroz (2013) notes that "affective sounds, moments or images suspend linear temporality" (p.5) and "affect is fundamentally related to flux and change"[89] (p. 6). Later, she summarizes that,

> what the concept of duration might be seen to bring to film theory, then, is the notion of temporal strands intertwining and braiding together in cinema, as well as the process of their unfolding and expanding (pp. 35-36).

This speaks to our coining the phrase *curating the flow of emotion and cognitive perception* that for us is at the core of film authorship. Mroz's temporal strands in flux are the rhythms that determine the dance of our flow of emotion and meaning-making. They operate inside the triangle of rhetorical stance and form currents on which we ride between Deleuze's "derangement of senses" and Merleau-Ponty's "dynamic passivity" as immersed spectators who *surrender*.

Genette notes that in literature, a dialogue exchange has a kind of conventional equality between narrative and story time (1983, p. 86). In film, however, a dialogue scene is constructed around pace, alterable both in performance and in the edit room. Moments of heightened tension in the narrative are frequently stretched out in time, not only by picture editing but sound as well, marking or scanning time with various techniques that Chion might fit into scansion and c/omission in particular, again from his "said/shown" paradigm modified to our *heard/seen*, and many other sound editing and mixing techniques. For example, in the often-studied schoolhouse scene of The Birds (Hitchcock, 1963), the repeated sounds of a cyclical song, with

[89] Mroz evokes Brian Massumi in defining "affect" as "intensity, something that moves beyond meaning" (Mroz, 2013, p. 5).

nonsense lyrics sung by innocent children while our protagonist sits unaware of the birds gathering impossibly quietly behind her – that never-ending sound cycle scans the scene rhythmically and stretches out the felt time of impending disaster in what Chion refers to as "the exasperation of waiting" while also providing "temporal continuity"[90] (2009, p. 168), both aspects of Genette's *tense*. Moreover, sound is frequently used to solidify the transitional, liminal space in an anachronistic shift, like sound treatment during a dream or flashback, and here the relationship between story and narrative time is substantially complex. In the anachronism, in the liminal space, we find Mroz's flux, and therefore affect. She cites Gregg and Siegworth (2010) noting that:

> affect arises in the midst of in-between-ness. . . affect is an impingement or extrusion of a momentary or sometimes more sustained state of relation as well as the passage (and the duration of passage) of forces or intensities[91] (p. 1).

Thus by carrying us into a liminal space, a place of flux, sound demonstrates via tense the ability to create affect, tied again to experience.[92]

A final and important aspect of Genette's narrative tense is *frequency*, which deals with the relationship between a story element and how many times it is uttered. There can be: *singulative narrative* where repeated utterances occur as the story moments also repeat; *repeating narrative* where the utterances repeat a number of times that do not reflect repetitions in the story world; or *iterative narrative* where a single utterance takes upon itself several occurrences together. All of these can be found in films and are frequently executed through sonic discursive elements. In Bresson's formidable masterpiece of phenomenology in film, *A Man Escaped* (1956), each time the protagonist goes to the gated window of his prison cell to communicate with men in the yard, the same sounds can be heard in the distance: a trolly, children playing, a distant train. Similarly, sections of Mozart's Mass in C Minor also appear at repeated story moments, whenever the men are circling or walking in file in the yard. With the repetition corresponding to actual repetition in the story, these are examples of singulative narrative. For repeating narrative, both SFX and music are frequently used as sonic motifs in

[90] Temporal Continuity (and physical) were shown in our content analysis in Chapter III to correlate with sound design.
[91] (Mroz, 2013, p. 6)
[92] Affect theory would likely be a very fruitful exploration for sound, but we shall save it for future work as it is beyond the scope of this these.

this way. In Lodge Kerrigan's *Clean Shaven* (1993), a portrait of schizophrenic Peter Winter, throughout the film in a variety of locations one hears the sound of angry dogs barking. This becomes a motif within the film and reflects the ever-angrier, aggressive and disturbed state of the protagonist, making the sound of dogs a *repeating narrative* element. One question posed here, however, is whether those dogs occur in the story world repeatedly. It is possible a listener-viewer might take it for granted that these acousmatic dogs really do exist each time they are heard, rendering it *singulative*. We would argue, however, that it would be highly unlikely that they do and that, through iteration, the metaphor of the motif comes into play, dominating any sense that they are of the literal story, particularly in the absence of any visual clue that they exist.[93] The *iterative narrative* case for sound is, perhaps, the most powerful and difficult to describe. It is here where a SFX can have its strongest narrative impact. Let me begin by sharing its most frequent use in written text: the imperfect tense found in romance languages, among others, to describe repetitive action in the past, for example *j'étudiais dans cette école* which may translate to *I used to study in that school*, something that occurred repeatedly in the past. Many instances of studying in that school are captured in a single phrase. In film, outside of dialogue spoken in the imperfect or qualified with words like "everyday" or "I used to", this can be harder to find. We must dig deeper for what it means for story to repeat and along these lines we can find powerful metaphorical use of sound. For instance, in the recent WWI film *1917* (Mendes, 2019), our protagonist, young British soldier Schofield, is at yet another decision point, pondering if he shall stay to help a young woman with a baby hiding out in a completely burned-out town with danger lurking everywhere, or continue on his assigned quest, one that is certainly more dangerous. At the peak of tension, a pause through relative silence is created, an *anachronism* of screen time against story time, and a distant church bell is sounded, serving as scansion, even though we have previously learned visually that the town's church, along with everything else, is destroyed. This single sonic utterance captures the many iterations of moral dilemmas, fate, and divine intervention that the protagonist has faced along his journey, and propels a decisive resolution. While each dilemma has thus far been a unique story point, the fact that each is a moral dilemma is a repetition of *story* and that repetition represents an iteration. This particular singular sound at a critical story moment stands out even more

[93] Interestingly, one dog near the beginning of the film is shown standing, wagging its tail happily and peacefully, and yet immediately before and after we hear the presumed same dog barking and growling ferociously, solidifying the idea that what we hear is perhaps not of the diegesis or story.

because it cannot be a story element, for no church bell can sound in the story, rendering it pure affect. The sound subverts all diegesis. It is, of course, narrative and well represents tense-as-iterative-narrative, signifying the repetition of moral dilemma, fate, and/or divine intervention. Note that in all three of these examples, SFX ultimately function as metaphor and, in all of them, undermine the diegetic construct, as it is not clear whether or not they are truly diegetic as classically defined in film theory. We also find several examples in our sound design for *Koyaanisqatsi*, noting the phone ring metaphor specifically. The theme of disconnectedness between humanity and nature, and the loneliness that generates, resonates throughout the film. Our single instance of the unanswered phone is, therefore, an *iterative narrative* sound.

Moving on to Genette's narrative aspect *mood*, it expresses the "type of discourse used by the narrator," or the "modalities…of narrative representation." Extrapolating further in semiotics language and sound, with *mood* Genette seems to be indicating the paradigmatic relationship of signifiers and signs, not only in terms of the choice of sounds – choice being executed by a narrator – but in the discourse they are meant to create. Mixing is critical in establishing mood and certainly mixing can be thought of as filmic discourse or a modality of narrative representation. Even the work to render a realistic sound of a car crash, for instance, achieved through a combination of synchresis, various sounds at appropriate levels and positions, equalization, reverb, delay and other mixing actions in order to render the sound pro-filmic, that work performed to create the illusion is of course filmic discourse.[94] The resulting car crash takes on a character, if you will, dependent upon this work: it may be light and ethereal or massively destructive or anywhere along that spectrum. This is *mood*. Earlier, Genette states that *mood* deals with the "opposition between showing and telling" (p. 30). In the car crash example we are both showing – sonically – a car crash but telling of its destructiveness through sonic detail, simultaneously. Rather than being in opposition, these are intertwined tracks of narrative. Furthermore, this echo's Chion's discussion of the "said/shown" relationship and the rhetorical figures that result: scansion, c/omission, counterpoint, contrast, contradiction. Chion provides a spectrum across counterpoint, contrast and contradiction in that sound and image, or sound and sound, can enter into conversation along a continuum from a new idea that adds

[94] Though a very helpful paradigm that we shall continue to use, the binary pro-filmic vs. filmic discourse is perhaps another overdetermined contrast in this writer's opinion. Film begins as a blank slate. Filmic discourse is used to create virtually every pro-filmic, diegetic object. Filmic discursive events seem to be defined by the degree of traces of the filmmaking, like a moving camera or a dissolve. But when those traces are not evident, they still exist.

counterpoint, to ideas that contrast in some way, to ideas that are in direct contradiction with one another. While Chion's analysis is still shackled by almost exclusively looking at the tension between dialogue and everything else, it is in this rich discussion where Chion refers to rhetorical construction and comes close to framing sound design in narrative discourse and literary rhetoric. But, while Ganette's *mood* encapsulates Chion's "said/shown", modified to our *heard/seen*, it firmly places it within *narrative*. Additionally, with sound mood, we are frequently concerned with other tensions, for instance concrete mimesis vs. abstract representation. Concrete mimesis occurs when, via synchresis and mixing, sounds are made to sound as if coming from the actions on screen and from the diegetic world, as exemplified above with the car crash mix as a modality of narrative representation. Meanwhile abstract representation is found when we go beyond that in both selection and mixing of sounds. For instance, in our *Koyaanisqatsi* sound design, the dinosaur-like scream at the explosion of the rocket is an abstract representation. But isn't the synchresis of that particular sound also a kind of mimesis? The sound effect "shows" an explosion while concurrently "telling" a greater idea through counterpoint. It is, Chion would argue, a rhetorical figure of the said and shown if he might include non-dialogue as a possible "said". Another example is found in the created *physical diegesis* in the *Koyaanisqatsi* sound design, using background sounds to unify an otherwise disconnected space, the emergency environment. The extensions of continuity sounds over these disparate images is a modality of narrative representation and, of course, impacts narrative processing or again, *the curated flow of emotion and cognitive perception*. Genette's *mood*, which encompasses both the showing and telling, seems to render irrelevant many sound-for-film binary characterizations for, as we've shown, sound effects in particular can range from said to shown, from mimesis to pure story, and simultaneously occupy both these spaces as well as points throughout a spectrum, and the same is true for the binary diegesis construct. *Mood* for Genette, if applied to film sound, seems rather to lie in the subtleties of design and mixing that enable the phenomenological experience, balance rhetorical stance, or create moments or durations of affect.

Lastly, Genette's third characteristic of narrative discourse is *voice*. *Voice*, he states, refers "to a relation with the subject…of the enunciation." It points towards the relationship that the narrating has to both narrative and story, or "the way in which the narrating itself is implicated in the narrative," the narrating including its "two protagonists: the narrator and his audience, real or implied" (p. 31). While narrating in literature is traditionally thought of as the

actions and utterances of a *narrator* in the sense of a character or protagonist, locking us back into diegesis for what is diegesis if not the world of the character, narrator, or protagonist, here Genette expands to include the hands of "narrating" where at times the creative traces can be identified. Working again with our car crash example, *voice* is found in intended impact on the audience by the filmmakers, be it to carry them into a subjective spiritual space with an unusually abstract representation of the sound of the crash, or to make them jump in their seats with violence. These are varying modes of enunciation. Genette also notes that within voice one can find narrative levels, a hierarchy of narrative traces as the enunciation and reception of story is communicated over time. Thus voice interacts with tense in terms of the unfolding of story, the flux of duration, and iteration, but while *tense* deals with permutations in time, *voice* deals with the mode of enunciation, a mode created by the artist and a fundamental reflection on art.

This author is particularly interested in the idea of narrative levels, as there may in fact be a direct correlation between this idea and the audio levels, sound fades or dissolves, sonic space, immersion, and sonic direction created by film mixing, arguably entering the listener-viewer into the hermeneutic circle. The narrative instance, the enunciation, or the moment that something is made clear within the "first narrative" in written literature, Genette would call "intradiegetic" or "diegetic" (p. 228) and for film sound Chion would agree, whether the sound is onscreen or offscreen. However, if a sound is acting in a second narrative, one that is not in the world of the character, it becomes metadiegetic in Gennette's language, or "the narrative within the narrative" as described in his *Figures II* (1979, p. 202). Audio levels, direction and reverberation, or space, alter the narrative impact here, for a low level off screen sound effect, like the cheering crowd at the stadium in our *Koyaanisqatsi* work, can be of the first narrative, whereas the ringing telephone over shots of lonely people is clearly metadiegetic, or a different level of narrative. And when the cheering crowd continues over night workers in office towers, the narrative of the story is put in discourse with greater, more universal ideas: a first narrative expands, or crosses into, a second. In these examples, the narrating is implicated in the narrative, and it is here where we more deeply find *voice*. Voice therefore is key in engaging the listener-viewer into the hermeneutic process whether interpretation happens in situ or is enabled through later evaluation and discussion.

Moreover, as a tool of narrative discourse, sound is a tricky and amorphous fellow for it can activate all three of Genette's rhetorical classes at once and operate at different narrative

levels of voice concurrently. For instance, to return to the playground example in a horror film, the moment that we perceive the innocent children is a specified time at which we enter into a discourse that contrasts violence with innocence, and this moment defines *tense*. *Mood* describes the contrast itself that we have created in is pure form – violence/innocence – as a rhetorical stance, and, perhaps, the discomfort that this contrast creates for the audience, or the affect, which is served by the overall sound of the mixing. The usage of children and its intended impact is of course *voice* in its simplest sense, as we might have instead used an angry barking dog, and either way we will treat it as part of the character's point of view or perception, particularly if we are led to believe that our character hears them. But if the level and spatial determination is such that it feels so far offscreen that we are unsure of the character's point of view with regard to the sound, it is operating only at a metadiegetic level. Moreover, if we were to apply reverberation to the sound, we might believe it to be a memory and step into a completely different discourse on both the diegetic and metadiegetic level, bringing us through altered duration, liminality and anachrony with respect to *tense* (past time). The most memorable moments of films with powerful sound design make use of these very different degrees of discourse operating concurrently through sound selection, placement, and mixing techniques, as we will see later in our film analysis of *Velvet Goldmine* in Chapter V.

It seems like a close look at narrative discourse offers a more compelling view of the function of sound design than image-based positioning found in much of the canon of film sound formalism that has developed. The traditional diegesis of film theory used as a framework by Chion, more or less locked into a visual paradigm and a binary system, is rather shown by Genette to operate in an overlapping and rhizomatic network of narrative discourse. Looking through sound we even articulate Genette's rhetoric with Bracco's "phenomenological hermeneutic approach" (Baracco, 2017, p. 76), where the interweaving and inseparability of experience and interpretation is seen to be driven by *tense*, *mood*, and *voice*. Through this prism, sound design in film looks a lot more like Genette's Proustian scene, "encumbered with digressions of all kinds, retrospections, anticipations, iterative and descriptive parentheses, didactic interventions by the narrator, etc., […] able to give […] a fully paradigmatic importance" (Genette, 1983, p. 111).

The *Koyaanisqatsi* sound design example is, perhaps, a Proustian sequence. The discourse created by the addition of sound effects can be seen to contain: digressions, like a phone sound over images of people; retrospections by way of extending a sound from one image

to another, like the birds over the city continuing to play over images of computer chips and tapestries, which asks for a retrospective look-back between visual signifiers; anticipations like the early ramp up of a background sound before a scene begins as with the emergency background prior to the cut to the scene; iterative and descriptive parentheses like the repeated phone ringing over images of lonely people; didactic interventions by the narrator like those same phone rings or the screech of a dinosaur on the explosion of the rocket – all creating a fully paradigmatic and phenomenological experience that goes beyond the discourse provided by images and music alone. Whether these sounds and mixing choices emanate from the world of the frame is not only difficult to define but, we conclude, largely irrelevant to the experience.

Genette's *Narrative Discourse* provides us with a more holistic mechanism for engaging with film sound predominantly because it allows us to formalistically break down sound rhetoric in overlapping paradigms rather than categorical ones. It removes the diegetic framework by including all sound as part of narrative, even while respecting the role of narrating and story. He finds it fruitful as a mechanism for uncovering rhetorical structures within Proust and, as we have stated, Proust serves for us as a parallel for film sound in its most creative forms. Proust's snaking Nile of shifting tense, mood, and voice captures the elusive nature of film sound and the surrender to experience that both encourage. Through Genette, film sound seems theoretically liberated from image paradigms, as he has provided us with discursive rhetorical tools to analyze how sound functions within its own paradigm and in conversation with image on equal ground. Having addressed our third area of inquiry, we will move on to Chapter V where we put narrative discourse to the test via a case study of Todd Hayne's *Velvet Goldmine* (1998), mixed by this author.

CHAPTER V. CASE STUDY: *VELVET GOLDMINE* THROUGH GENETTE

To recap our work thus far, we address foundational sound theories and in many instances their relationship to other film theory in Chapter I, exposing what we feel to be weaknesses in the predominance of screen-based or image-based film sound theory. Most of our discussion was in the realm of diegesis due to the way in which sound diegesis permeates so much of film theory whenever sound is addressed. We also were very interested in Chion's rhetorical figures as they represented something of a first important step towards recognizing that rhetorical positioning is possible within the construction of sound. However, these theories and figures remain screen-based as they mostly tie to dialogue and, in fact, fall short of describing many functions that sound undertakes. Taping into results of that research, we moved to create an artistic output, a sound edit and mix, that might illustrate shortcomings in existing theory and, during that analysis, introduced some other film theory that is not traditionally applied to sound. From this process, we came to identify desired outcomes of filmmakers with sound that could be tested first by examining sound design through the prism of a technology and then via content analysis to find correlation between sound design and desired outcomes, thereby providing the research community with fuel to support statements about sound's functions and to feed further work. Having shown sound to bring narrative, emotional and thematic change to the listening-viewing experience, and hoping for an expanded theory of rhetoric to apply to sound design, we moved to Genette's literary rhetorical theory to find a different approach to the rhetoric of film sound that might better encapsulate the various levels, angles, and rhetorical stance with which sound works in film. We used moments from our artistic work to illustrate our mapping of Genette upon film sound and found that his serves as a fruitful mechanism that puts sound in conversation with image on equal ground. Now we seek to test out a Genette-type approach in the context of a case study to see how it might work against a film other than our the artistic work created for this study. Rather, we iteratively go back in time to a film mixed by this author years ago, cycling our artistic research elliptically to a former work, putting past and present in conversation.

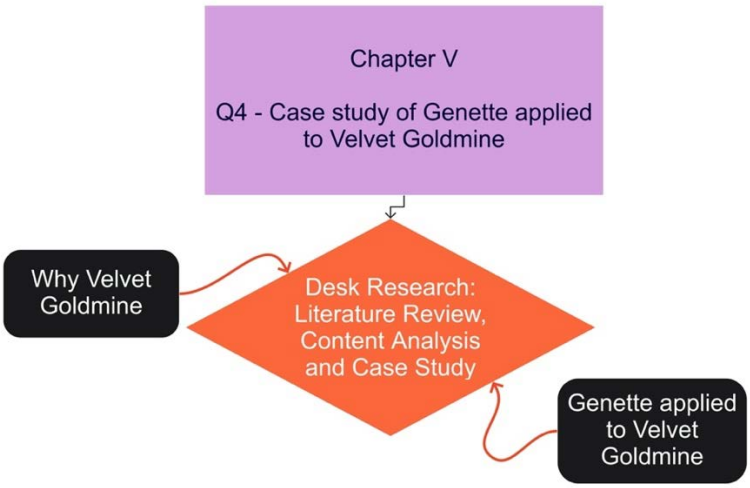


Figure 20: Chapter V Roadmap

To examine ways in which Genette's *Narrative Discourse* might function in a film sound analysis, we move to the 1998 film *Velvet Goldmine*, directed by Todd Haynes, produced by Christine Vachon and Michael Stipe, edited by James Lyons, musical score by Carter Burwell and Craig Wedren, with supervising sound editor Paul P. Soucek and re-recording mixing by the author of this paper (Haynes, 1998).

Velvet Goldmine Image 1: Title Shot

There are three primary reasons why we chose this film. Firstly, it is a Proustian narrative for it is certainly, as previously mentioned (and now *singular-iteratively*), similar to Proust in that it is "encumbered with digressions of all kinds, retrospections, anticipations, iterative and descriptive parentheses, didactic interventions by the narrator, etc., […] able to give […] a fully paradigmatic importance" (Genette, 1983, p. 111). Second, through a maximalist approach to sound design and image, it presents a Wagnerian aesthetic and sensory experience, in a kind of cinema that James Lastra (2008), based on his sensory-focused analysis

of *Apocalypse Now*, might agree is

> a resolutely audio-visual one, a cinema born not out of painting or theater but out of the spirit of publicly performed music, where drama and audience exist in a shared space united by a sonic envelope. This cinema's nearest relative is, perhaps, the Wagnerian musical drama […] because both harbor the wish for a reintegration of sensory experience with the whole of human being (p. 127).

Not only do the visual and sonic elements of *Velvet Goldmine* align with Lastra's take on the technologically-empowered modern sound film, but the very content of the film – *story* and *narrative* and *narrating* – is grounded in culture and pop-music-as-Wagnerian-opera glam rock-and-roll, which presented a reflexively theatrical musical space of immersion defined not only by the concert stage but by the transformative immersion, into the lives of countless fans, of the values and societal subversions that glam represented. This deeply reflexive quality is firmly expressed in the opening text card on black and in silence via a statement that points to sensory experience, "Although what you are about to *see* [my italics] is a work of fiction, it should nevertheless be played at maximum volume." The sensorial, particularly the sound, frames the experience while the experience is glam, Wagner, in the cinematic audiovisual form.

The third reason for choosing *Velvet Goldmine* is the unique perspective that this writer can offer. As the re-recording mixer of the film, I am able to deconstruct the sound work, exposing the traces of narrating and therefore shedding light on targeted narrative techniques and functions of sound in the film. These might otherwise be quite off limits to the general public who, due to the overwhelmingly Wagnerian discursive elements, most commonly *surrenders*, just as Munday was released by Proust from the "typical pressures of reading." The *Velvet Goldmine* audience is compelled via the sound to let go and be enveloped in an experience that can be thought of as phenomenological in nature because, consciously or not, our work honored "the founding gesture of sound design" according to Lastra (2008), which was first a separation of sensory experience with representation and then,

> [a] compulsive linking of the two in an indissoluble unity that appeared to efface representation, and by effacing it restore reality and to offer it to us in its fullness (p. 90).

Our previous discussions of affect, phenomenology, dynamic passivity, interimplication and temporality all find resonance in Lastra's words.

To summarize the film, *Velvet Goldmine* is set mostly in the world of 1970s glam rock, with characters loosely based on known icons such as David Bowie, Iggy Pop and others. Much of the story is told through the eyes and ears of the protagonist Arthur Stuart (Christian Bale), a journalist who, in 1984 is charged with researching the life and disappearance of his own teen-idol obsession, glam star Brian Slade (Jonathan Rhys Meyers). He accomplishes this by interviewing Brian's former managers Cecil (Michael Feast) and Jerry (Eddie Izzard), Brian's ex-wife Mandy (Toni Collette), and Brian's co-performer and sometimes sexual partner Curt Wild (Ewan McGregor). Thus an investigative structure going back in time, à la *Citizen Kane* (Welles, 1941) frames the *story* and sets us up for a journey in time that is beyond non-linear, but rather kaleidoscopic.

We will first present the opening 16 minutes of the film. Since we are looking at sound that is extraordinarily layered and detailed, this analysis requires quite a bit of textual description to be as comprehensive and clear as possible. We encourage readers to access the film with a good listening apparatus, first watching this section, and then rewinding to hear/watch again, along with this text. Still, we strive to provide description here that will make the analysis clear, even without screening. We will then move on to a brief sequence near the very end of the film.

Mixers know that often, the most difficult and most rewarding moments of mixing are in the liminal spaces between scenes, referred to as *transitions*. These opening sequences of *Velvet Goldmine* feel as if they form a near constant transition of approximately 16 minutes. *Story* elements are there, but the *narrative* (Genette's narrative) and the *narrating* dominate the experience, so much so that many listener-viewers are not quite certain of the story unfolding before them. Surrendering to a kaleidoscopic experience seem to be the goal. As previously mentioned, the film opens with a text card stating that, "Although what you are about to see is a work of fiction, it should nevertheless be played at maximum volume." As it fades away, acousmatic sounds over black grow, building a sound montage, with various sounds mixed in and out of the perceptual limits, at once featured, later pulled back. Still over black, these sounds include feedback, radio frequency tuning, a light kind of patriotic marching band, radio static, magical chimes, whooshes, low tones, submarine sonar sounds, and voices that are distorted as if playing over radio-frequency broadcasts, but undiscernible in language and positioned around

the 360º theatrical space for full immersion.[95] After some time, a lulling musical score creeps in. Stars appear, and a shooting star is highlighted with sound through synchresis, via a few descending notes in the score and a SFX of a mid-frequency-range whoosh. The musical score, thus far without a discernable tempo or rhythm, builds with several overlapping and rising notes into one, landing on a tonic with the introduction of a narrator's voice.

Let us pause and look at just a bit of the narrative discourse thus embedded in the opening 30 seconds. Visually we have text, presumably coming from the writer/filmmaker and therefore representing a voice that is not part of the story, *extra-diegetic* to Genette. Then we have black, then stars, and that's it visually. It is sound who enters into discourse, primarily with itself, as *narrative*, beginning with sound in our mind as we read the text, our imagined sound activated in the secondary auditory cortex as discussed earlier in this work, and possibly the chuckles of other people in the audience. The subsequent sound montage is a collage of signifiers, all of which likely have targeted meanings, although the interplay woven by sound editing and mixing alters, or controls, the discursive *tempo* by creating the precise moments when one signifier or another is registered, or rather, enunciated and perceived. The utterance, therefore, is in the hands of the mixer. This also speaks to various levels of narrative as captured by Genette's *voice*, as modes of enunciation. Ideas invoked include: technology, sound recording, radio broadcasting, communication, authoritarianism, lurking, exploration, distance, time, history and magic. Note that these "ideas" function only as fleeting essences[96] materialized or concretized via sound alone. As with some of the sound montage in our *Koyaanisqatsi* sample, and as with Proust for Genette, the sounds are both showing and telling simultaneously, *memetic* and *diegetic*. The *mood* is derived by the overall soundscape, including selection and mixing of music and SFX, the gentle pace, the dark screen, all modes of representation. That these sounds, particularly the distorted spoken words, are placed around 360 degrees is a technique used to create Lastra's lofty "shared space, united by a sonic envelope" with the goal of "reintegration of sensory experience with the whole of human being." And with that, the voice of a proper narrator, a woman (Janet McTeer), is simultaneously heard through three speakers – left, center and right – disassociating her voice from the grounded nature of center-screen,

[95] The film was mixed in Dolby digital, 5.1.
[96] Again, looking towards phenomenology.

> Histories, like ancient ruins, are the fictions of empires [long pause] while everything forgotten hangs in dark dreams of the past [pause] ever threatening to return (timecode 01:24).

Voice, *tempo* and *mood* are all critically at play here. First, we have a new extra-diegetic narrator who seems to be different than the one who gave us the earlier text card, for no logic would support that she is read earlier yet heard here. Next, there is the paradigmatic choice of the voice. On the mixing stage, we had two voices from which to choose, Janet McTeer's and a man's. His offered a sound trope of a voice-of-god documentary narrator with authority but some poetry as well, while hers, equally convincing and poetic, offered something much more subtle; a smoothness that is nearly seductive, an inflection that makes one feel she is letting you in on a secret, and an alto that approaches gender neutrality. This sonic choice made during the mix, as we tested both in the room, represents a completely invisible filmic discourse that defines the *mood* of this moment, the modality of narrative representation. That her mellifluous voice emerges from multiple speakers further defines *mood*. She is at once showing, via spoken text, and telling, via performance and mixing, and it is in this tension that *mood* is established, as Genette would argue. Of course, *voice* is paramount as well, for in the "relationship that the narrating has to both narrative and story" McTeer's voice establishes absolutely nothing about story, but quite a bit about narrative, as the entire statement is a metaphor for themes embedded in the film, given by a voice that already knows the secret. This, therefore, reflects a rhetorical stance of the filmmakers that applies to the overall film, embedded in just a few seconds. Genette might further add that a secret, hinted at but not fully revealed, is a *prolepsis*. Moreover, in the text above, we note the *pauses* and return to *tense,* raising a critical question about pace and relative silence. Dialogue and sound effects are meticulously placed, and pauses are created for many reasons but, first and foremost, they establish rhythms that can render discursive time rigid or fluid. In this case, it is fluid, as time seems to slow over pauses, and accelerate when sounds cascade[97]. This evokes Mroz's notion of "temporal strands intertwining and braiding […] unfolding and expanding" (2013, pp. 35–36). McTeer's pauses, fine-tuned in the mix when put in discourse with the music and SFX, are but one example of countless moments of anachrony throughout the opening 16 minutes; some

[97] Even in film music, composers conduct tempos in a fluid manner, often driven by visual cues in the form of a strip that scrolls across the screen alerting the conductor to an oncoming "sync" point.

subtle and others obvious. Moreover, we have entered in a Gricean talk-exchange via the cooperative principle. The all-knowing voice will enlighten us, we believe, and we trust that we will come to understand. But it is not just the words, but rather the ways they are being performed and mixed, with pauses and tonalities and multiple speaker placement beyond center screen, rhetorical of course, that provide the inference of all-knowingness that will eventually unfold. Thus, layers of meaning are constructed via sound.

Under her final words, a low rumble builds along with score, joined by a technological warble sound almost like an alarm, another magical chime sound, accelerating from beneath her final word to the visual arrival of a passing spaceship, with subwoofer and mid-range whooshes. The buildup of sound is a prolepsis not simply because it precedes the arrival on

Velvet Goldmine Image 2: Spaceship soars
TimeCode 01:45

screen but because we don't know what it is until we see it.[98] Time is also accelerated out of McTeer's final words as "threatening to return" ushers in the spacecraft. As the ship descends, a harp becomes prominent in the score and moves to arpeggiated 16th-note triplets, marking time strongly and for the first time since the opening of the film. We journey down, ushered by a harp, a sound often associated with things angelic, or heavenly, a new *voice*. As we descend on nighttime Dublin in 1854, indicated by visual text, the sounds of crickets appear.

[98] Chion describes this as an acousmatic sound becoming visualized.

Velvet Goldmine Image 3: Oscar Wilde Home, Dublin
TimeCode 02:10

This is the first moment we feel we are entering some semblance of a location, a scene. From above we see a baby in a blanket on the ground, the silence of the night is broken by the sound of an opening door. The musical triplets fade away and then the score, which ends on an unresolved chord[99], provides a suspension of time and a time of suspense. A maid, appears. In ADR[100] she states, "lord in heaven" in a nighttime whisper, with the camera POV still up in the air above the street. From camera position, one could not hear such a whisper, but it is mixed louder than realism as a guarantee for the enunciating moment. This exemplifies the discursive hand behind the pro-filmic dialogue. Horse hooves are heard in the distance, off-screen sounds that do more than signify "horses" but in discourse with the visuals roughly establish or support historical period. They serve up meaning as icons, symbols and index simultaneously; iconic as horse-on-cobblestone, symbolic as a learned symbol of a time period, indexical implying that people are out at this time of night who did not witness a spaceship and a baby. As the image cuts in closer on the maid lifting the baby, a church bell is heard that crosses the cut, joining the crickets to establish continuity of time, as did the emergency sounds in our *Koyaanisqatsi* design. It rings three times. But is this diegetic? It is the middle of the night. Church bells generally do not announce 03:00 in the morning. Chion would call this a border space, a SFX straddling diegetic and non-diegetic. But while the audience might perceive it to be diegetic, it can't be. This is much more like the bell discussed with respect to Mendes' *1917* (2019), an iterative-narrative device that captures in one SFX a theme or idea that appears repeatedly in

[99] A dominant V7 chord which is a musical chord that begs to lead to a different chord, usually the I chord or tonic of the key of the music.
[100] Automated Dialogue Replacement, which is recorded in a studio, sometimes by a different actress. It can be mixed to sound as "of the scene" as possible or it can be elsewhere on a spectrum between mimesis and abstraction.

the film – divine intervention. Providence tells us it is inevitable that "dark dreams of the past" are "ever threatening to return", *ever* evoking the past and future imperfect tense, and the *return* embodied in the parallel forms of aliens and Jesus via a baby from heaven, all this put in discourse by image and sound, sounds including the word "ever" from the narrator, "Lord in heaven" from the maid, a spaceship with technological sounds including an alarm, and an impossible church bell. The iterations of divine providence across the entire film, captured in this bell but yet to appear, include various human essences – queerness, liberation, glam and pop culture – all of which are laid bare before us in the opening 16 minutes.

The first bell is tuned to the major 3rd of the established tonic in the music. It is therefore in harmony, or consonance, which is a discursive sound design choice since dissonance would alter Genette's narrative through a change in *mood*. The maid continues, "Madam Wilde, Richard, come quickly!" The second church bell sounds. A whoosh sound draws the maid's attention to the nighttime sky, with a different kind of warble air sound, timed with a small flash of light in the clouded backdrop. This *tells* us the spaceship has left both by sound and by drawing our attention to the subtle flash to make sure we notice it, an act of *tempo and voice*. A major fifth is added to the score such that there is now a major chord created by the two tones of the music and the church bell, a *mood* of resolution. Foleys pre-announce the arrival of the Wilde's, giving time continuity which is of course *tempo* since it establishes unity between screen time and story time, regardless of image discontinuities. The church bell continues for its third and last chime, louder than the other two, punctuating the divine gift of the child. As the Wilde's arrive at the doorstep and pull back the baby's wrapping, a sliding musical tone finds its landing just as an emerald broach on the baby's clothing is revealed.

Velvet Goldmine Image 4: Emerald on Baby Wilde
TimeCode 02:32

A fully resolved tonic major chord is established in the score. The foley of the blanket

and even a subtle pop of the catching of the cloth on the edge of the broach highlight the reveal, along with notions of magic in the score, including choral voices which return us to the divine in the human. The discourse in this moment is rich. Crystal clear foleys of sounds that shouldn't be heard function as materializing sound indices that concretize texture and pull us into some locus between Deleuze's "derangement of the senses" and Merleau-Ponty's "interimplication between body and world". The magical and human musical voice, and a slide to a beautiful chord, join the soft textiles and collectively point towards a jewel in close-up, keeping us in a space of divinity, providence, delicacy, and intimacy. Throughout this short-lived nighttime location, the *mood* established, the *tempo* played with, and a *voice* that seems to shift to the divine sets us up for what's next, the idea that Oscar Wilde, through divine providence, came to us from outer space as a savior who, through his art, will liberate us in presently unimaginable ways. With images of this scene rooted almost purely in *story*, this full discourse is provided by sound.

A new low frequency tone or chord comes in ahead of the dissolve, the transition, adding to the sounds as the choral vocals leave, we dissolve to images of school children standing up one by one to announce their hopes for the future, dressed in 19th century clothing, with name tags. The first few voices fade in along the slow dissolve, along with arpeggiated 8th notes in the score. Again, the arpeggios disappear just before, at full level, the last of the five boys, wearing a name card that says "O. Wilde" and with the emerald broach on his collar, announces "I want to be a pop idol." The boy exchanges a look with the teacher, and his emerald sparkles, with a magical sound.

Velvet Goldmine Image 5: Young Oscar Wilde
Timecode: 02:50

It is the suppression of the voices of the other boys and the suspension of the music for Oscar's words that are most noteworthy here. The *mood* established by the mixing allows

music to dominate during the entering transition, connecting us to the previous scene in our first major ellipsis, an anachronism that serves as a jump forward of around 10 years of story time. This leap is complemented by the slowing down or lengthening of duration through musical suspension and a pause, plus a relatively higher mixing level of dialogue that allows Oscar's line to land firmly and cement itself in our heads.

The moment accelerates out with a new musical entrance, pads of tones that continue over the title card "One hundred years later…" with the sound of children playing in a yard. From above, we see and hear that they harass a classmate. In a narrative act, an offscreen hand-rung bell is heard that both establishes a school yard and announces the end of recess, serving two functions related to story. The boys are in uniform. Still overhead, the camera moves up through a tree, and the sound of leaves rustling in the breeze highlights the visual and fills the high frequency space as children's voices continue. The breeze is materialized, the texture of the leaves felt. All along, the foleys of the children hitting one another and their feet on pavement have crystal clarity, rendering the scene very real and discursively highlighting the oppressive nature of their actions. There are layers upon layers of *mood* in this moment including the children's voices and school bell that at once evoke youth, innocence, and bullying, triggering timeless memories good and bad[101] and implicating the audience in the moment. The breeze, which is invisible, something we often only know of indexically, is beautiful, cleansing, a sound of oxygen, and carries us to the next cut where we arrive at the ground, the harassed boy laying over a sewer grate, alone. We still hear the rustle of the leaves, but not the kids, which is taken over by the sound of water rushing in the sewer and the boy's foleys as he notices something in the dirt, clears away earth and sticks, and finally digs something out.

[101] A discussion of audience memory as an example of both anachrony and the use of tense is beyond the scope of this discussion, but certainly may be a fruitful path to pursue.

Velvet Goldmine Image 6: Emerald in Young Jack's Hand
Timecode: 03:28

Choral voices anticipate – *prolepsis* - the reveal of the emerald broach, an audiovisual *singular iteration* having been previously established and returning with both image and sound in the form of chimes. He has pricked his finger on the broach. Only musical score is heard at this point, the shift out of background sounds signaling a move forward, *tense*. The narrator (Janet McTeer) returns: "Childhood, adults always say, is the happiest time in life." Although this line continues into the next visual scene, its placement here before the dissolve provides a great sense of irony and joins the two spaces in a type of thematic continuity as shown in the *Koyaanisqatsi* example. Moreover, this is another example of Todorov's *tempo* as it relates to the precise placement of the enunciating, generating irony through perception via the timing.

Visuals dissolve to a painted fantasy space of a path in a forest, clearly designed as a surrealistic theatrical set. The same boy in school uniform, seen from behind, walks towards sunlight upstage, and we hear layers of forest birds in reverb, but no foley for walking, no indexical sounds. The inclusion of dreamy ambient birds without the indexing of the character discursively places us in an oneiric space, liminal, surreal, as the narration has paused for us to take it in, allowing a hypnotic loop of music to move us forward. She continues, "But as long as he could remember, Jack Fairy knew better."

Velvet Goldmine Image 7: Shots from Jack Fairy Fantasy Sequence
TimeCodes: 03:37 / 04:06 / 04:17

The image dissolves to deep blue, a new liminal space, no more birds, another pause

in the voice-over. The boy's close-up profile appears in blue light and as he walks gentle quarter notes in the music mark time like a march to destiny. "Until one mysterious day when jack would discover that somewhere there were others quite like him," she says, a small pause as music builds, then "singled out for a great gift." There is another pause, the music now in a vamping crescendo, a loop that increases with volume. The boy applies blood from his pricked finger to his lips. "And one day," – another pause, with all time-marking features in the music disappearing and only a suspended musical chord remaining – "the whole sinking world" – another small pause, and then almost silence as music fades away – "would be theirs." Silence is left for about a second as the boy, having put blood on his lips as lipstick, smiles in a mirror. The pauses throughout, or rather the spacing out of the voice-over, coupled with the rhythmic and melodic repetition of the music, are discursively intoxicating and dreamlike while accomplishing Chion's *exasperation of waiting* as previously discussed with respect to music and the very determined song in Hitchcock's *The Birds*. The combination of modes of representation and enunciation of the various elements put us in an oneiric surrealist trance while heading towards something that we anticipate. To Genette, this effect is created by a kind of prolepsis that forges suspense, while also giving us an *anachrony* between the time-marked pattern in the music and the greatly stretched out reading of the words, two simultaneous sonic temporal vectors, then another shift in *tempo* as rhythms disappear for the last line.

We pause a moment here to bring into conversation other theories previously discussed. This one moment serves as a good case to do so, although certainly these paradigms exist throughout the entire case, even if we do not point out each instance in the interest of focusing on Genette's rhetoric. For starters, *duration* here is altered repeatedly via the pauses in voice against the gentle rhythm of music, creating affect for the listener-viewer which "arises in the midst of in-between-ness," according to Siegworth and Gregg. Taken into this fantasy or dream world, into an abstraction, we engage in Walsh's "dynamic passivity" between "losing ourselves" (Deleuze) and "interimplication between body and world" (Merleau-Ponty), in the *surrender* zone, forged by the abstraction of image and rhetoric of sound, particularly embedded in the music and the quality and pacing of the voice-over. Inside this story world, especially with the slow pace offered, we can also vary our state, as with Baracco's phenomenological hermeneutical perspective, by magnifying or diminishing our intentionality of perception (Baracco, 2017, p. 48), reading the code of narrative layers, or not, back in the hermeneutic circle begun at the beginning of the film and reentered a few times already, or breaking from

the circle in favor of surrendered experience. We elaborate here to point out that the rhetorical figures described, which are dense and varied, are not mutually exclusive with theories of reception, hermeneutics, phenomenology, etc., because sound offers, or rather invites, rhizomatic engagement.

Continuing, the silence on the boy's broad smile and the overall dreamlike state is suddenly penetrated by a hard sound cut, the entrance of a loud, saturated music piece, what Lucy Fischer (1985) would call an abrupt tonal contrast in her article on foundational sound editing techniques and Vertov. The first four screeching 8th notes of 70s rock guitar play while the image is still on the smiling boy in his liminal space, a mismatch and anticipation, a clear prolapse in terms of preceding the cut, and an ellipse as image cuts to 1974. That this ellipse happens in sound before picture presents a discursive connection from the lip-painted boy to his older self, whom we see not right away but a few cuts into this credit sequence montage.

Velvet Goldmine Image 8: Image from Title Sequence
TimeCode: 04:26

The music dominates at full volume and saturation, feeding all speakers in the theater, with occasional foleys of running shoes on pavement and happy excited voices of teenagers running through the streets. The mix of the song, done with analogue tube amplifiers from the 1970s which allows for heavy distortion in an aesthetically acceptable way, is fed to all speakers to completely immerse the listener-viewer in the rock-and-roll. This is highly discursive rhetoric on many levels. The "sound" of the music places us in a period, and the pace and surround aspect create an embodiment into the frenetic energy of the teenagers running on screen. It is temporal, metaphorical, and overwhelming enough to ask for *surrender*. We carry the smile of the boy in us from the first early beats before the time cut, to the smile of the running teenagers in outrageous glam-like period costumes, to our own smiles. Eventually there are four quick breaks, built into the music, which allow for a little ADR dialogue to establish

that a dandy walking through the streets with the emerald broach is Jack Fairy. Music picks up right as it left off, setting a cadence in which the ADR dialogue rhythmically fits into the musical break, almost as if it is part of the song. This challenges the notion of the song as nondiegetic and, rather, the song becomes what these youth hear in their hearts. More than six minutes in, the audience has not yet been given a chance to breathe, and the kaleidoscope continues.

As the song ends with teens standing in a line for a concert, a male voice comes in, clearly a TV reporter-type performance, “The streets of London are ablaze….” The camera ultimately transitions to, and lands on, a TV, although we are in a liminal space until the 4x3 framing and curvature of an old TV screen become clear with a reporter facing the camera providing important *story* information. There has been a subtle shift in the reporter’s voice. It was initially nearly full, but not as full as a voice would be in person. Once clearly emanating from a television speaker, the voice is thinner, but still not as far as that mixing effect might be taken. As the reporter describes the performer Brian Slade (Jonathan Rhys Meyer) whom all the teens have rushed to see, the camera cuts to Brian in close-up with the TV visible in the background, and the TV sound gradually fades to proper perspective with room reverb and an even thinner frequency range.

Velvet Goldmine Image 9: Brian Back Stage
TimeCode 07:40

The pace of this transition is all about *tempo* and the enunciating of narrative, but achieved through mixing; equalization, volume level, and reverb techniques. Through this transition we have arrived in a new space, we are back stage at the anticipated concert. We know in part from the mise-en-scène, but more due to new music off at a distance, drowned in reverb, awash nondirectionally from all speakers, with no high frequencies and lots of subwoofer, which together establishes that a band is playing onstage elsewhere, nearby. Here

is an important point to mention regarding mixing, particularly reverb and equalization. Embedded in how we hear are sonic signifiers that relate space, material, room size, barriers, gender, age, and many other things. The gradual change of EQ and reverb on the reporter is a spectral shift of signifier[102] and this exemplifies mixing techniques as a source of signification. For example "rock-music-in-church" and "rock-music-at-a-club-heard-from-backstage" are different signs with discernably different signifiers, even if the music is the same.[103] Backstage, the music vocals, undecipherable in all the reverb and equalization, become a kind of wailing tension sound as does the pounding of low-frequency percussion. Thus they act on an emotional level due to the mixing, an example of *mood*, and the *flow of emotion and cognitive perception*, a reminder that to call the music "offscreen diegetic" or "source music" is incomplete and irrelevant to how it functions. Next, a foley of hairspray highlights the actions of Brian's assistant. The TV announcer tells us the performer has had premonitions of being assassinated on stage and Brian instructs his assistant to turn it off. Rhythms, *tempo*, are supported by SFX with a sudden banging knock on the door, a door open, and music from the stage made louder and clearer as the door is opened, another mode of narrative expression. Brian is told "it is time." Offscreen, the door closes for time continuity and to signify that Brian and his assistant are left alone. Thus the door close acts both as tempo and as a *narrative* sound act that conveys *story*. As Brian turns, nervously, towards his assistant, foley's of his sequin-filled costume make a punctuated sound. This exaggerated material-sound-indexing of the character highlights the drama while stimulating us in a tactile way. This is not realism at all but rather another mode of narrative expression, even if synchresis gives us the sense that it is pure story or pure mimesis. Then an ellipsis jolts us with a hard cut on picture and sound to moments later, Brian walking in the hallway toward the stage with accentuated indexical footsteps at high volume and pace, sounds of fans screaming, and an MC announcing the band live but with a voice buried in all the loudness that awaits.

Let us pause and address *voice*. We first felt the voice of the filmmaker with the text

[102] We might additionally argue that the slow change is an act of scansion and voice, marking a transition less of scene or space, but of perception as the sweeping change is gradually enunciated with a flow of time. As a mixer, these changes are executed rhythmically, instinctively, sometimes with eyes closed. There is a curvilinearity to the sweep that is only "right" when it simply "is".

[103] The treatment of offscreen music signifies "through walls, hard rock, club, loud" even though the level is not loud. Is this syntax or paradigm? Another duality that may be overdetermined, for it seems to be both, hence challenges for sound semiotics.

card on black, then we were put in the space of a female narrator as the storyteller. Then perhaps we entered the world of Oscar Wilde but we leapt forward to Jack Fairy. At that point, with the narration over the boy's story, it seemed that Jack Fairy might land as a central protagonist, particularly as we also see him decked out in 1974. We are then in news land, via narrative and story delivered by a reporter reflecting the voice of middle age at a time of cultural revolution, but then land with yet another character, Brian Slade. At this point, we are pretty sure he is our protagonist. But it shall shift again. This all represents voice as in a standard, literary sense. But with sound and Genette, *voice* is more than that. As noted earlier, it is "the way in which the narrating itself is implicated in the narrative," the narrating including its "two protagonists: the narrator and his audience, real or implied." Thus, the creative traces that can be found through all these transitions, along with their intended emotional and perceptive impacts, ushering along the baton that is passed from one voice to another, which happens to define the entire narrative structure of the film analogously to the passing of the emerald, are all *voice* just as much as they represent *mood* and *tense* as they are uttered, evoking *various narrative levels*. Genette does not imply that these categorical narrative buckets of *tense*, *mood* and *voice* are mutually exclusive, as they fully implicate one another through discourse, which is, after all, why all this falls under *narrative* and neither *story* nor *narration*. It is in this sense of simultaneity and discourse that we find the application of Genette to film sound so useful, as it begins to smash down hierarchical walls that separate sound and image and layers of signification and function. Sound tools like prelaps, reverbs, fader moves, drone tones, rhythmically timed sound effects and myriad others, which are often thought of as style by those dismissive of the discursive work that sound does, have a home in Genette's narrative constructions of tense, mood and voice, both on their own terms and in relationship to image.

The band starts playing, loudly, feathers flying in the air. Five layers of crowds cheering and screaming have been mixed: two drowned in reverb, two in less reverb, one reserved for the closer-up shots of crowds. They are placed fully around the room in stereo pairs that crisscross in the theater for maximum immersion and depth, forging an embodied experience for audiences.

Velvet Goldmine Image 10: Images from Concert
TimeCode: 08:32 & 08:43

Again, design and mixing looks towards Lastra, towards the Wagnerian "reintegration of sensory experience." The music has a strong 4/4 beat, marking time and story duration. Image cuts back to Brian, still offstage, with no reduction in levels, no accurate perspective shift, for the sensory immersion given by sound is primary. As Brian takes center stage and poses for the crowd, a boost in some of the crowd punctuates. It is we who are cheering. A teenager, whom we've seen highlighted in the running of the teens during the credit sequence, is shown in profile among the crowd. We recognize him, as the actor Christian Bale, later learning his name is Arthur. He notices a strange figure lurking to the side, but he turns back to the stage and starts shouting. There is no change in sound through this reveal and his individual screams of joy are barely perceptible in spite of our camera proximity, right at the edge of perception. Three levels, three narratives, three *voices*: the music of Brian is the event and narrative level one; the crowds experienced via sonic immersion embody fandom, our audience, young liberated revolutionaries whose cheers rise and fall above and below the music as narrative two; the singular teen who seems to know something, and yet his voice is buried by the other narratives serves as narrative three and will soon takeover. Genette's narrative levels of *voice* are well represented here. While intertwined throughout the entire 16-minute opening we analyze, here is a moment with three concurrent narrative levels, represented by mix levels of the sonic signifiers they produce.

Now in another transition, though in the same diegetic scene, the music begins to fade very slowly and some earlier tones, both musical score and SFX, that were largely unnoticed begin to grow, also very slowly. The SFX of an alarm, like a school alarm for recess or a fire drill, rings, fully mixed into the other sounds and at the edge of perception, but there. The band's music is now gone, eerie score dominates, a heartbeat pounds in the subwoofer, occasional crowd screams poke through, drowned in reverb, all on alternating closeups of Arthur and the

lurking man in the wings. Arthur is in a different head space now, and the sound reflects that as we have transitioned to a new liminal space, and duration lengthens. The discursive narrative elements are a function of what is chosen, their sonic linguistics, where they are placed in time and space, and how they are mixed, all in conversation with one another and image.

Velvet Goldmine Image 11: Concert Shooting Sequence
TimeCode: 09:18 - 09:53

In this very subjective, tense, anticipatory head space or *prolapse*, we see a close-up of a gun as it is cocked for shooting, a metallic sound with much reverb, and the heart beat grows, cut off by a sharp gun shot and a shift in the muted sounds of crowds into a kind of screaming, but still distant. The teen's heart? Brian's heart? Once again in an oneiric space, what metaphysical sense does the heart take on in this discourse? Brian is hit and goes flying backwards, landing on the stage with a thud, also at the edge of perception. These sounds at the edge of perception make us lean in, draw our attention through a near but incomplete omission. Very dark and eerie tones enter, the alarm sound is still there within the mix, more exposed now, evoking narrative counterpoint of continuing danger. People screaming are still drowned in reverb, so we are not permitted to reside in a realistic space, but rather still in the subjective head of the witnessing teen, and therefore his *voice*. Sound effects include sliding tones that don't seem to land, contributing to pure affect, scansion with signifiers with no knowable signification. An overhead shot over Brian, blood on torso, spins down towards him as crowd sounds fade away and tones, largely from SFX but also from musical score, dominate. There is no real harmonic tonic, no sense of musical key but rather lots of dissonance. We find ourselves transitioning again, as synthetized choral voices "ah" enter gently, over a cut to a chess board, and dominate the darker dissonant tones. "Ah" is human again, and divine. An iteration. The dissonance disappears. A chess piece is moved, and image cuts to teen Arthur standing outside, reading a headline of Brain Slade's shooting. The "ah" tones dominate more, fed by warm winds that sneak in. Image dissolves to a continuous wide shot of Arthur on the outdoor steps of a home, the vocal synth is the only sound remaining until piano enters elegantly, in the key that the synth "ah" established, and the voices fade. The music has been handed off from rock-

concert distorted guitar saturation to foreboding darkness to angelic divinity to a new song that begins with an acoustic piano intro as a warm wind blows, the camera tilting up to the sky.

From the moment the total immersive concert sound changed to something more subjective, through to this new song, we have been in and out of an everchanging relationship of *tempo*, an anachrony between screen time and story time, "temporal strands intertwining and braiding together" (Mroz, 2013, p. 36). Time has sped up and slowed down again and again. We've been overwhelmed and then given space to breath. We have been taken to loudness, heightened dissonance and tension then back to sublime harmony and resolution. We have been in the concert space, in Arthur's head, at a completely discontinuous and didactic chess board, and out in the streets a day later. Literary voice seems to have shifted fully to Arthur, for the moment, but the dance of levels of sound, levels of narrative, and the constant and elegant hand-off of *voice* as discursive narrative has been ever-changing in a weave, on which it is difficult to elaborate in text. But again, the rhizomatic experience forged by sound *narrative* has us surrender to the ebbs and flows of abstraction and embodiment, hermeneutics and experience, a dance. The Proustian narrative continues.

Tilting up to the sky along Arthur's gaze, the piano intro still playing in a steady rhythm, a young boy's face in the clouds whispers, "put out the torches, light the moon, light the stars." It feels like a dream, with soft sync, a whispered secret. In fact, this was done in ADR, intentionally not matching the more projected performance done for the camera by the young actor. *Voice* is implicated here, as is *mood*.

Velvet Goldmine Image 12: Arthur on Stoop, Boy in Cloud
TimeCode: 10:18 / 10:37

Then, over shots of boys and girls, faces up, in dappled sunlight, our narrator McTeer returns:

> For once there was an unknown land –
> full of strange flowers and subtle perfumes –
> a land of which it is joy of all joys to dream –
> a land where all things are perfect, and poisonous.

Each of these phrases is lined up with a musical phrase of the piano intro, allowing the musical phrase to begin before the voice does. This pacing and placement, this enunciation or utterance, this aspect of *tempo* that includes juxtaposed positioning, renders the lines as spoken lyrics, discursively unifying them with the song, imprinting the primacy of glam rock's, or pop music in general's, ability to evoke dreams and subvert, particularly for the young. Chion (2009) would note, as mentioned much earlier in his discussion on the rhetoric of nondiegetic songs, that as lyrics, these words provide an "acknowledgment of the hand of fate" (p. 404). Thus the positioning of lyrics through sound editing, universalizes the content. But are they nondiegetic, the words and the song – in the story world? The song continues with a cut to Brian Slade singing at a microphone.

Velvet Goldmine Image 13: Brian Singing 'Grand One'
TimeCode: 11:04

Thus, given a visual concretization, the song appears to be diegetic. Or is it? We are in a new space, back to Brian, but when? He's dead, or is he not? Where is he singing? He wears different clothes. What story-world could this be? Again, the diegesis question is irrelevant. And as for *tense*, have we gone forward or backward? This anachrony is, for the moment, undefined and again we are being asked to *surrender*, for the descriptive richness is unending, and the story logic has not yet been given form.

Going into the second verse of music, a television reporter is heard before the cut to a news exposition piece on "glam", with people interviewed on the streets, images of cross-dressers and same-sex kissing, all while the song of the music video continues. Brian's song dips where necessary to allow the announcer or an interview subject to be clearly heard. Are the music, and accompanying video, therefore part of this news piece? It sometimes seems that, and other times not. It also surges back to full volume when a large enough gap in other content

allows it. The news piece goes on to talk about Brian Slade, providing exposition, *story*, as the song continues, partly in the voice of the reporter, but mostly the interview subjects: a disgusted old man, an enthusiastic old woman, members of the band. While one of the band members is interviewed about bisexuality, we visually cut to Arthur, as a younger teenager in a movie theater, an old man putting his hand on Arthur's knee and smiling, and then cut back to the interview, with the interview and music continuing through this visual digression.

Velvet Goldmine Image 14: Interviews Cut To Arthur's Past
TimeCodes: 12:21 / 12:37 / 12:32

Is this a diegetic or nondiegetic insert? It can't be part of the video, Arthur has nothing to do with the band or the news piece, but his encounter is scored by the song, unbroken. Levels of narration, time periods, and voice dance about.

Time is clearly marked by the music, mirrored by the picture editing rhythms, giving us a sense that this is a music video. Yet we are given a different narrative, that it is news piece, at the same time with clues of both, fully integrated because of the mixing of the music around the interviews. Yet, the moment of Arthur in the movie theater is not of either. The discourse spins, rambles, supports multiple narrative levels, or *voices*, concurrently. We have also jumped back in time to before the shooting, and back to some unknown time of Arthur feeling challenged by a sexual advance, and then finally forward in time, to an image of a newspaper announcing that the Slade shooting was a hoax. We return to a close-up of Brian's face performing, presumably punctuating the end of the video song, although we honestly have no indication of where he is, and he states, "'I knew I should create a sensation,' gasped the rocker, and he went out."[104] With a freeze frame on Brian, the remaining music is only a dark tone that continues while a SFX of a 16mm projector enters. Then the frozen image of Brian melts away as if it is on a film strip that is stuck in a projector and burning. Unseen voices in appropriately-created room reverb discuss what they just saw. Anachrony, a large ellipse, is evident. The

[104] This line also plays with voice, as he is speaking of himself first in the first person and then in the third.

projector sound fades away as we flash-fade visually to Arthurs eyes in extreme CU, with no sense of diegetic space other than what is delivered sonically. *Tempo* and *duration* had previously been rhythmically structured by the beat of a song, but are now structured without an underlying tempo while embellished by the contrast of a rapid-fire pulse of a projector over a long, static shot. Layers of narrative, in tense, voice, and mood, are again at play simultaneously.

Velvet Goldmine Image 15: Arthur's Eyes
TimeCode: 13:38

We are in a drab office, we can hear the occasional typewriter and telephone ring. It's established to be a newsroom sonically first and Arthur (Christian Bale) is a reporter. We have proper camera-based perspectives on both voices and off-screen background SFX and yet we hear foleys of pocket change each time Arthur's boss' hand goes in and out of his pocket, and Arthur's glasses nervously clicking in his hands. At this point it seems we have finally landed in a story of which we can make sense. Arthur appears older, SFX and the mixing of them are generally realistic to the space, although counterpoint is exemplified by the foley sounds of change in the pocket, change that does not really exist in the actor's pocket, and eyeglass clicking, which made little to no sound on set.

In our *Koyaanisqatsi* design we were concerned with targeted meaning and demonstrated that SFX impact interpretation, although not always as intended, not quite so predictably. Both these foley sounds of change and eye glasses are heightened as materializing sound indices and provide narrative or thematic information as well. Is the jingling change in the boss' pocket representative of his obsession with money? Is he cheap? Or is it Arthur's perspective that he is ruled by his job, by the money he has to have to survive, which calls to him as he looks at his boss? Are the eyeglasses a symbol of Arthur's reluctance to look, to see something he is avoiding? Or of being seen? Or simply nervousness? Any or all of these are

possible outcomes of counterpoint, or none, or countless others. We are reminded of Baracco (2017) addressing the Ricœurian hermeneutic position that is in opposition to any pre-determined definitive meaning embedded in text, insisting that the process of interpretation "remains always open-ended and in constant search of meanings" (p. 91).

Moreover, a few moments later in the next scene, Arthur gives voice to his state of mind via narration, which retrospectively alters or rather solidifies the meaning of the pocket change and eyeglasses. Thus while they are enunciated in one moment, clarity to meaning occurs in another, as experienced by the audience. Much beyond Chion's counterpoint, these are unique kinds of *iteration*, metaphoric ones, when a SFX in one place echoes dialogue in another. Perhaps this exemplifies Lastra's "compulsive linking" of sensory and representation in a way that effaces representation, "to restore reality and to offer it to us in its fullness" (Lastra, 2008, p. 135).

Arthur is assigned a story to research, "10-years later, whatever happened to Brian Slade." Steady, dark guitar tones enter as Arthur's discomfort with the assignment grows – they are dissonant, become distorted and land on a minor and slow-paced musical theme with a city image and text on screen, "1984 New York City", as Arthur walks along the street.

Velvet Goldmine Image 16: Arthur Walks 1984, NYC
TimeCode: 15:16

The setting, and the date, are meant to feel Orwellian, reflecting the rhetorical stance of the filmmakers, pulling the listener-viewer into the hermeneutic circle and "dynamic passivity" to access various narrative levels and abstractions.[105] Sound provides footsteps of crowds but no voices, a Chion c/omission perhaps, but more of a *voice* that registers

[105] As was discussed by me with director Todd Haynes in the mix.

oppression.[106] A voice is heard imploring people to join an organization to support the president, filtered and delayed to sound as if emitted from a public address-type speaker, therefore mixed indexically and an act of filmic discourse. Sirens in the distance, so drowned in reverb they are more of a feeling than a discrete presence, suggest danger, or police. A cold arctic wind cuts through as well, though Arthur is not dressed for winter. There are unrecognizable screeches in reverb, like a metro train, followed by a siren that is now more discernible. The scene has been rendered cold, authoritarian, by sounds and mixing techniques that have established *mood*. All of these play through and under a voice-over by Arthur in which he wrestles with being dragged back in time to look at things he gave up for money and security; hence the echo, the *iteration*, back to the heightened pocket change and the clicking glasses from the scene before. Arthur by now is the *first narrative voice* as depicted through voice-over and the tracking shot of his walking. The cold sounds along with image details form the *second narrative layer*. Arthur approaches an urban plaza and the sounds of what was the public address system transitions to pop music and a person shouting into a microphone. This is eventually visualized, shown to be a giant screen, a rock concert, broadcast live via satellite to screens everywhere. The performer thanks corporate sponsors and the presidential committee. The sound from that screen has been rendered in two layers – that of a live concert and that of a speaker in a plaza, in order to construct a sense of reality, to parse out the right information, both *mood* and *tense* at play. Arthur walks away, ignoring it. He steps down onto the stairs to the subway and the subwoofer booms like a timpani as someone passes him and they look at each other, familiar. The weight of the moment is felt from the subwoofer. Not told, but shown through sound, via *mood*. Musical score has returned, in a slow rhythm with slide guitar.

Velvet Goldmine Image 17: Subway Sequence
TimeCode: 16:14 / 16:30 / 16:36

Subway screeches fill the upper frequencies, drowned in reverb, as Arthur continues down the steps and a child's voice is heard in voice-over, another poem:

[106] We remind the reader of our proposed expansion of Chion's rhetorical figures to be between the seen and heard, rather than between dialogue and everything else.

Yesterday upon the stair
I saw a man that wasn't there.

We cut to inside the subway, with the characteristic flash of light as if the electricity is going on and off, but with a lightning strike SFX on the flash. The child continues,

He wasn't there again today,
How I'd wish he'll go away.

It is revealed that a boy on the subway reads aloud from a book. He looks up, presumably at his mother. Our sense of *voice* has again been manipulated, with a voiced poem beginning on the steps and continuing without a pause over a visual ellipsis – *tense, tempo* – a time cut forward, to the interior of the subway, feeling completely as yet another extra-diegetic narrator and mixed to sound as one, but revealed to be part of the "diegesis", the story, a boy on a train whom Arthur notices. And yet, the boy never sounds in situ, for he is simultaneously both – operating on multiple narrative levels reinforced by the mixing of the voice, where diegesis is poorly defined and irrelevant. As for the text of the poem, it operates on yet another narrative level, opening more questions than answering. What is it to see a man who isn't there and wish he would disappear? Why do the words come from a child? This is an *iterative* theme that reappears, not the first time a child's presence reaches for abstractions, hinting at the rhetorical stance of the filmmakers, with meaning that is difficult at this juncture to determine. We are again, as with many moments since the beginning of the film, asked to *surrender*, to forfeit meaning-making for experience, where the full breadth of discursive vectors is provided by sound in its choice, placement and mixing.

We are still on the boy, as another blinking of subway lights is sonified with the sound of a lightning bolt SFX. There are still more momentary flashes and then image cuts to Arthur sitting in the subway, yet all but one of these flashes are silent. It is illogical to have subway light flashes sonified with thunder, but synchresis provides a certain logic with repetition. Yet, the logic within the illogical has been broken with flashes that don't sound. This plays as metaphor, as scansion, as *mood*, and as the *voice* that tells of an impending storm for Arthur that began earlier on his extreme close-up in the office, highlighted by the sound of the clicking eyeglasses. It is therefore *iterative* and working in multiple sound vectors; foleys for props and

impossible lightning strikes. The long transition from the street to the train once again feels oneiric, liminal, and continues with new, gentle, slow sounds of music on a shot of Arthur on the train, thinking, and looking out the subway window. His voice-over enters, and with Arthur's returned state-of-mind homo-diegetic voice, underscored by the gentle train, fading lightning strike sounds, and music picking up pace into arpeggiated triplets of a song, not actually score, we dissolve visually to kids running in the street, back in the 1970s, in the colors of the 1970s, and cut to Arthur in a classroom.

Velvet Goldmine Image 18: Young Arthur in Classroom
TimeCode: 17:30

The tempo for this transition, driven by music and gradual crossovers of ambience and other sound, carries us through a liminal "in-between-ness," (Siegworth and Gregg). We are taken for a ride, with Arthur on a train, back to his past, and we engage in Walsh's "dynamic passivity" between "losing ourselves" (Deleuze) and "interimplication between body and world" (Merleau-Ponty)," willingly going for the ride that finally lands, for a while anyway, in a time and place. We are now 16 minutes into the film.

The first 16 minutes the film has seen 1854 Dublin, approximately 1863 Dublin, "a hundred years later" in England so perhaps 1963, and 1974 London at the time of the shooting hoax. Then earlier in that rough period, then later, then earlier again, then later again, then a jump to 1984 New York, and a transition back to 70s London, to Arthur's teen years as he is discovering Brian Slade and his own sexuality. Additionally, moments of timeless intervention, rupture, the rare nondiegetic image, include the chessboard (10:01) and a bright pink title card providing a Norman O. Brown quote, "meaning is not in things but in between them," (13:06) referencing meaning-making and abstractions. Repeating Genette, Proust is,

> encumbered with digressions of all kinds, retrospections, anticipations, iterative and descriptive parentheses, didactic interventions by the narrator, etc., […] able to give […] a fully paradigmatic importance (1983, p. 111).

Overwhelmingly, in *Velvet Goldmine*, these characteristics of Proustian prose or rhetorical devices are embedded in the sonic discourse. Moreover, Genette further explains that Proust is almost all "scenes" of a narrative form that is most rich in information and therefore mimetic, but on the other hand, the narrator's presence is constant and intense, especially as the producer of the metaphors. So, the narrator is simultaneously at the extreme of showing and telling. The creators of *Velvet Goldmine* as narrators, as well as the shifting protagonists as narrators, often with voice-over, are in a similar place, though the degrees of showing vs. telling shift dramatically in a connected dance, an ever-unraveling discourse, a paradigmatic kaleidoscope. Genette's *narrative*, broken down into *tense*, *mood* and *voice* and further broken down as earlier discussed, serve as a meaningful structure to unravel the sonic discourse in *Velvet Goldmine* and its relation to image and story.

We wish to briefly address one of the final sequences of the film, for it reflects deeply on this discussion. It may be a fantasy, or a memory, this is unclear at the point of enunciation although later implied to have been real. One hour and 37 minutes into the film, Young Arthur is shown at a "death of glam" concert where Curt Wild is performing, so we are roughly in 1974, the exact time is never stated. After, backstage in a dressing room, Arthur walks in and Curt Wild is there. Arthur, in voice-over that comes from the 1984 Arthur, says "It's only now, looking back, I see how you patched through my walls, and entered my life…in waves." On a shot of Curt, looking at Arthur and not speaking, we hear Curt's voice-over, "come closer". Image cuts back to young Arthur, walking towards the camera, towards Curt, but no longer in the backstage dressing room space, but rather an undefined liminal space, like the childhood Jack Fairy who was to put blood on his lips in defiant self-acceptance. In voice-over, Curt continues, "don't be frightened...what's your name…what's your favorite song, movie?...don't be nervous…are you high?" In sync, Arthur answers, "I'm on a button" as he arrives on a roof, Curt waiting in the distance. Then back to voice-over, sonically to 1984 Arthur, "He was waiting for me. I followed his signals, slipped away, and now suddenly…"

Velvet Goldmine Image 19: Images from Roof Scene
TimeCodes: 1:41:04 / 1:41:41 / 1:42:12

Time cuts forward a beat, a beer opening and spraying punctuates the time cut. "Hey," Curt says (in situ, in sync) looking at the sky at a shooting star while we hear the same sonic woosh from the beginning of the film. Curt seduces him, removes Arthur's shirt. A spaceship flies over, dropping glitter, in which they bask together.

Velvet Goldmine Image 20: Additional Images From Roof Scene
TimeCodes: 1:42:42 / 1:42:45

A piano intro begins, and Curt says, lips moving but mixed dreamy like voice-over, "make a wish…" The camera pulls back, and in the distance they make love on the roof, and Curt's disembodied voice continues "and see yourself, on stage, inside out, a tangle of garlands in your hair, of course you are pleasantly surprised…" Then Arthur's 1984 voice narrates, "softly, he said…" Back to Curt's voice, "I will mangle your mind." This final line of the scene utilizes delay[107] and all speakers in the room so that we hear Curt's overlapping voice from five different points in the room, not quite simultaneously, such that each is registered separately, but not as distinct repetition.[108] The effect of this moment represents an iteration that is at once *singulative narrative* (the moment), *repeating narrative* (the multiple iterations we perceive), and *iterative narrative* (capturing many instances thematically in one utterance). The intended

[107] Delay is a process that delays a sound signal in time, frequently added to the original position as well.
[108] Typically, a delayed repeat of 75ms or more is registered as a separate utterance, while less than 75ms is perceived as a fatter sound. This varies between individuals and frequencies. I made sure all instances were perceptible made more possible by immersion, by feeding each delayed signal to a different speaker give not only time delay but a change in directionality.

impact is a different kind of immersion than one would feel if all speakers were used simultaneously, without the delays. More than just an immersion, it is a penetration that calls for *surrender*, a mangling of the mind. In *voice*, Genette reminds us that the narrating includes its "two protagonists: the narrator and his audience, real or implied." The voice is Curt's, the receiver is not only Arthur, but us all. Similarly, in our look at pragmatics we implicated the audience in the Gricean talk-exchange and cooperative principle. Throughout Velvet Goldmine, the filmmakers offer moments where they directly speak to us, sometimes from outside the story world like the Norman O. Brown quote, sometimes from voice-over that seems at once diegetic and nondiegetic, and here, more subtly via a mixing choice to penetrate with what is almost a rupture of the fourth wall, received that way due to the iterative nature (*tempo*) of the direct address already embedded in the film. Genette's rhetorical tools articulate well here with pragmatics and with the power that immersion and other sound techniques offer to our concept of *surrender*, and as Curt does for Arthur, the sound design *seduces surrender*.

From here, the music becomes the music video "Tumbling Down," (Harley, 1998) with Brian Slade in an outrageously colorful, flamboyant costume on a baroque set that is entirely Wagnerian, with a cascading staircase, a chandelier to climb on, grand theatrics, flying feathers and shimmering glitter.

Velvet Goldmine Image 21: Brian in Wagnerian Performance
TimeCode: 1:45:23

Lastra, in "Film and the Wagnerian Aspiration: Thoughts on Sound Design and the History of the Senses" (2008), notes that *Apocalypse Now*

> uses form to point to what is not entirely contained by that term. In short, it argues by way of sensation…that affect, emotion, and sensation have a history and, beyond that, a meaning as well (p. 137).

By subverting form in richly complex narrative discourse and especially through sound design,

Velvet Goldmine also finds its deepest meaning via sensation, "in between" things, in the liminality it creates that oftentimes pays little mind to story and invests everything in Genette's narrative, in modalities of representation and enunciation, in the *surrender* to the swings between abstraction from self via Deleuze's multisensorial and the interimplication between body and world, including story-world, of Merleau-Ponty's phenomenology.

We see, through this case study, that Genette's narrative discourse applies well to sound design. It exposes myriad roles of sound design in a much more complex discourse, as part of a full partnership with image in a broader definition of narrative. As a Proustian film that relies on sound deeply, *Velvet Goldmine* serves as an interesting case study for examining Genette simply due to the complex nature of the sound-image discourse. We had sought to explore new ways of thinking of film-sound rhetoric after interrogating existing theories and exposing, through our own research-led artistic output and content analysis, both desired outcomes from sound effects and the complexities of narrative and thematic communication brought about by sound. The expansive rhetorical structures found in Genette proved promising as we extrapolated them from literary analysis to film sound. Having now applied Genette to *Velvet Goldmine*, and work that is part of this author's repertoire as mixer, we find deep resonance and, most importantly, an ease of vocabulary available to describe how sound design functions, even in a very complex discursively constructed film.

CHAPTER VI. CONCLUSIONS & FUTURE WORK

This story began around a protagonist, this researcher, a film sound mixer of many years and professor in two academic film cultures – at the University of Pennsylvania in the United States followed by Lusófona University in Portugal. My artistic practice as a mixer has directly informed my teaching and the courses I have developed. Those theories that are most often taught, as well as others that have jumped off the page and resonated with me during my personal theoretical explorations, helped to frame my courses over the years. But time and time again, the sound artist in me challenged the theoretician, as did many of my students. It seemed that established theories that developed a vocabulary for discussing film sound do a rather surface job at describing, addressing, or analyzing many of the ways in which film sound works. While many authors have written of film sound in less accessible but more philosophical terms, and those discourses are often very rewarding – often reaching towards the ethereal, the experiential, the psychological, the embodied – other books and articles that have established an analytical vocabulary that is easily accessible for teaching merely scratch the surface and remain overwhelmingly attached to image or to the screen, grounded in a cultural hegemony that favors image and diminishes sound, even while they attempt to overcome it. The tension between theory and practice is what led us on an exploratory artistic research design with hybrid methods that might steer us towards our fundamental goal: to liberate film sound from image paradigms, and bring it back into discourse on equal grounds. The silver lining of confronting limiting established views is that, as discussed in the introduction, the field is rather young, with sound theory representing only a fraction of theoretical work in film.

6.1 THE EXPLORATORY RESEARCH PATH

Our exploratory path began in Chapter I by addressing our first of four inquiries: are current ubiquitous structural techniques of discussing film sound adequate? This initially involved discussing existing sound theories, especially Chion's diegesis and rhetorical structures; the former because it is ubiquitously taught and understood as a way to categorize sound, yet quite incomplete, and the latter because, although Chion's figures of rhetoric remain tied to the screen via dialogue, he grants permission to look at sound and its relationship to image as a rhetorical one, a very compelling idea. Here we brought into discourse the problem

of diegesis as a concept, other schools of film theory in the context of sound, for instance semiotics and rhetoric, as well as fields beyond film theory that have often been brought to it, like phenomenology and abstraction, hermeneutics, pragmatics and utterance, embodiment and affect, and more. Along the way, these shed light on the surface nature of existing film-theory vocabulary by offering tools to elaborate on sound function through interdisciplinarity and with greater depth. To further demonstrate shortcomings in sound theory, based on this literature review we created an artistic object from which limitations in theory could emerge through analysis. Not only did this offer criticism of existing theories, but it gave us a forum for expressing, via interdisciplinarity, a richer discourse, further justifying our explorations. Our artistic research and case study brought these dynamics to light. In the aggregate, the answer to our first inquiry became clear; not only are diegesis and rhetoric, particularly as expressed by Chion, not adequate but there are many other avenues to pursue. It is also worth noting that it is here where we established our use of the term *surrender* – not to suggest Deleuze's disappearance of self, but to enunciate surrendering control of cognition to a flow of anonymous intensities and cognitive awareness.

In reading literature on sound theory, both popular and academic, an underlying discomfort was felt as we came upon many assumptions of sound design impacts on audiences, assumptions that could be rationalized through literary discourse, but that do not seem to have been defined and tested. This formed our second area of inquiry: are these assumptions of sound function defendable? We felt the need to establish a mechanism for testing sound functions and to use such a test to open dialogue about cognitive and experiential mechanisms related to sound, including the generation of meaning, the creation of thematic understanding, and the overall experience of listening-viewing. Chapters II and III were dedicated to this search. In Chapter II, we began with an understanding that sound design is now ubiquitous in film creation and, as such, must offer something to those involved along the lines of a technology. We looked closely at theories of technology and innovation, particularly Winston and Rogers, to establish that both engineered sound solutions for film and sound design are technologies. However, for sound design, what remained was defining the advantages that it offers the social sphere, which encompass the assumptions in the literature that gave us discomfort. By looping back to our research-led artistic output and the discussions that ensued, we ultimately proposed that sound design provides creators and audiences, or the social sphere, the following seven advantages:

1. Increased attention and focus by audiences;

2. Support for narrative understanding and/or thematic interpretation;
3. Physical and temporal continuity – hiding of the apparatus;
4. Thematic continuity;
5. Imagined sound – engagement of mechanisms of phantom sound;
6. Experiential engagement or impression;
7. Emotional engagement.

Chapter III was dedicated to content analysis that could test for correlation between sound effects and the proposed advantages, using both quantitative and qualitative tools. Overall, we determined that assumptions are defendable, but with nuance and the need for modified study design in some cases. We were not able to show correlation for #1, attention and focus, and we propose a new study design due to issues with environmental scaffolding that might overcome the fact that attention was already very high in the version without sound design, giving skewed results. For #2, support for narrative understanding and/or thematic interpretation, we found that for some shots or sequences with sound design, correlation with intended meaning-making was demonstrated. For others it was not. In discussing those that did not correlate, we found in particular that shifting away from targeted meaning-making was accompanied by indications of greater experience, or impact, moments perhaps of *surrender* in lieu of interpretation. Evidence of such surrender was found in questions on memory and thematic continuity. Advantages #3 and #4 had strong correlation between sound effects and continuity, including both *physical and temporal continuity* and *thematic continuity*, which we defined as connectivity between disparate shots or sequences in terms of interpretive themes. #5, imagined sound, also known as phantom sound, showed a strong correlation with sound effects, demonstrating that the secondary auditory cortex is active and better able to imagine sound given a visual stimulus in the context of expecting sound effects. Hypothesized advantage #6, experiential engagement or impression, was examined using questions about the memorability of images. Results for this were mixed, but certainly changes in memorability did occur with sound effects. Put in conversation with advantage #2 dealing with impression, this yielded an interesting analytical process and is suggestive of the dance between embodied experience and cognitive processing discussed in earlier theories, paralleling our concept of *surrender*. This certainly warrants further investigation, perhaps on a single shorter segment rather than a long piece. Finally, hypothesized advantage #7, looking at emotional engagement, did exhibit correlation between sound design and emotional engagement that serves the social

sphere, even if defining specific causes of this shift is rather complex. Overall sound-design does meet the requirements of an innovative technology, which further implies that: (i) it is diffusible; (ii) it evolves; and (iii) the likelihood of desired outcomes of filmmakers being met is increased with sound design.

Having put to rest that sound effects do, in fact, support many of the hypothesized advantages and certainly have some impact on those that showed non-significant statistical correlation, in Chapter IV we returned to address our third area of inquiry that asks if there is a place outside of film theory where a more holistic theoretical framework might offer a stronger analytical discourse. Due in part to Chion's opening of the idea of rhetoric, and to Genette's use of Proust as a literary example, we turned to Genette's *Narrative Discourse* as a potential model, drawn to it by his idea that anything subject to textual analysis in literature is narrative, and all sound is part of the text of film, thereby broadening the scope of what is traditionally defined as narrative in film to include all sound. We therefore examined Genette and used our research-led artistic output, as well as other films, as a case study to place various sound design strategies and outcomes into Genette's narrative categories of *tense, mood,* and *voice*, and their sub-categories, finding rich and flexible possibilities in this model. We found that Genette's rhetorical structures adapted remarkably well to our concerns around sound design both in terms of text and multiple vectors of function.

We then moved to Chapter V where we tested Genette against a complex sound design mixed by this author, *Velvet Goldmine* (1998). This case study demonstrated that the discursive role within narrative as defined by Genette is a fruitful way to look at sound for film, both in terms of the discursive role sound plays against image, and against other sound. Moreover, the rhetorical approach focuses on the authorial intent and inference, leaving open a vast space for theories of reception and, in our case study, those articulated very well with rhetoric and further supported the need to broaden discourse on sound design to reconcile theory and practice.

6.2 Revisiting Liberation

Having addressed our four area of inquiry, we now circle back to our overall aspirational goal as expressed in the introduction. Can we theoretically liberate film sound - dialogue, sound effects, music and mixing – from strict image paradigms, and then bring it back into conversation with image on equal ground? For us, the flexibility offered in Genette's narrative discourse gives us an important tool that does liberate film sound from image

paradigms precisely because it is based on textual analysis with ample space for nuance and allows, perhaps for the first time, for the discursive and performative role of mixing to find a home in both affect and meaning-making. We are reminded that narrative, for Genette, is concerned with the analysis of texts and that all sound in film is textual, it is all part of the universe of cinematic signifiers. In this construction, two fundamental ruptures from screen-based sound theory hold true: diegesis is not a determining factor of the role of sound, and rhetorical framing need not be structured around a hegemony of dialogue, as it is for Chion. Using narrative structures like *tense, mood, and voice* we find that sound operates around the triangle that encloses Booth's *rhetorical stance*, which positions available arguments, audience and voice[109] not in opposition to one another but in discourse with one another, the rhetorical stance emerging from that discourse. But with Genette we can delve deeper. The arguments available to us are found in the whole text, all of narrative, including *tense, mood and voice*. The audience is implied in Genette's narration as is the implied character of the enunciator. Genette's voice is not Booth's, but rather it is in relation to the subject of enunciating, which in film is most frequently a character, or an object that makes sound. This is perhaps one of the critical differences between sound and image. If we see a clock on screen, the enunciator of that image is the filmmaking apparatus that created the representation of a clock on screen. If the clock chimes, however, the enunciator of the sound is *thought* to be the clock. It is here where sound defies the typical representational characteristics of image. For image, voice resides outside the narrative. For sound, it emerges from within and this is equally true for a sound whose diegetic placement is challenging. We question where does it come from, assuming it is from within, and if it is music score, it is most easily associated with *mood*, defining the mood of a scene, part of Genette's narrative signification, a "modality of narrative representation." The role sound plays in temporal anachrony and continuity, duration in terms of the relation between story time and felt time, and frequency in terms of iteration, is ubiquitous across most if not all films. It is only through equal weight to sound and image that these emerge in the sound film. Plus, as sound can and frequently does operate on multiple *voice* narrative levels concurrently, as we have shown in our content analysis of thematic/narrative interpretation and our discussion of Genette and *Velvet Goldmine*, variations in *tense* play strongly between sound and sound as well, not only sound and image.

[109] For Booth, we are reminded that voice is the implied character of the enunciator.

But what of Chion? His foundational constructs of diegesis and rhetoric are certainly not to be discarded, nor do they represent even close to his full body of work. These and many sound theories of his and of others that have emerged over the last few decades frame part of an evolution, not a revolution, in how we look at sound for film. Time and time again they have broken ground, and in those fissures have planted seeds that ultimately challenge image-dominated film theory. We conclude that this canon of work serves as a minor literature, utilizing the language of image and the positioning of sound within it to ultimately subvert that positioning. We find further evidence of image-based dominance in many of the contemporary thinkers whose great work fueled expanded ideas in our research. For instance, one would imagine that theories on temporality in film might offer strong positions on sound, and yet we find oversights, or shall we call them oversounds, as in Mroz's extensive and insightful *Temporality and Film Analysis* where, in its first chapter entitled "Time, In Theory", Mroz discusses Sobchack's work in *Carnal Thoughts* which addresses sensory in cinema and embodiment, in which Sobchack asserts that we bring not only sight but all our senses to the cinema: touch, smell and taste.[110] Sound is conspicuously absent from Mroz's link made here between embodiment and temporality (Mroz, 2013, p. 27). In the context of such omissions even in the contemporary canon, Chion's subversive minor literature work 30 years earlier shines formidably. That subversion of film theory, however, may still be in its infancy. As Deleuze and Guattari (1986) suggest,

> The problem of a minor literature but also a problem for all of us: how to tear a minor literature away from its own language, allowing it to challenge the language and making it follow a sober revolutionary path? … Kafka answers: steal the baby from its crib, walk the tight rope (p. 19).

Chion and others have formalistically molded film sound into Kafka's baby as a product of her mother's DNA, namely image-based film theory, then stolen her, and only just begun to walk the tight rope towards revolutionary, or evolutionary independence. Our work today is to continue along that rope and ultimately reintegrate a mature child, a young adult, into an evolved family structure in which reunion implies both independence and interdependence. For us, Genette's narrative discourse is one such point along the rope and an attempt at reintegration. Our research is relevant in the field precisely because it represents a

[110] Mroz cites (Sobchack, 2004, p. 59)

break from the screen, illustrated with case study and content analysis, and looks to new theoretical underpinnings that have not been, or rarely are, put to the test of sound for film. Our content analysis in particular, and further brought forth in our application of Genette to *Velvet Goldmine* in the case study, bridges a gap between authorial construction and reception. This bridge is quite relevant because it is here where practitioners highlight the assertion that the theoretical discourse, particularly with respect to diegesis as discussed, often misses on what sound actually does. For us, the bridge is traversed in the articulation that begins with a rhetoric encompassing narrated instance and rhetorical stance, utterance, inference, and the cooperative principal of pragmatics, and carries us through pre-cognition, dynamic passivity, sensory derangement, interimplication and *surrender*. Such an interdisciplinary approach seems to be where sound theory for film and other applications may thrive, and from which we may springboard new ideas, innovative technologies, and innovative practices.

6.3 SEDUCING SURRENDER

Along the tight trope we have entered our space of seduced *surrender*, and theoretical questions arise in this space, highlighting the relevance of our work to other researchers. Here, in *surrender*, we are at once cognitive and experiential, we are engaged in the hermeneutic circle, tapping into pre-consciousness, reading signs consciously (or not), and dancing between a pure, out-of-self abstraction and interimplication. It is no wonder, therefore, that multiple interpretations in the film world are always possible, as they are in the real world in terms of individual, political and social narrative. In Ricœur's symbolic interpretation, according to Baracco (2017),

> Different interpretations are not occasional expressions of an individual freedom, but are considered constitutive and necessary in order to understand being. Through these interpretations, human existence reveals itself as a purely hermeneutic experience, which in a deep relationship with a symbolic world must answer to a continuous and inescapable question of meaning (p. 111).

By this, within the symbolic world of film, the question of meaning is unanswerable but rather lies in a dialectic among and between various interpretations. How that reconciles with integrating the cognitive unconscious dimension into cognitive film studies (Coëgnarts, 2015) remains unclear. For film audiences, is pre-conscious cognition societal or individual? Is it determinant in interpretation or merely a point of origin into the hermeneutic circle? When

sound is executed in its most interesting ways, these questions come to light precisely because of the abstractions, or access to them, that sound brings, forging multiple interpretations via the rhetorical devices discussed and the *surrender* it often brings. We return to Proust's (2004) autobiographical fiction and his character Marcel's phenomenological moment, his surrender, seduced by words on a page, that

> filled me with […] a joy I felt I was experiencing in a deeper, vaster, more unified region of myself, from which all obstacles and partitions seemed to have been removed (p. 96).

To us, this joy defines the full force of the *surrender*, into *the flow of emotion and cognitive perception* without boundaries. Sound design, especially when Proustian, takes us to this liberated space where multiple experiences and interpretations are possible.

To conjure Proust, Wagner and Oscar Wilde along this artistic research and exploratory journey has not only been an exciting engagement but has underlined the very subversion that we, as practitioners, understand implicitly and intuitively to be a primary role of sound in film. In part, this subversion is rooted in the departure from, perhaps deterritorialization of, the frame by the physics and temporality of the sound wave and the immersion of sensation. Sound design, in its power as an element of discourse, ruptures the two-dimensional visual frame of the screen as it is rendered in three-dimensional sound pressure waves that travel around and through us, and are only marginally different from the sounds they are meant to represent, whether concrete or abstract. These sound waves activate the physical, psychological, emotional and cognitive centers of the listener-viewer who sits, filmmakers hope, alone in the dark and glued to the screen. Those same filmmakers use sound pressure waves to fundamentally curate *the flow of emotion and cognitive perception,* by penetrating defenses and "mangling minds" as Curt Wild does to us all in *Velvet Goldmine* (Haynes, 1998), coming at us from all directions iteratively. This sound is, in the word of Kittler, the *Real* (1999). It is a rupture of the diegetic territory of the story, pulling narrative out, off the screen and entering it into discourse between the real and the representational – subverting image with real physical stuff on its own terms. Lastra's "founding gesture of sound design," as mentioned earlier, places sensory experience independently from representation and rejoins them, effacing representation and thereby restoring reality. This is Kafka's baby, more than a representation of her parents, but her own, very real DNA which is, like any human, a total work of art.

Lastra (2008) writes, of the Vietnam War theater depicted in *Apocalypse Now*, that it is an example of gesamtkunstwerk[111] (p. 128), a total sensory immersion for the participants; those participants in the drama itself and those in the movie theater. He echoes Genette in the sense that sound design creates layers of voice in a way that opens a path of vulnerability for an audience in their "relation with the subject […] of enunciation" (Genette, 1983, p. 31). If we are to look at film, therefore, through the prism of Wagnerian aesthetics as elaborated by Lastra and exemplified in our application of Genette to *Velvet Goldmine*, what becomes clear is that sound design, or even sound in general, does not enter into an image discourse but is fundamentally part of the total work of art and, in the modern theater or home screening area with multiple speakers, is often the more powerful component in terms of sensorial and narrative immersion. As our analysis shows, ultimately the discourse within the audiovisual text and its reception for an audience are far more complex than what can simply be regarded as a visual text supported and enhanced by sound. This begs a response to an inherent assumption embedded in many questions often posed by film theorists, directors and others, including our colleagues, when we discuss our work. It goes something like this.

> So what you're saying is that sound can alter the interpretive process or the emotional experience of the audience given what is happening on screen – sound design can change the meaning and impact of the image?

Our answer to this question is two-fold. One the one hand, our content analysis strongly suggests that sound design can, and often does, alter or contribute to interpretation and emotion. However, and more fundamentally, our linguistic and philosophical discussions lead us to answer the question with another, more probing question. Why is it that sound is positioned in the original question as the catalyst that alters image? *Why is it not that image alters the meaning-making and emotional signaling of sound?* This is, perhaps, the ultimate question to have arisen from our exploratory journey. Our visually-dominated culture ironically (or cynically) blinds us to the full richness of our interaction with sound in the audiovisual experience. Perhaps it is due to this blindness that sound can play a subversive role, and perhaps it is due time that the sound film is theoretically reapproached from this positioning, with sound on *at least* equal ground as image.

[111] A total work of art, or "an art work produced by a synthesis of various art forms (Merriam-Webster, n.d.-a), and a term used by Wagner in at least two essays on aesthetics.

Therein lies the greatest strength of the audio component of this audiovisual medium. At its most creative, sound design fuels a Deleuzian map, which can break, augment, or rupture the diegetic construct, engaging in a Proustian dance of narrative discourse that exploits our tenuous sense of diegetic, physical and metaphysical boundaries. It indeed makes all filmgoing interactive, an argument we have earlier made. Carbone (2014) connects immersion and interactivity via Merleau-Ponty's "Eye and Mind" in a discourse on the electronic image, which we wish to share here but further articulate with sound. The "character" of the electronic image, Carbone says, lies in immersion, not representation, "and in the consequent 'interactive' rather than 'contemplative performance' that it demands" (pp. 231-232). We have already argued that sound, too, is based on immersion and not representational – due to the real physics, the technology behind it, its indexical nature, plus our entire discourse around *surrender*. Carbone further explains that "'rather than seeing [the painting], I see *according to*, or with it',[112] while the 'interactive performance' that it requires seems to be evoked when it is affirmed that 'vision is *the meeting'*, as at a crossroad, of all the aspects of Being".[113] Thus to Carbone the sensory apparatus defines the meeting point of interaction. We earlier explained that all across the value chain of filmmaking, there is an interaction with listener-viewers from the perspective of product creation. But here we see that, especially with the real or non-representational aspect of sound design in an experience, there is quite a literal in-the-moment interaction as well which is tangible, as we have shown, in the hermeneutics of interpretation and in the surrender. Furthermore, we find a theoretical parallel to sound's Deleuzian map, a rhizomatic one, in the myriad theories of communication and theories of being that the sound film permits, as we have seen throughout this discourse.

To conclude we turn finally to Lastra (2008) who seems to elicit the intent of sound design to heal a world so lost in representation, so detached from the real, that it seeks to provision the core essence of phenomenology. Lastra asserts that, in its post-Dolby development,

> [sound design] aspired, in effect, to the status of total artwork, which would redeem a world no longer capable of being experienced in itself. Though born of, and in part

[112] Carbone cites (Merleau-Ponty, 1964, p. 126), emphasis Carbone's.
[113] Carbone cite (Merleau-Ponty, 1964, p. 147), emphasis Carbone's.

> complicit with, a crisis in perception, it harbored within itself the utopian desire to restore to the senses the ability to experience (p. 135).

6.4 Future Work

Our exploratory work has yielded some questions or concerns that warrant further work both as an extension of this research and in other areas. Sound bears a strong relationship to mediated representation, diegesis, rhetoric, embodiment, and the suspension of disbelief, which are all important in terms of the current state of media studies, notably in the immersive media space.

Content Analysis:

As earlier stated, the artistic research piece which we used in content analysis only truly establishes correlation between sound design and desired outcomes for *this* sound design and *this* film excerpt as experienced by our *specific* samples, which are less random than we would have hoped for. Further analysis with different films is warranted, or even a film created specifically for study. For instance, adding sound effects to a film that originally has only dialogue and musical score (less dominant music than Phillip Glass), may yield a new set of data around audience responses to sound effects due to the narrative grounding of dialogue. In addition, we may wish to create short sound designs that specifically test for only one desired outcome at a time, with much shorter excerpts, so that audience memory becomes a smaller variable in data collection. We would also like to revisit the population samples with our original, pre-Covid design, namely a more randomized audience acquired in a shopping mall, experiencing the two versions under the same controlled conditions in a movie theater, and with a larger cohort. All these re-approaches to artistic research and content analysis should provide data that either more robustly supports our conclusions or challenges them, and would therefore be very worthwhile moving forward.

Suspension of Disbelief in Immersive Media:

Jumping off from our conclusions that sound is of the *real* and often creates experience above representation, this author wishes to investigate one specific implication of this assertion. The suspension of disbelief is generally thought of as the cognitive process by which we

consciously, on some level, agree to believe a film or any story as true, willingly overlooking those clues that tell us it cannot be, the clues that otherwise remind us of its representative or mimetic nature. This suspension is thought to be a natural part of receiving any kind of storytelling and especially audiovisual entertainment. Given that sound anchors audiences to some sort of *real*, to what extent does sound already bypass the suspension of disbelief mechanism in the immersive and hi-fidelity cinematic context? Does this work differently in different viewing environments: small screen with headphones, television, or virtual reality apparatus. Horstmann posits that the embodied experience of film spectating is defined by multi-channel reception of representation with sound the only sensory faculty implied in not only one, but two of those channels (Horstmann, 2018, p. s15). Moreover, he separates sound into "verbal" and "auditory", using *verbal* to describe the sound we see and *auditory* for what we don't see, therefore pointing to diegesis as a defining wall that separates channels of audio reception. Meanwhile Karkulahti examines Coleridge's suspension of disbelief and posits that we engage in a suspension of virtual disbelief in the virtual realm, the goal of which is constructing narrative coherence (Karhulahti, 2012, pp. 5–7).[114] In those types of immersive space with more limited interactivity and hyperreal representation with a constructed narrative, for instance 360º cinema or goggled 3D cinema, is such suspension required? This question eats at me because of unexamined implications of new technologies. If, through content analysis and case study in this exploration, sound is shown to have persuasive and emotional power in terms of its role in cross-modal perception (Bordwell & Thompson, 2010, p. 308) and its relationship to the real (Kittler, 1999), and we are to extrapolate that power to a new image paradigm with immersive hyperreal media, perhaps crossing the representational or symbolic boundary from imaginary to real, than the persuasive and psychological power of immersion can be brought into question and debate, particularly with respect to *representations* of trauma and propaganda, as well as healing and myriad other applications or effects from which the suspension of disbelief protect. In other words, if in hyperreal media we no longer need to construct Coleridge's narrative coherence, and therefore do not need to suspend disbelieve, are we subject to real multimodal experience, not only through sound. If so, what does this imply? To test this, we would start by partnering with a cognitive neuroscientist to see if we can isolate the mechanism of suspension most likely using fMRI, and thereby devise a study to see how

[114] Here Karhulahti defines virtuality as requiring interactivity.

sound does or does not engage in this mechanism, and then extrapolate further. We are currently engaged in the development of a European University for Film and Media Arts, FilmEU, established between four universities in four countries and taking on many associated partners elsewhere. In the context of FilmEU, we have developed a research initiative in neurocine in which we will partner with neuroscientists via an interdisciplinary model of exploration. This thesis will feed directly into that work, particularly in the investigation of the suspension of disbelief as both a tool to elaborate on the experience of the real and its ramifications.

A New Rhetoric of Film Sound

This research only scratches the surface of using literary technique, namely rhetoric, as a mechanism for textual analysis of sound design. Constructing a new theory based on narrative discourse and rhetorical techniques of sound editing and mixing, one that generates a comprehensive glossary of rhetorical devices, presents itself as a needed addition to the current cannon and would provide a new vocabulary for further critical studies. We believe this could be rooted in Genette and expanded by articulating it with Van Dyke's critical discourse analysis (Dijk, 1993) as an exploration of ideology in language, engaging in lexicon, propositions and semantics as a possible triangulation with Genette's rhetoric and hermeneutics. Thus a new rhetoric would attempt to reconcile both sides of the conversation.

BIBLIOGRAPHY

CITED SOURCES

Alber, J., & Fludernik, J. (2014). *Mediacy and narrative mediation*. The Living Handbook of Narratology. http://www.lhn.uni-hamburg.de/node/28.html

Alsobrook, R. T. (2000). Machines That Made the Movie, Part 1. *International Cinematographers Guild*.

Altman, R. (1985). The Evolution of Sound Technology. In E. Weis & J. Belton (Eds.), *Film Sound: Theory and Practice* (pp. 44–53). Columbia University Press.

American Psychological Association. (2020). *APA Dictionary of Psychology: Narrative psychology*. American Psychological Association. https://dictionary.apa.org/narrative-psychology

Antunes, L. (2014). Neural Correlates of the Multisensory Film Experience. In *Neuroscience And Media: New Understandings And Representations.* Routledge.

Ashby, H. (Director). (1971). *Harold and Maude*. Paramount Pictures.

Avid Technology. (2018). *Pro Tools* (2018.7) [Computer Software].

Avildsen, J. (Director). (1976). *Rocky*. United Artists.

Baracco, A. (2017). *Hermeneutics of the Film World*. Springer International Publishing. https://doi.org/10.1007/978-3-319-65400-3

Beasley-Murray, J. (1997). Whatever Happened to Neorealism? Bazin, Deleuze, and Tarkovsky's Long Take. *Iris*, *23*, 37–52.

Beck, J. (2008). The Sounds of "Silence": Dolby Stero, Sound Design, and The Silence of the Lambs. In J. Beck & T. Grajeda (Eds.), *Lowering The Boom: Critical Studies in Film Sound* (pp. 68–83). University of Illinois Press.

Beck, J., & Grajeda, T. (Eds.). (2008). *Lowering The Boom: Critical Studies in Film Sound*. University of Illinois Press.

Behrens, L. (1979). The Argument in Film: Applying Rhetorical Theory to Film Criticism. *Journal of the University Film Association*, *31*(3), 3–11. JSTOR.

Bellucci, F. (2021). Peirce on Symbols. *Archiv Für Geschichte Der Philosophie*, *103*(1), 169–188. https://doi.org/10.1515/agph-2017-0087

Belton, J. (2008). The Phenomenology of Film Sound: Robert Bresson's "A Man Escaped." In J. Beck & T. Grajeda (Eds.), *Lowering The Boom: Critical Studies in Film Sound*. University of Illinois Press.

Benjamin, W. (1968). *Illuminations* (H. Arendt, Ed.; H. Zohn, Trans.). Schocken Books.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*(20), 11818–11823. https://doi.org/10.1073/pnas.191355898

Boillat, A. (2009). La « diégèse » dans son acception filmologique. Origine, postérité et productivité d'un concept. *Cinémas*, *19*(2–3), 217–245. https://doi.org/10.7202/037554ar

Booth, W. C. (1963). The Rhetorical Stance. *College Composition and Communication*, *14*(3), 139–145. JSTOR. https://doi.org/10.2307/355048

Bordwell, D., & Thompson, K. (2010). *Film art: An introduction* (9th ed). McGraw-Hill.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. MIT Press.

Bresson, R. (Director). (1956, November 11). *A Man Escaped*. Gaumont Film Company.

Bresson, R. (Director). (1974, September 26). *Lancelot du Lac*. Gaumont Film Company.

Brickman, P. (1980). A Social Psychology of Human Concerns. In R. Gilmour & S. Duck (Eds.), *The Development of social psychology*. Academic Press.

Buckland, W. (2013). The Acousmatic Voice And Metaleptic Narration in Inland Empire. In C. Vernallis, J. Richardson, & A. Herzog (Eds.), *The Oxford handbook of sound and image in digital media* (pp. 236–249). Oxford University Press.

Burton, T. (Director). (1992). *Batman Returns*. Warner Bros.

Cambridge Dictionary. (n.d.). Rhetoric. In *Cambridge Dictionary.org*. Retrieved October 22, 2021, from https://dictionary.cambridge.org/dictionary/english/rhetoric

Cameron, J. (Director). (2009). *Avatar*. 20th Century Fox.

Carbone, M. (2014). The Philosopher And The Moviemaker. In K. Novotný, P. Rodrigo, J. Slatman, & S. Stoller (Eds.), *Corporeity and affectivity: Dedicated to Maurice Merleau-Ponty* (pp. 220–232). Leiden.

Castelvecchi, S. (2020). On "Diegesis" and "Diegetic": Words and Concepts. *Journal of the American Musicological Society*, *73*(1), 149–171. https://doi.org/10.1525/jams.2020.73.1.149

Chaplin, C. (Director). (1923, September 26). *A Woman of Paris*. United Artists.

Chion, M. (1985). *Le son au cinéma*. Editions de l'Etoile : Diffusion, Seuil.

Chion, M. (1991). Quiet Revolution... And Rigid Stagnation (B. Brewster, Trans.). *October*, *58*, 69–80. https://doi.org/10.2307/778798

Chion, M. (1994). *Audio-vision: Sound on screen* (C. Gorbman, Ed. & Trans.). Columbia University Press.

Chion, M. (2009). *Film, a sound art* (English ed.). Columbia University Press.

Coëgnarts, M. (2015). *Embodied Cognition and Cinema: The Sensory-Motor Grounding of Abstract Meaning in Film* [Phd, Universiteit Antwerpen].

https://www.academia.edu/20109660/2015_Embodied_Cognition_and_Cinema_The_Sensory_Motor_Grounding_of_Abstract_Meaning_in_Film

Collins, P. (1981). *In The Air Tonight* [Single record]. Virgin-Atlantic.

Cooper, L. (Ed.). (1988). *The Rhetoric of Aristotle* (34. print). Prentice-Hall.

Coulthard, L. (2013). Dirty Sound. In C. Vernallis, J. Richardson, & A. Herzog (Eds.), *The Oxford handbook of sound and image in digital media* (pp. 115–126). Oxford University Press.

Crosland, A. (Director). (1927, October 6). *The Jazz Singer*. Warner Bros. Pictures.

Cuadrado, F., Lopez-Cobo, I., Mateos-Blanco, T., & Tajadura-Jiménez, A. (2020). Arousing the Sound: A Field Study on the Emotional Impact on Children of Arousing Sound Design and 3D Audio Spatialization in an Audio Story. *Frontiers in Psychology*, *11*, 737. https://doi.org/10.3389/fpsyg.2020.00737

Darabont, F. (Director). (1994). *The Shawshank Redemption*. Columbia Pictures.

Dascal, M., & Gross, A. G. (1999). The Marriage of Pragmatics and Rhetoric. *Philosophy and Rhetoric*, *32*(2), 107–130. https://doi.org/10.1353/par.1999.0001

Daud, S. S., & Sudirman, R. (2016). Discovering sound effect on visual memory performance based on electroencephalography. *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 210–215. https://doi.org/10.1109/IECBES.2016.7843444

Dearing, J., & Meyer, G. (2006). Revisiting diffusion theory. In A. Singhal & J. W. Dearing (Eds.), *Communication of Innovations: A Journey with Ev Rogers* (pp. 29–60). SAGE Publications India Pvt Ltd. https://sk.sagepub.com/books/communication-of-innovations

Deleuze, G. (1986). *Cinema 2: The Time-Image*. University of Minnesota.

Deleuze, G., & Guattari, F. (1986). *Kafka: Toward a minor literature*. University of Minnesota Press.

Deleuze, G., & Guattari, F. (1987). *A Thousand Plateaus*. University of Minnesota Press.

Demme, J. (Director). (1991). *Silence of The Lambs*. Orion Pictures.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2005). *The SAGE handbook of qualitative research* (3rd ed). Sage Publications.

Dienstfrey. (2016). The Myth of the Speakers: A Critical Reexamination of Dolby History. *Film History*, *28*(1), 167. https://doi.org/10.2979/filmhistory.28.1.06

Dijk, T. A. V. (1993). Principles of critical discourse analysis. *Discourse & Society*, *4*(2), 249-283,. https://doi.org/10.1177/0957926593004002006

@Dolby. (2021, June 8). *"This is a sonic revolution." Hear music producers and engineer*. Twitter. https://twitter.com/dolby/status/1402294605668827146

Dufourt, H. (1999). Pierre Schaeffer: Le son comme phénomène de civilisation. In J.-C. Thomas (Ed.), *Ouïr, entendre, écouter, comprendre après Schaeffer* (pp. 69–82). INA-Buchet/Chastel.

Eisenstein, S. (Director). (1925, April 28). *Strike*. Goskino.

Eisenstein, S., Pudovkin, V., & Alexandrov, G. (1988). *Statement on Sound* (R. Taylor & I. Christie, Eds.). Harvard Univeresity Press.

Eom, J.-S., Kim, S.-H., Jung, S.-S., & Sohn, J.-H. (2014). Low Frequency Noise and It's Psychological Effects. *Journal of the Ergonomics Society of Korea*, *33*(1), 39–48. https://doi.org/10.5143/JESK.2014.33.1.39

Fano, M. (1987). 2. Musique et film: Filmusique. *Vibrations*, *4*(1), 23–27. https://doi.org/10.3406/vibra.1987.978

Fischer, L. (1985). “Enthusiasm”: From Kino-Eye to Radio-Eye. In E. Weis & J. Belton (Eds.), *Film Sound: Theory and Practice* (pp. 247–264). Columbia University Press.

Flans, R. (2005, May 1). *Classic Tracks: Phil Collins’ “In the Air Tonight.”* Mix: Professional Audio & Music Production. https://web.archive.org/web/20070317065230/http://mixonline.com/mag/audio_phil_collins_air/index.html

Francis Ford Coppola (Director). (1979). *Apocalypse Now*. United Artists.

Frayne, J. (1985, April). History of Disk Recording. *Journal of the Audio Engineering Society*, *33*(4), 263–270.

Gasnier, L. J. (Director). (1936). *Reefer Madness*. Motion Picture Venutres. https://www.youtube.com/watch?v=zhQlcMHhF3w&ab_channel=Inter-Path%C3%A9

Genette, G. (1979). *Figures II*. French & European Pubns.

Genette, G. (1983). *Narrative Discourse: An Essay in Method* (J. E. Lewin, Trans.; illustated reprint, Vol. 3). Cornell University Press.

Genette, G. (1988). *Narrative discourse revisited*. Cornell University Press.

Genette, G. (2009). *Codicille*. Seuil.

Gerra, G., Zaimovic, A., Franchini, D., Palladino, M., Giucastro, G., Reali, N., Maestri, D., Caccavari, R., Delsignore, R., & Brambilla, F. (1998). Neuroendocrine responses of healthy volunteers to `techno-music’: Relationships with personality traits and emotional state. *International Journal of Psychophysiology*, *28*(1), 99–111. https://doi.org/10.1016/S0167-8760(97)00071-8

Godfrey Reggio, Cinematic Seer. (2019). Harvard Film Archive. https://harvardfilmarchive.org/programs/godfrey-reggio

Gregg, M., & Seigworth, G. J. (Eds.). (2010). *The affect theory reader*. Duke University Press.

Grice, H. P. (1975). Logic and Conversation. In J. P. Kimball, J. L. Morgan, & P. Cole (Eds.), *Syntax and semantics 3* (pp. 41–58). Academic press, Harcourt Brace Jovanovich.

Hanlon, L. (1985). Sound in Bresson's Mouchette. In E. Weis & J. Belton (Eds.), *Film Sound: Theory and Practice* (pp. 323–331). Columbia University Press.

Harley, S. (1998). *Tumbling Down* [Recorded by The Venus in Furs].

Haynes, T. (Director). (1998, October 26). *Velvet Goldmine*. Lionsgate.

Hayward, S. (2000). *Cinema studies: The key concepts* (Second ed). Routledge.

Heckerling, A. (Director). (1989). *Look Who's Talking*. Tri-Star Pictures.

Herzog, A. (2013). Charm the air to give a sound. In C. Vernallis, J. Richardson, & A. Herzog (Eds.), *The Oxford handbook of sound and image in digital media* (pp. 191–213). Oxford University Press.

Hitchcock, A. (Director). (1963, March 28). *The Birds*. Universal-International Pictures.

Holman, T. (2010). *Sound for film and television* (3rd ed). Focal Press.

Horstmann, J. (2018). Narrative representation and fictionality in performative media. *Frontiers of Narrative Studies*, *4*(s1), s5–s23. https://doi.org/10.1515/fns-2018-0030

Hughes, H. C. (2001). *Responses of Human Auditory Accociation Cortex to the Omission of an Expected Acoustic Event*. https://www.ncbi.nlm.nih.gov/pubmed/11352613.

Hven, S. (2019). The affective niches of media. *NECSUS*, *Spring 2019 #Emotions*. https://necsus-ejms.org/the-affective-niches-of-media/

IBM Corp. (2020). *IBM SPSS Statistics for Macintosh* (27.0.1.0). IBM Corp.

Illuminated Integration. (2020, October 26). Sound Design for Theatre. *Sound Design for Theatre*. https://illuminated-integration.com/blog/sound-design-for-theatre/

Inayat Khan. (1996). *The mysticism of sound and music* (Rev. ed., 1st Shambhala ed). Shambhala ; Distributed in the United States by Random House.

Karhulahti, V.-M. (2012). Suspending Virtual Disbelief: A Perspective on Narrative Coherence. In D. Oyarzun, F. Peinado, R. M. Young, A. Elizalde, & G. Méndez (Eds.), *Interactive Storytelling* (Vol. 7648, pp. 1–17). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34851-8_1

Kellogg, E. (1955). History of Sound Motion in Pictures. *The Journal of the Society of Motion Picture Engineers*, *64*, 295,.

Kerins, M. (2011). *Beyond Dolby (stereo): Cinema in the digital sound age*. Indiana University Press.

Kerrigan, L. (Director). (1993, September). *Clean Shaven*. Strand Releasing.

Kirsten, G. (2018). Barthes' Early Film Semiology and the Legacy of Filmology in Metz. In M. Tröhler & G. Kirsten (Eds.), *Christian Metz and the Codes of Cinema: Film Semiology and Beyond*. Amsterdam University Press. https://doi.org/10.2307/j.ctv62hdcz

Kittler, F. A. (1999). *Gramophone, film, typewriter*. Stanford University Press.

Koyaanisqatsi. (n.d.). Criterion Collection. Retrieved February 20, 2022, from https://www.criterion.com/films/28034-koyaanisqatsi

LaGravenese, R. (Director). (1998). *Living Out Loud*. New Line Cinema.

Lastra, J. (2008). Film and the Wagnerian Aspiration: Thoughts on Sound Design and the History of the Senses. In J. Beck & T. Grajeda (Eds.), *Lowering The Boom: Critical Studies in Film Sound* (pp. 123–138). University of Illinois Press.

Lee, S. (Director). (2006). *When The Levees Broke: A Requium in Four Acts* [DVD]. HBO Documentary Films.

Lehman, P. (1988). Politics, Film Theory, And The Academy. *Journal of Film and Video*, *40*(2), 43–51. JSTOR.

Lumière, A., & Lumière, L. (Directors). (1896). *L'arrivée d'un train en gare de La Ciotat*. Société Lumière.

Lynch, D. (Director). (1986). *Blue Velvet*. De Laurentiis Entertainmnet Group.

Mancini, M. (1985). The Sound Designer. In E. Weis & J. Belton (Eds.), *Film Sound: Theory and Practice* (pp. 361–368). Columbia University Press.

Marks, L. U. (2000). *The skin of the film: Intercultural cinema, embodiment, and the senses*. Duke University Press.

Marshall, C., & Rossman, G. B. (2016). *Designing qualitative research* (Sixth edition). SAGE.

Mendes, S. (Director). (2019, December 4). *1917*. Universal Studios.

Merleau-Ponty, M. (1962). *Phenomenology of perception* (C. Smith, Trans.). Routledge and Kegan Paul.

Merleau-Ponty, M. (1964). Eye and Mind. In *The Primacy of Perception* (pp. 159–190). Northwestern University Press.

Merleau-Ponty, M., Dreyfus, H. L., & Merleau-Ponty, M. (1964). The Film and the New Psychology. In *Sense and non-sense* (pp. 48–59). Northwestern University Press.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. Jossey-Bass.

Merriam-Webster. (n.d.-a). Gesamtkunstwerk. In *Merriam-Webster.com dictionary*. Retrieved August 27, 2021, from https://www.merriam-webster.com/dictionary/Gesamtkunstwerk

Merriam-Webster. (n.d.-b). Metacognition. In *Merriam-Webster.com dictionary*. Retrieved August 31, 2021, from https://www.merriam-webster.com/dictionary/metacognition

Metz, C. (1974a). *Film Language: A Semiotics of the Cinema* (M. Taylor, Trans.). University of Chicago Press.

Metz, C. (1974b). The Cinema: Language or Language System? In M. Taylor (Trans.), *Film language: A semiotics of the cinema* (University of Chicago Press ed, pp. 31–91). University of Chicago Press.

Metz, C. (1977). Le perçu et le nommé. In *Essais sémiotiques* (pp. 153–159). Klincksieck.

Metz, C. (1985). Aural Objects. In E. Weis & J. Belton (Eds.), *Film Sound: Theory and Practice* (pp. 154–161). Columbia University Press.

Miller, M. (2004, September 24). *The History of Surround Sound*. InformIT. https://www.informit.com/articles/article.aspx?p=337317

Miranda, E. R. (2005). *Computer sound design: Synthesis techniques and programming* (2. ed). Focal Press.

Mroz, M. (2013). *Temporality and film analysis* (Paperback edition). Edinburgh University Press.

Neorealism. (n.d.). Film Reference. Retrieved January 25, 2022, from http://www.filmreference.com/encyclopedia/Independent-Film-Road-Movies/Neorealism-HISTORICAL-ORIGINS-OF-ITALIAN-NEOREALISM.html

Neskov, B. (2021, March 5). *Interview with Branko Neskov* [Personal communication].

Novack, D. (Director). (2015). *Finding Babel* [Documentary]. Odessa Films, Video Project.

Olivier, L. (Director). (1948). *Hamlet*. Rank Film Distributors Ltd.

Ondaatje, M., & Murch, W. (2004). *The conversations: Walter Murch and the art of editing film* (1st pbk. ed). Knopf.

Peirce, C. S., Houser, N., & Kloesel, C. J. W. (1998). *The essential Peirce: Selected philosophical writings* (Peirce Edition Project, Ed.; Vol. 2). Indiana University Press.

Perrin, D., Michaelian, K., & Sant'Anna, A. (2020). The Phenomenology of Remembering Is an Epistemic Feeling. *Frontiers in Psychology*, *11*, 1531. https://doi.org/10.3389/fpsyg.2020.01531

Petro, L. S., Paton, A. T., & Muckli, L. (2017). Contextual modulation of primary visual cortex by auditory signals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714), 20160104. https://doi.org/10.1098/rstb.2016.0104

Proust, M. (2004). *Swann's way: In search of lost time. Book 1 Book 1* (C. Prendergast, Ed.; L. Davis, Trans.).

Punchdrunk. (2011, March 7). *Sleep No More* [Interactive Theater Performance].

QSR International Pty Ltd. (2021). *NVivo* (1.4.1.4361). QSR International Pty Ltd. https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

Radway, J. A. (2006). *Reading the romance: Women, patriarchy, and popular literature*. The University of North Carolina Press.

Reggio, G. (Director). (1982). *Koyaanisqatsi: Life Out Of Balance*. Institute for Regional Education; American Zoetrope.

Reggio, G. (Director). (2002, September 17). *Koyaanisqatsi* [DVD]. MGM.

Ricœur, P. (1981). *Hermeneutics and the human sciences: Essays on language, action, and interpretation* (J. B. Thompson, Ed.). Cambridge University Press ; Editions de la Maison des sciences de l'homme.

Ricœur, P. (1990). *Time and narrative. Vol. 1* (K. McLaughlin & D. Pellauer, Trans.; Repr, Vol. 1). Univ. of Chicago Press.

Rogers, E. M. (1983). *Diffusion of Innovations*. The Free Press.

Ross, A. (2005, June 27). Glass's "Koyaanisquatsi" and the art of film scoring. *The New Yorker*, *June 27, 2005*. https://www.newyorker.com/magazine/2005/06/27/sound-and-vision

Routledge (Firm) (Ed.). (2000). *Concise Routledge encyclopedia of philosophy*. Routledge.

Schreger, C. (1985). Altman, Dolby, and the Second Sound Revolution. In E. Weis & J. Belton (Eds.), *Film Sound: Theory and Practice* (pp. 348–355). Columbia University Press.

Sera, M. (n.d.). *The value of hermeneutics to film phenomenology: Michelangelo Antonioni's L'eclisse.* Retrieved June 6, 2021, from https://www.academia.edu/17522351/The_value_of_hermeneutics_to_film_phenomenology_Michelangelo_Antonioni_s_L_eclisse

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

Smith, D. (2018, Summer Edition). *Phenomenology*. The Stanford Encyclopedia of Philosphy. https://plato.stanford.edu/entries/phenomenology/

Smith, H., & Dean, R. T. (Eds.). (2009). *Practice-led research, research-led practice in the creative arts*. Edinburgh Univ. Press.

Sobchack, V. C. (2004). *Carnal thoughts: Embodiment and moving image culture*. University of California Press.

Sonnenschein, D. (2001). *Sound Design: The Expressive Power of Music, Voice and Sound Effects in Cinema*. Michael Wiese Productions.

Sound Ideas. (n.d.). [Sound FX Library Files].

Spielberg, S. (Director). (1977). *Close Encounters of The Third Kind*. Columbia Pictures.

Sullivan, G. (2005). *Art practice as research: Inquiry in the visual arts*. Sage Publications.

Sullivan, G. (2009). Making Space: The Purpose and Place of Practice-led Research. In H. Smith & R. T. Dean (Eds.), *Practice-led research, research-led practice in the creative arts*. Edinburgh Univ. Press.

Synessios, N. (2001). *Mirror*. I.B. Tauris Publishers.

Tarkovsky, A. (1986a). *Sculpting in time*. Bodley Head.

Tarkovsky, A. (Director). (1986b). *The Sacrifice*. Sandrew.

Taylor, H., Norden, M., Kerins, M., Jordan, R., & Thom, R. (2007). Discourses on Diegesis: On the Relevance of Terminology. *Off Screen*, *11*(8–9). https://offscreen.com/view/soundforum_2

Teruggi, D. (2007). Technology and musique concrète: The technical developments of the Groupe de Recherches Musicales and their implication in musical composition. *Organised Sound*, *12*(3), 213–231. https://doi.org/10.1017/S1355771807001914

The Institute for Regional Education. (n.d.). *Koyaanisqatsi*. Koyaanisquatsi. Retrieved June 7, 2021, from https://www.koyaanisqatsi.org/films/koyaanisqatsi.php

The Official Academy Awards® Database. (n.d.). Oscars. Retrieved February 5, 2022, from https://awardsdatabase.oscars.org/

Tomasulo, F. P. (1988, January 1). The Text-in-the-Spectator: The Role of Phenomenology in an Eclectic Theoretical Methodology. *Journal of Film and Video*, *40*(2), 20–32.

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, *53*(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Valle, A. (2015). Towards a Semiotics of the Audible. *Signata*, *6*, 65–89. https://doi.org/10.4000/signata.1063

Vertov, D. (Director). (1931). *Enthusiasm: Symphony of the Donbas*. UkrainFilm.

Vitaphone: True to its Mark. (1928). *Motion Picture Magazine*, *December*, 4. Library of Congress.

Walsh, M. (2004). The immersive spectator: A phenomenological hybrid. *Angelaki*, *9*(3), 169–185. https://doi.org/10.1080/0969725042000307709

Weis, E., & Belton, J. (Eds.). (1985). *Film Sound: Theory and Practice*. Columbia University Press.

Welles, O. (Director). (1941). *Citizen Kane*. RKO, Warner Bros., Paramount.

Welles, O. (Director). (1962, December 22). *The Trial* [35mm]. Astor Pictures Corporation.

Whittington, W. (2013). Lost In Sensation: Reevaluating the Role of Cinematic Sound in the Digital Age. In C. Vernallis, J. Richardson, & A. Herzog (Eds.), *The Oxford handbook of sound and image in digital media* (pp. 61–74). Oxford University Press.

Williams, A. (1985). Godard's Use of Sound. In E. Weis & J. Belton (Eds.), *Film Sound: Theory and Practice* (pp. 332–345). Columbia University Press.

Willig, C. (2017). *The Sage handbook of qualitative research in psychology, 2e* (2nd edition). SAGE Inc.

Winston, B. (1998). *Media technology and society: A history: from the telegraph to the Internet*. Routledge.

Yewdall, D. L. (2012). *Practical art of motion picture sound* (4th ed). Focal Press.

CITED FILMS, PERFORMANCES & MUSIC

Ashby, H. (Director). (1971). *Harold and Maude*. Paramount Pictures.

Avildsen, J. (Director). (1976). *Rocky*. United Artists.

Bresson, R. (Director). (1956, November 11). *A Man Escaped*. Gaumont Film Company.

Bresson, R. (Director). (1974, September 26). *Lancelot du Lac*. Gaumont Film Company.

Burton, T. (Director). (1992). *Batman Returns*. Warner Bros.

Cameron, J. (Director). (2009). *Avatar*. 20th Century Fox.

Chaplin, C. (Director). (1923, September 26). *A Woman of Paris*. United Artists.

Collins, P. (1981). *In The Air Tonight* [Single record]. Virgin-Atlantic.

Crosland, A. (Director). (1927, October 6). *The Jazz Singer*. Warner Bros. Pictures.

Darabont, F. (Director). (1994). *The Shawshank Redemption*. Columbia Pictures.

Demme, J. (Director). (1991). *Silence of The Lambs*. Orion Pictures.

Eisenstein, S. (Director). (1925, April 28). *Strike*. Goskino.

Francis Ford Coppola (Director). (1979). *Apocalypse Now*. United Artists.

Gasnier, L. J. (Director). (1936). *Reefer Madness*. Motion Picture Venutres. https://www.youtube.com/watch?v=zhQlcMHhF3w&ab_channel=Inter-Path%C3%A9

Harley, S. (1998). *Tumbling Down* [Recorded by The Venus in Furs].

Haynes, T. (Director). (1998, October 26). *Velvet Goldmine*. Lionsgate.

Heckerling, A. (Director). (1989). *Look Who's Talking*. Tri-Star Pictures.

Hitchcock, A. (Director). (1963, March 28). *The Birds*. Universal-International Pictures.

Kerrigan, L. (Director). (1993, September). *Clean Shaven*. Strand Releasing.

LaGravenese, R. (Director). (1998). *Living Out Loud*. New Line Cinema.

Lee, S. (Director). (2006). *When The Levees Broke: A Requium in Four Acts* [DVD]. HBO Documentary Films.

Lumière, A., & Lumière, L. (Directors). (1896). *L'arrivée d'un train en gare de La Ciotat*. Société Lumière.

Lynch, D. (Director). (1986). *Blue Velvet*. De Laurentiis Entertainmnet Group.

Mendes, S. (Director). (2019, December 4). *1917*. Universal Studios.

Novack, D. (Director). (2015). *Finding Babel* [Documentary]. Odessa Films, Video Project.

Olivier, L. (Director). (1948). *Hamlet*. Rank Film Distributors Ltd.

Punchdrunk. (2011, March 7). *Sleep No More* [Interactive Theater Performance].

Reggio, G. (Director). (1982). *Koyaanisqatsi: Life Out Of Balance*. Institute for Regional Education; American Zoetrope.

Reggio, G. (Director). (2002, September 17). *Koyaanisqatsi* [DVD]. MGM.

Spielberg, S. (Director). (1977). *Close Encounters of The Third Kind*. Columbia Pictures.

Tarkovsky, A. (Director). (1986b). *The Sacrifice*. Sandrew.

Vertov, D. (Director). (1931). *Enthusiasm: Symphony of the Donbas*. UkrainFilm.

Welles, O. (Director). (1941). *Citizen Kane*. RKO, Warner Bros., Paramount.

Welles, O. (Director). (1962, December 22). *The Trial* [35mm]. Astor Pictures Corporation.

APPENDIX A – LINKS TO *KOYAANISQATSI* VIDEO CLIPS

Sample A - Original Excerpt
https://vimeo.com/308267921

Sample B - Excerpt with Novack Sound FX
https://vimeo.com/308269202

APPENDIX B – SURVEY

Survey - Instructions

Please find a quiet place to screen the 15 min video, using the link below. Ideally, a large screen with good sound or good headphones. Good sound is essential for this survey.

You may TAKE NOTES about what you are thinking, if anything, as the video plays but DO NOT REWIND.

Also, do not view the survey until after you have finished the video.

Answer as honestly as possible. We are counting on your participation according to the instructions. THANK YOU!

VIDEO LINK: https://vimeo.com/308267921
* Required

Attention and Focus

1. In general, how would you rate your LEVEL of attention during the screening? *

Mark only one oval.

	1	2	3	4	5	
Weak Attention	◯	◯	◯	◯	◯	Strong Attention

2. How CONSISTENTLY were you able to keep your attention? *

Mark only one oval.

	1	2	3	4	5	
Inconsistent Attention	◯	◯	◯	◯	◯	Consistent Attention

3. During the rocket rise and fall, and to the end, how would you rate your LEVEL of attention? *

Mark only one oval.

	1	2	3	4	5	
Weak Attention	◯	◯	◯	◯	◯	Strong Attention

Impressions

4. How would you rate the overall viewing experience of the entire video? *

Mark only one oval.

	1	2	3	4	5	
Weak	◯	◯	◯	◯	◯	Powerful

Overall Story

5. What is the film saying? Do you feel you have a sense? *

Mark only one oval.

	1	2	3	4	5	
No understanding at all	◯	◯	◯	◯	◯	A strong understanding

6. Please describe your overall understanding, if any (what the film is saying). *

Impressions - #1

7. The clip opens with a sequence of aerial shots of the city followed by other still shots. Do you remember these images? *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

8. How strongly do you feel there is a connection between these shots? *

Mark only one oval.

	1	2	3	4	5	
Weak	◯	◯	◯	◯	◯	Extremely Strong

9. Please describe your understanding of what this sequence might be trying to communicate. *

Impressions - #2

In the next sequence, there is a nighttime aerial shot over a sports stadium. This is followed by images of office buildings at night.

10. Do you remember these specific images? *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

11. How strongly do you feel there is a connection between these images? *

Mark only one oval.

	1	2	3	4	5	
Weak	◯	◯	◯	◯	◯	Extremely Strong

12. Please describe your understanding, if any, of these shots. *

Impressions - #3

There is an image of a man sitting in a green space, smoking.

13. Do you remember this specific image? *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

14. Please describe your thoughts during this image, if any. *

Impressions - #4

There is a shot of people loading into an elevator.

15. Do you remember this specific shot? *

Mark only one oval.

	1	2	3	4	5	
No	◯	◯	◯	◯	◯	Yes, very clearly

16. Please describe your thoughts during this image, if any. *

Impressions - #5

There is an image of a transportation vehicle moving and descending.

17. Do you remember this specific image? *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

18. Please describe your thoughts during this image, if any. *

People Sequence

There is a series of images of people. Do you remember these specific images listed below?

19. A sightseeing guide *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

20. A black man in close-up, who smiles a bit *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

21. An old man shaving his neck *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

22. A bouncy woman walking and smiling a lot *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

23. An old man in a hat with a very droopy face, who turns to the camera *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

24. A man deep in a crowd with glasses *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

25. An older woman trying to light a cigarette *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

26. If you remember the older woman with the cigarette, did you hear the cigarette lighter? *

Mark only one oval.

	1	2	3	
No, not at all	◯	◯	◯	Yes, very clearly

27. At the end of the shot, she looks up. Did you know what she is looking at? *

Mark only one oval.

	1	2	
No	◯	◯	Yes

28. If yes, what do you think she was looking at? *

_____

29. Did you feel these people were in the same time and location or are they random, independent images? *

Mark only one oval.

	1	2	3	4	5	
Very Random	◯	◯	◯	◯	◯	Exact Same Time and Place

30. How strongly do you feel there is a connection between these images? *

Mark only one oval.

	1	2	3	4	5	
Weak, no connection	◯	◯	◯	◯	◯	Extremely strong connection

31. Do you remember your overall thoughts during these shots? *

Mark only one oval.

	1	2	3	4	5	
Not at all	◯	◯	◯	◯	◯	Completely

32. Please describe your understanding of story or message, if any, from this sequence of people shots. *

Sequence - stretcher to fireman

There is a series of images that may or may not be directly related to each other in physical space and time. (1) a man lifted and placed on a stretcher, (2) a blond woman sitting in a car who closes the car window, (3) a shirtless man sitting in a window, (4) a long wide shot of people walking around in rubble, (5) a fireman walking through a scene and stopping.

33. Did you feel these images were in the same time and location or random? *

Mark only one oval.

	1	2	3	4	5	
Very random	◯	◯	◯	◯	◯	Exact same time and space

34. Do you feel there is a story connection between these images? *

Mark only one oval.

	1	2	3	4	5	
Weak connection	◯	◯	◯	◯	◯	Strong connection

35. Do you remember your overall thoughts during these images? *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, completely

36. If you remember at all, please describe your thoughts, if any, during these images. *

37. If you remember at all, did this sequence seem to be connected to the previous sequence of "people" or not? *

Mark only one oval.

	1	2	3	4	5	
No, not at all connected	◯	◯	◯	◯	◯	Yes, completely connected

Do you remember?

38. There is an image of a hospital bed showing the shaking hand of an old woman. Do you remember this moment? *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

39. An older man walks past the camera in slow motion, directly staring into the camera while walking. Do you remember this moment? *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

40. There is an image of the stock market, with time-lapse and superimposed images. Do you remember this moment? *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, very clearly

41. Please describe your thoughts, if any, during this stock market shot. *

Rocket Sequence

42. Do you remember your thoughts when the rocket was taking off? No, somewhat, or Yes? *

Mark only one oval.

	1	2	3	
No	◯	◯	◯	Yes

43. If somewhat or yes, what were you thinking? *

44. Do you remember your thoughts when the rocket was soaring through the sky? *

Mark only one oval.

	1	2	3	
No	◯	◯	◯	Yes

45. If somewhat or yes, what were you thinking? *

46. Do you remember your thoughts when the rocket exploded? *

Mark only one oval.

	1	2	3	
No	◯	◯	◯	Yes

47. If somewhat or yes, what were you thinking? *

48. Do you remember your thoughts when the piece of the rocket slowly fell? *

Mark only one oval.

	1	2	3	
No	◯	◯	◯	Yes

49. If somewhat or yes, what were you thinking? *

Recommendation

50. This is a small part of a longer film. Having seen this small part of the film, would you consider seeing the entire film? *

Mark only one oval.

	1	2	3	4	5	
No, not at all	◯	◯	◯	◯	◯	Yes, definitely

MTurk Worker ID

51. Please enter MTurk Worker ID *

APPENDIX C – QUANTITATIVE RAW DATA

Appendix C – Quantitative Raw Data

```
FREQUENCIES VARIABLES=Survey_1FxNo_2FxYes Loc_1Theater_2Mturk Gender_1Male_2Fem_3X Age
```

Frequencies - Entire Sample

Statistics

		Survey_1FxNo_2FxYes Screening Version	Loc_1Theater_2Mturk Screening Location	Gender_1Male_2Fem_3X Gender	Age
N	Valid	324	324	324	324
	Missing	0	0	0	0

Frequency Table

Version of Film

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 No SFX	165	50.9	50.9	50.9
	2 With SFX	159	49.1	49.1	100.0
	Total	324	100.0	100.0	

Location of Screening

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Theater	141	43.5	43.5	43.5
	2 MTurk	183	56.5	56.5	100.0
	Total	324	100.0	100.0	

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	141	43.5	43.5	43.5
	2 Female	170	52.5	52.5	96.0
	3 Non-Conforming	7	2.2	2.2	98.1
	4 No Answer	6	1.9	1.9	100.0
	Total	324	100.0	100.0	

Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-25	135	41.7	41.7	41.7
	26-35	55	17.0	17.0	58.6
	36-50	90	27.8	27.8	86.4
	51 and over	44	13.6	13.6	100.0
	Total	324	100.0	100.0	

Frequencies - In Theater Only

Sample Size

		Survey_1FxNo_2FxYes Screening Version	Loc_1Theater_2Mturk Screening Location	Gender_1Male_2Fem_3X Gender	Age
N	Valid	141	141	141	141
	Missing	0	0	0	0

Frequency Table

Film Version

	N	%	%	%
1 No SFX	71	50.4%	50.4%	50.4%
2 With SFX	70	49.6%	49.6%	100.0%

Sample Size

	N	%	%	%
1 Theater	141	100.0%	100.0%	100.0%

Gender

	N	%	%	%
1 Male	49	34.8%	34.8%	34.8%
2 Female	80	56.7%	56.7%	91.5%
3 Non-Conforming	6	4.3%	4.3%	95.7%
4 No Answer	6	4.3%	4.3%	100.0%

Age

	N	%	%	%
18-25	133	94.3%	94.3%	94.3%
26-35	5	3.5%	3.5%	97.9%
36-50	2	1.4%	1.4%	99.3%
51 and over	1	0.7%	0.7%	100%

Frequencies - MTurk Only

Statistics

		Survey_1Fx No_2FxYes Screening Version	Loc_1Theater_2Mturk Screening Location	Gender_1 Male_2Fem_3X Gender	Age Age
N	Valid	183	183	183	183
	Missing	0	0	0	0

Film Version

	N	%	%	%
1 No SFX	94	51.4%	51.4%	51.4%
2 With SFX	89	48.6%	48.6%	100.0%

Sample Size

	N	%	%	%
2 MTurk	183	100.0%	100.0%	100.0%

Gender

	N	%	%	%
1 Male	92	50.3%	50.3%	50.3%
2 Female	90	49.2%	49.2%	99.5%
3 Non-Conforming	1	0.5%	0.5%	100.0%

Age

	N	%	%	%
18-25	2	1.1%	1.1%	1.1%
26-35	50	27.3%	27.3%	28.4%
36-50	88	48.1%	48.1%	76.5%
51 and over	43	23.5%	23.5%	100.0%

Sample Group

	N	%	%	%
Theater A no SFX	71	21.9%	21.9%	21.9%
Theater B + SFX	70	21.6%	21.6%	43.5%
MTurk A no SFX	94	29.0%	29.0%	72.5%
MTurk B + SFX	89	27.5%	27.5%	100.0%

Age * Sample Group Crosstabulation

Count

		Sample Group				
		Theater A no SFX	Theater B + SFX	MTurk A no SFX	MTurk B + SFX	Total
Age	18-25	68	65	1	1	135
	26-35	3	2	23	27	55
	36-50	0	2	44	44	90
	51 and over	0	1	26	17	44
Total		71	70	94	89	324

Gender * Sample Group Crosstabulation

Count

		Sample Group				
		Theater A no SFX	Theater B + SFX	MTurk A no SFX	MTurk B + SFX	Total
Gender	Male	28	21	49	43	141
	Female	35	45	45	45	170
	Non-Conforming	5	1	0	1	7
	No Answer	3	3	0	0	6
Total		71	70	94	89	324

T-Test - Entire Sample (All Groups) - All Questions - RED for p<.50

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Attention_level_general_1	Equal variances assumed	0.530	0.467	-0.511	322	0.610	-0.049	0.097	-0.240	0.141
	Equal variances not assumed			-0.510	314.045	0.611	-0.049	0.097	-0.240	0.141
Attention_consistency_2	Equal variances assumed	0.111	0.739	0.201	322	0.841	0.021	0.107	-0.189	0.232
	Equal variances not assumed			0.201	319.702	0.841	0.021	0.107	-0.189	0.232
Attention_rocket_3	Equal variances assumed	0.110	0.740	-0.197	322	0.844	-0.026	0.131	-0.284	0.232
	Equal variances not assumed			-0.197	321.928	0.844	-0.026	0.131	-0.284	0.232
Narrative_Clarity_overall_5	Equal variances assumed	0.005	0.946	-1.088	322	0.278	-0.136	0.125	-0.383	0.110
	Equal variances not assumed			-1.087	320.839	0.278	-0.136	0.125	-0.383	0.110
Memory_aerials_7	Equal variances assumed	0.213	0.645	-0.096	322	0.924	-0.010	0.103	-0.212	0.192
	Equal variances not assumed			-0.096	321.796	0.924	-0.010	0.103	-0.212	0.192
Memory_statium_city_10	Equal variances assumed	0.028	0.868	-0.008	322	0.993	-0.001	0.121	-0.240	0.238
	Equal variances not assumed			-0.008	321.247	0.993	-0.001	0.122	-0.240	0.238
Memory_factoryman_13	Equal variances assumed	0.246	0.620	-1.260	322	0.208	-0.227	0.180	-0.581	0.127
	Equal variances not assumed			-1.261	321.995	0.208	-0.227	0.180	-0.581	0.127
Memory_elevator_15	Equal variances assumed	0.749	0.387	-0.120	322	0.904	-0.015	0.124	-0.258	0.228
	Equal variances not assumed			-0.120	318.588	0.904	-0.015	0.124	-0.258	0.228
Memory_subway_17	Equal variances assumed	6.726	0.010	-1.363	322	0.174	-0.226	0.166	-0.553	0.100
	Equal variances not assumed			-1.361	316.705	0.174	-0.226	0.166	-0.553	0.101
Memory_guide_19	Equal variances assumed	0.530	0.467	0.060	322	0.952	0.010	0.170	-0.325	0.345
	Equal variances not assumed			0.060	321.876	0.952	0.010	0.170	-0.325	0.345
Memory_blackman_closeup_20	Equal variances assumed	28.586	0.000	2.922	322	0.004	0.274	0.094	0.089	0.458
	Equal variances not assumed			2.899	264.924	0.004	0.274	0.094	0.088	0.459
Memory_oldman_shave_21	Equal variances assumed	23.287	0.000	2.349	322	0.019	0.200	0.085	0.033	0.368
	Equal variances not assumed			2.322	227.817	0.021	0.200	0.086	0.030	0.370
Memory_bouncy_woman_22	Equal variances assumed	4.632	0.032	-0.812	322	0.417	-0.123	0.152	-0.422	0.175
	Equal variances not assumed			-0.814	320.394	0.416	-0.123	0.151	-0.421	0.175
Memory_droopyMan_23	Equal variances assumed	0.667	0.415	-0.030	322	0.976	-0.004	0.147	-0.293	0.284
	Equal variances not assumed			-0.030	321.746	0.976	-0.004	0.146	-0.293	0.284
Memory_man_glasses_incrowd_24	Equal variances assumed	1.459	0.228	-0.194	322	0.846	-0.033	0.171	-0.370	0.304
	Equal variances not assumed			-0.194	319.969	0.846	-0.033	0.171	-0.371	0.304
Memory_cigarette_lady_25	Equal variances assumed	0.018	0.892	-0.018	322	0.985	-0.002	0.119	-0.236	0.232
	Equal variances not assumed			-0.018	321.596	0.985	-0.002	0.119	-0.236	0.232
Memory_hospital_38	Equal variances assumed	0.096	0.757	0.084	322	0.933	0.012	0.139	-0.261	0.285
	Equal variances not assumed			0.084	321.995	0.933	0.012	0.139	-0.261	0.284
Memory_walking_staring_39	Equal variances assumed	2.603	0.108	0.620	322	0.536	0.107	0.172	-0.232	0.445
	Equal variances not assumed			0.621	321.630	0.535	0.107	0.172	-0.231	0.444
Memory_stockmarket_40	Equal variances assumed	0.051	0.821	0.421	322	0.674	0.071	0.168	-0.259	0.400
	Equal variances not assumed			0.421	321.878	0.674	0.071	0.168	-0.259	0.400

ImaginedSound_26	Equal variances assumed	21.348	0.000	-2.616	322	0.009	-0.171	0.065	-0.300	-0.042
	Equal variances not assumed			-2.605	301.956	0.010	-0.171	0.066	-0.300	-0.042
Phys_diegesis_cigarette_stretcher_27	Equal variances assumed	1.348	0.247	-0.580	322	0.562	-0.026	0.045	-0.115	0.063
	Equal variances not assumed			-0.580	319.744	0.562	-0.026	0.045	-0.115	0.063
Phys_diegesis_people_29	Equal variances assumed	0.471	0.493	-1.465	322	0.144	-0.214	0.146	-0.502	0.074
	Equal variances not assumed			-1.464	320.305	0.144	-0.214	0.146	-0.503	0.074
Phys_diegesis_emergency_33	Equal variances assumed	0.000	0.984	-2.269	322	0.024	-0.339	0.149	-0.633	-0.045
	Equal variances not assumed			-2.269	321.468	0.024	-0.339	0.149	-0.633	-0.045
Theme_Diegesis_aerials_8	Equal variances assumed	0.047	0.829	-1.568	322	0.118	-0.206	0.131	-0.464	0.052
	Equal variances not assumed			-1.569	321.988	0.118	-0.206	0.131	-0.464	0.052
Theme_Diegesis_stadium_city_11	Equal variances assumed	0.162	0.688	-0.920	322	0.358	-0.126	0.137	-0.395	0.143
	Equal variances not assumed			-0.920	321.542	0.358	-0.126	0.137	-0.395	0.143
Theme_diegesis_people_30	Equal variances assumed	1.973	0.161	-2.784	322	0.006	-0.358	0.129	-0.611	-0.105
	Equal variances not assumed			-2.789	320.940	0.006	-0.358	0.128	-0.610	-0.105
Theme_diegesis_emergency_34	Equal variances assumed	0.203	0.652	-1.851	322	0.065	-0.275	0.149	-0.568	0.017
	Equal variances not assumed			-1.851	321.464	0.065	-0.275	0.149	-0.568	0.017
Theme_diegesis_emergency_to_people_37	Equal variances assumed	0.096	0.756	-1.561	322	0.119	-0.208	0.133	-0.469	0.054
	Equal variances not assumed			-1.561	320.993	0.120	-0.208	0.133	-0.469	0.054
Impression_overall_4	Equal variances assumed	0.516	0.473	-0.322	322	0.748	-0.039	0.120	-0.275	0.198
	Equal variances not assumed			-0.321	317.053	0.748	-0.039	0.120	-0.275	0.198
Impression_people_31	Equal variances assumed	1.350	0.246	-1.441	322	0.151	-0.192	0.133	-0.454	0.070
	Equal variances not assumed			-1.442	322.000	0.150	-0.192	0.133	-0.453	0.070
Impression_emergency_35	Equal variances assumed	0.018	0.893	0.041	322	0.967	0.006	0.139	-0.268	0.280
	Equal variances not assumed			0.041	321.351	0.967	0.006	0.139	-0.268	0.280
Impression_rocket_takeoff_42	Equal variances assumed	2.224	0.137	-0.768	322	0.443	-0.063	0.082	-0.224	0.098
	Equal variances not assumed			-0.769	321.571	0.443	-0.063	0.082	-0.223	0.098
Impression_rocket_soar_44	Equal variances assumed	0.095	0.759	-0.283	322	0.777	-0.024	0.085	-0.192	0.143
	Equal variances not assumed			-0.283	320.446	0.777	-0.024	0.085	-0.192	0.143
Impression_rocket_explode_46	Equal variances assumed	0.829	0.363	0.495	322	0.621	0.034	0.068	-0.101	0.169
	Equal variances not assumed			0.494	319.655	0.622	0.034	0.069	-0.101	0.169
Impression_rocket_fall_48	Equal variances assumed	0.002	0.969	0.041	322	0.967	0.003	0.083	-0.161	0.167
	Equal variances not assumed			0.041	321.557	0.967	0.003	0.083	-0.161	0.167
Impression_watch_full_film_50	Equal variances assumed	0.007	0.934	-1.254	322	0.211	-0.215	0.172	-0.553	0.122
	Equal variances not assumed			-1.254	321.803	0.211	-0.215	0.172	-0.553	0.122
PhysDiegAvg Overall Physical Diegesis	Equal variances assumed	0.712	0.399	-2.273	322	0.024	-0.19314	0.08495	-0.36027	-0.02600
	Equal variances not assumed			-2.276	321.896	0.024	-0.19314	0.08487	-0.36010	-0.02617
ThemDiegAvg Overall Thematic Diegesis	Equal variances assumed	0.904	0.343	-2.487	322	0.013	-0.23456	0.09430	-0.42008	-0.04903
	Equal variances not assumed			-2.485	319.988	0.013	-0.23456	0.09438	-0.42023	-0.04888
MemoryAve Overall Memory	Equal variances assumed	0.284	0.594	0.033	322	0.974	0.00206	0.06283	-0.12154	0.12566
	Equal variances not assumed			0.033	318.922	0.974	0.00206	0.06290	-0.12169	0.12580
ImpressionAve Overall Impression	Equal variances assumed	0.401	0.527	-0.862	322	0.389	-0.06119	0.07095	-0.20078	0.07840
	Equal variances not assumed			-0.862	320.063	0.389	-0.06119	0.07101	-0.20089	0.07851

T-Test - Entire Sample (All Groups) - Significant Results Only

Notes

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Memory_blackman_closeup_20	Equal variances assumed	28.586	0.000	2.922	322	0.004	0.274	0.094	0.089	0.458
	Equal variances not assumed			2.899	264.924	0.004	0.274	0.094	0.088	0.459
Memory_oldman_shave_21	Equal variances assumed	23.287	0.000	2.349	322	0.019	0.200	0.085	0.033	0.368
	Equal variances not assumed			2.322	227.817	0.021	0.200	0.086	0.030	0.370
ImaginedSound_26	Equal variances assumed	21.348	0.000	-2.616	322	0.009	-0.171	0.065	-0.300	-0.042
	Equal variances not assumed			-2.605	301.956	0.010	-0.171	0.066	-0.300	-0.042
Phys_diegesis_emergency_33	Equal variances assumed	0.000	0.984	-2.269	322	0.024	-0.339	0.149	-0.633	-0.045
	Equal variances not assumed			-2.269	321.468	0.024	-0.339	0.149	-0.633	-0.045
Theme_diegesis_people_30	Equal variances assumed	1.973	0.161	-2.784	322	0.006	-0.358	0.129	-0.611	-0.105
	Equal variances not assumed			-2.789	320.940	0.006	-0.358	0.128	-0.610	-0.105
Theme_diegesis_emergency_34	Equal variances assumed	0.203	0.652	-1.851	322	0.065	-0.275	0.149	-0.568	0.017
	Equal variances not assumed			-1.851	321.464	0.065	-0.275	0.149	-0.568	0.017
PhysDiegAvg Overall Physical Diegesis	Equal variances assumed	0.712	0.399	-2.273	322	0.024	-0.19314	0.08495	-0.36027	-0.02600
	Equal variances not assumed			-2.276	321.896	0.024	-0.19314	0.08487	-0.36010	-0.02617
ThemDiegAvg Overall Thematic Diegesis	Equal variances assumed	0.904	0.343	-2.487	322	0.013	-0.23456	0.09430	-0.42008	-0.04903
	Equal variances not assumed			-2.485	319.988	0.013	-0.23456	0.09438	-0.42023	-0.04888

T-Test - Location: Theater Only

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Attention_level_general_1	Equal variances assumed	0.109	0.741	-1.973	139	0.050	-0.283	0.143	-0.566	0.001
	Equal variances not assumed			-1.972	136.927	0.051	-0.283	0.143	-0.567	0.001
Memory_factoryman_13	Equal variances assumed	0.703	0.403	-2.414	139	0.017	-0.681	0.282	-1.238	-0.123
	Equal variances not assumed			-2.415	138.827	0.017	-0.681	0.282	-1.238	-0.123
Theme_diegesis_people_30	Equal variances assumed	1.684	0.197	-2.509	139	0.013	-0.491	0.196	-0.877	-0.104
	Equal variances not assumed			-2.511	137.287	0.013	-0.491	0.195	-0.877	-0.104
Impression_people_31	Equal variances assumed	0.013	0.908	-2.210	139	0.029	-0.502	0.227	-0.952	-0.053
	Equal variances not assumed			-2.210	139.000	0.029	-0.502	0.227	-0.952	-0.053
ThemDiegAvg Overall Thematic Diegesis	Equal variances assumed	0.770	0.382	-2.003	139	0.047	-0.27650	0.13803	-0.54941	-0.00359
	Equal variances not assumed			-2.002	138.421	0.047	-0.27650	0.13808	-0.54952	-0.00348

T-Test - Location: MTurk Only

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Memory_blackman_closeup_20	Equal variances assumed	21.877	0.000	2.673	181	0.008	0.355	0.133	0.093	0.617
	Equal variances not assumed			2.643	148.861	0.009	0.355	0.134	0.090	0.620
Memory_oldman_shave_21	Equal variances assumed	25.612	0.000	2.426	181	0.016	0.350	0.144	0.065	0.635
	Equal variances not assumed			2.385	126.309	0.019	0.350	0.147	0.060	0.641
ImaginedSound_26	Equal variances assumed	19.615	0.000	-2.205	181	0.029	-0.164	0.074	-0.311	-0.017
	Equal variances not assumed			-2.180	147.943	0.031	-0.164	0.075	-0.312	-0.015
Phys_diegesis_emergency_33	Equal variances assumed	1.145	0.286	-2.375	181	0.019	-0.448	0.188	-0.819	-0.076
	Equal variances not assumed			-2.379	180.971	0.018	-0.448	0.188	-0.819	-0.076
Theme_diegesis_emergency_34	Equal variances assumed	1.447	0.231	-1.992	181	0.048	-0.399	0.200	-0.793	-0.004
	Equal variances not assumed			-1.987	177.761	0.048	-0.399	0.201	-0.794	-0.003
PhysDiegAvg Overall Physical Diegesis	Equal variances assumed	0.216	0.643	-1.861	181	0.064	-0.20029	0.10765	-0.41270	0.01211
	Equal variances not assumed			-1.864	180.990	0.064	-0.20029	0.10746	-0.41234	0.01175

T-Test - Gender: Male only

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Memory_blackman_closeup_20	Equal variances assumed	38.200	0.000	3.475	139	0.001	0.472	0.136	0.203	0.740
	Equal variances not assumed			3.287	88.815	0.001	0.472	0.144	0.187	0.757
Memory_oldman_shave_21	Equal variances assumed	15.820	0.000	1.922	139	0.057	0.308	0.160	-0.009	0.625
	Equal variances not assumed			1.817	88.497	0.073	0.308	0.169	-0.029	0.645
Phys_diegesis_emergency_33	Equal variances assumed	0.231	0.631	-2.375	139	0.019	-0.500	0.210	-0.916	-0.084
	Equal variances not assumed			-2.383	135.947	0.019	-0.500	0.210	-0.915	-0.085
Theme_diegesis_people_30	Equal variances assumed	2.628	0.107	-2.374	139	0.019	-0.477	0.201	-0.874	-0.080
	Equal variances not assumed			-2.408	138.871	0.017	-0.477	0.198	-0.869	-0.085
PhysDiegAvg Overall Physical Diegesis	Equal variances assumed	1.125	0.291	-2.076	139	0.040	-0.26062	0.12551	-0.50878	-0.01246
	Equal variances not assumed			-2.104	138.716	0.037	-0.26062	0.12389	-0.50558	-0.01566

T-Test - Gender: Female Only

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Memory_bouncy_woman_22	Equal variances assumed	9.860	0.002	-1.958	168	0.052	-0.411	0.210	-0.826	0.003
	Equal variances not assumed			-1.935	152.970	0.055	-0.411	0.212	-0.831	0.009
Memory_cigarette_lady_25	Equal variances assumed	11.072	0.001	-1.758	168	0.081	-0.269	0.153	-0.572	0.033
	Equal variances not assumed			-1.716	131.958	0.089	-0.269	0.157	-0.580	0.041
ImaginedSound_26	Equal variances assumed	31.388	0.000	-2.875	168	0.005	-0.250	0.087	-0.422	-0.078
	Equal variances not assumed			-2.948	152.738	0.004	-0.250	0.085	-0.418	-0.082
Theme_Diegesis_stadium_city_11	Equal variances assumed	0.991	0.321	-1.769	168	0.079	-0.324	0.183	-0.685	0.038

	Equal variances not assumed			-1.764	163.705	0.080	-0.324	0.183	-0.686	0.039
ThemDiegAvg Overall Thematic Diegesis	Equal variances assumed	0.096	0.757	-1.943	168	0.054	-0.24722	0.12721	-0.49835	0.00391
	Equal variances not assumed			-1.937	163.130	0.054	-0.24722	0.12762	-0.49922	0.00477

T-Test - Male in Theater Only

Survey_1FxNo_2FxYes Screening Version		N	Mean	Std. Deviation	Std. Error Mean
Attention_level_general_1	1 No SFX	28	3.82	0.863	0.163
	2 With SFX	21	4.00	1.049	0.229

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Memory_blackman_closeup_20	Equal variances assumed	25.783	0.000	2.116	47	0.040	0.143	0.068	0.007	0.279
	Equal variances not assumed			1.826	20.000	0.083	0.143	0.078	-0.020	0.306
Theme_diegesis_people_30	Equal variances assumed	1.603	0.212	-1.857	47	0.070	-0.667	0.359	-1.389	0.055
	Equal variances not assumed			-1.919	46.794	0.061	-0.667	0.347	-1.366	0.032

T-Test - Male on MTurk Only

Notes

Group Statistics

Survey_1FxNo_2FxYes Screening Version		N	Mean	Std. Deviation	Std. Error Mean
Attention_level_general_1	1 No SFX	49	4.67	0.555	0.079
	2 With SFX	43	4.58	0.932	0.142

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Memory_blackman_closeup_20	Equal variances assumed	23.433	0.000	3.180	90	0.002	0.618	0.194	0.232	1.005
	Equal variances not assumed			3.067	62.735	0.003	0.618	0.202	0.215	1.021
Memory_oldman_shave_21	Equal variances assumed	14.614	0.000	1.871	90	0.065	0.441	0.236	-0.027	0.909
	Equal variances not assumed			1.801	61.069	0.077	0.441	0.245	-0.049	0.930
Phys_diegesis_emergency_33	Equal variances assumed	0.751	0.389	-3.225	90	0.002	-0.774	0.240	-1.251	-0.297
	Equal variances not assumed			-3.235	89.326	0.002	-0.774	0.239	-1.250	-0.299
Theme_diegesis_emergency_34	Equal variances assumed	0.610	0.437	-1.977	90	0.051	-0.509	0.258	-1.021	0.003
	Equal variances not assumed			-1.962	85.019	0.053	-0.509	0.260	-1.025	0.007
PhysDiegAvg Overall Physical Diegesis	Equal variances assumed	0.046	0.831	-2.202	90	0.030	-0.29837	0.13551	-0.56759	-0.02915
	Equal variances not assumed			-2.210	89.524	0.030	-0.29837	0.13498	-0.56656	-0.03018

T-Test - Female in Theater Only

Notes

Group Statistics

Survey_1FxNo_2FxYes Screening Version		N	Mean	Std. Deviation	Std. Error Mean
Attention_level_general_1	1 No SFX	35	3.89	0.758	0.128
	2 With SFX	45	4.13	0.842	0.126

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Memory_factoryman_13	Equal variances assumed	0.321	0.572	-2.259	78	0.027	-0.838	0.371	-1.577	-0.099
	Equal variances not assumed			-2.247	71.669	0.028	-0.838	0.373	-1.582	-0.094
ImaginedSound_26	Equal variances assumed	8.925	0.004	-1.975	78	0.052	-0.283	0.143	-0.567	0.002
	Equal variances not assumed			-2.035	77.982	0.045	-0.283	0.139	-0.559	-0.006
Theme_Diegesis_stadium_city_11	Equal variances assumed	0.094	0.760	-2.528	78	0.013	-0.676	0.268	-1.209	-0.144
	Equal variances not assumed			-2.547	75.132	0.013	-0.676	0.265	-1.205	-0.147
ThemDiegAvg Overall Thematic Diegesis	Equal variances assumed	0.035	0.852	-2.062	78	0.043	-0.35302	0.17123	-0.69392	-0.01211
	Equal variances not assumed			-2.068	74.079	0.042	-0.35302	0.17069	-0.69311	-0.01292

T-Test - Female in MTurk Only

Notes

Survey_1FxNo_2FxYes Screening Version		N	Mean	Std. Deviation	Std. Error Mean
Attention_level_general_1	1 No SFX	45	4.71	0.626	0.093
	2 With SFX	45	4.56	0.867	0.129

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Memory_guide_19	Equal variances assumed	5.626	0.020	1.742	88	0.085	0.467	0.268	-0.066	0.999
	Equal variances not assumed			1.742	80.342	0.085	0.467	0.268	-0.066	1.000
Memory_oldman_shave_21	Equal variances assumed	13.822	0.000	1.746	88	0.084	0.289	0.165	-0.040	0.618
	Equal variances not assumed			1.746	57.178	0.086	0.289	0.165	-0.042	0.620

APPENDIX D – NVIVO CODES AND SUBCODES

Code Name
Emotion
Negative Emotional Response
Positive Emotional Response
Narrative Clarity - Meaning
High Clarity
High Narrative Understanding
High Thematic Understanding
Some Narrative Clarity
Zero Understanding
Narrative Sentiment
Negative
Neutral
Positive
Sound
Themes
Destruction
Death
Tragedy
Environment
Ghosts
Humanity
Connectivity
Disconnectedness
Interconnectedness
Diversity
General Humanity
Civilization
Futility of Life
Life Struggles
Overpopulation

Code Name
Routine
Urbanism
Work
Money
Nihilism
Technology
Time
Fleeting Life
Progress - change

APPENDIX E – QUALITATIVE FULL DATA

Themes – frequency of codes for individual questions and samples

			Codes (weighted)		Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	Struggles	Overpopulation	Routine	Urbanism	Work	Ghosts
					Weighted	Weighted	Weighted	Weighted	Weighted	Weighted	Weighted	Weighted	Weighted	Weighted	Weighted	Weighted	Weighted	Weighted	Weighted
		Mturk A	117.00	No SFX	4.00	1.00	1.00	12.00	3.00	77.00	10.00	1.00	25.00	24.00	9.00	9.00	8.00	0.00	0.00
	Q#6 Please	Port (theater) A	100.62	No SFX	6.62	3.97	3.97	22.00	17.21	62.23	10.59	7.94	19.86	7.94	0.00	9.27	6.62	0.00	0.00
	describe your	Mturk B	128.85	With SFX	5.28	2.11	14.79	23.24	17.21	80.27	14.79	1.06	35.91	14.79	0.00	7.39	3.17	3.17	0.00
	overall	Port (theater) B	110.11	With SFX	9.40	1.34	5.37	8.06	17.46	68.49	13.43	2.69	29.54	4.03	0.00	14.77	4.03	0.00	0.00
	understanding.				Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	StrugglesOverpopulation		Routine	Urbanism	Work	Ghosts
	If any (what the	All A	217.62	No SFX	10.62	4.97	4.97	18.62	39.21	139.23	20.59	8.94	44.86	31.94	0.00	18.27	14.62	0.00	0.00
	film is saying).	All B	238.97	With SFX	14.68	3.46	8.54	22.84	40.69	148.76	28.22	3.74	65.45	18.82	0.00	17.94	11.42	3.17	0.00
		ALL Change:	10%		38%	-31%	72%	23%	4%	7%	37%	-58%	46%	-41%	#DIV/0!	-2%	-22%	#DIV/0!	#DIV/0!
			9%		42%	-66%	35%	22%	1%	10%	27%	-66%	49%	-49%	#DIV/0!	59%	-39%	#DIV/0!	#DIV/0!
		Theater Change:																	
		Mturk Change:	10%		32%	111%	217%	23%	6%	4%	48%	6%	44%	-38%	#DIV/0!	-65%	-8%	#DIV/0!	#DIV/0!
			Total		Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	StrugglesOverpopulation		Routine	Urbanism	Work	Ghosts
		Mturk A	111.00	No SFX	0.00	0.00	0.00	34.00	6.00	71.00	19.00	2.00	10.00	0.00	0.00	9.00	39.00	0.00	0.00
	Q#9 Please	Port (theater) A	55.61	No SFX	2.65	0.00	0.00	10.59	3.97	37.07	1.32	2.65	7.94	0.00	0.00	5.30	19.86	0.00	0.00
	describe your	Mturk B	119.29	With SFX	2.11	1.32	1.06	39.02	2.11	74.99	10.56	2.11	19.01	5.28	3.17	1.06	32.74	1.06	0.00
	understanding	Port (theater) B	57.74	With SFX	4.03	0.00	1.34	9.40	9.40	33.57	2.69	1.34	8.06	1.34	0.00	2.69	17.46	0.00	0.00
	of what this				Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	StrugglesOverpopulation		Routine	Urbanism	Work	Ghosts
	sequence	All A	166.61	No SFX	2.65	1.32	0.00	44.59	9.97	108.07	20.32	4.65	17.94	0.00	0.00	5.30	58.86	0.00	0.00
	might be trying	All B	176.03	With SFX	6.34	1.34	1.06	47.42	11.51	108.56	13.25	3.46	27.07	6.62	3.17	3.74	50.20	1.06	0.00
	to																		
	communicate.	ALL Change:	6%		132%	1%	#DIV/0!	6%	15%	0%	-35%	-26%	51%	#DIV/0!	#DIV/0!	-41%	-15%	#DIV/0!	#DIV/0!
		Theater Change:	4%		52%	1%	#DIV/0!	-11%	137%	-9%	103%	-49%	1%	#DIV/0!	#DIV/0!	-49%	-12%	#DIV/0!	#DIV/0!
		Mturk Change:	7%		#DIV/0!	#DIV/0!	#DIV/0!	12%	-65%	6%	-44%	6%	90%	#DIV/0!	#DIV/0!	6%	-16%	#DIV/0!	#DIV/0!
			Total		Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	StrugglesOverpopulation		Routine	Urbanism	Work	Ghosts
		Mturk A	72.00	No SFX	1.00	0.00	0.00	3.00	4.00	64.00	5.00	1.00	18.00	2.00	0.00	4.00	10.00	24.00	0.00
		Port (theater) A	41.04	No SFX	2.65	0.00	1.32	1.32	0.00	35.75	5.30	0.00	3.97	1.32	0.00	5.30	7.94	11.92	0.00
	Q#12 Please	Mturk B	72.88	With SFX	0.00	1.06	3.17	2.11	3.17	63.37	5.28	2.11	16.90	2.11	0.00	10.56	9.51	16.90	0.00
	describe your	Port (theater) B	63.11	With SFX	4.03	1.34	1.34	0.00	0.00	56.40	5.37	1.34	16.11	1.34	1.34	6.71	9.40	14.77	0.00
	understanding				Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	StrugglesOverpopulation		Routine	Urbanism	Work	Ghosts
	if any of these	All A	113.04	No SFX	3.65	1.32	4.51	4.32	4.00	99.75	10.30	1.00	21.97	3.32	0.00	9.30	17.94	35.92	0.00
	shots.	All B	135.99	With SFX	4.03	2.40	4.51	2.11	3.17	119.77	10.65	3.46	33.01	3.46	1.34	17.28	18.91	31.67	0.00
		ALL Change:	20%		10%	#DIV/0!	241%	-51%	-21%	20%	3%	246%	50%	4%	#DIV/0!	86%	5%	-12%	#DIV/0!
		Theater Change:	54%		52%	#DIV/0!	1%	-100%	#DIV/0!	58%	1%	#DIV/0!	306%	1%	#DIV/0!	27%	18%	24%	#DIV/0!
		Mturk Change:	1%		-100%	#DIV/0!	#DIV/0!	-30%	-21%	-1%	6%	111%	-6%	6%	#DIV/0!	164%	-5%	-30%	#DIV/0!
			Total		Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	StrugglesOverpopulation		Routine	Urbanism	Work	Ghosts
		Mturk A	126.00	No SFX	4.00	0.00	3.00	1.00	3.00	115.00	23.00	16.00	22.00	11.00	2.00	18.00	19.00	4.00	0.00
		Port (theater) A	88.70	No SFX	2.65	0.00	1.32	0.00	1.32	83.41	14.56	27.80	11.92	3.97	0.00	15.89	9.27	0.00	0.00
	Q#32 Please	Mturk B	106.67	With SFX	5.28	0.00	0.00	0.00	1.06	99.28	20.07	19.01	25.35	8.45	2.11	8.45	15.84	0.00	0.00
	describe your	Port (theater) B	88.63	With SFX	2.69	0.00	1.34	0.00	1.34	83.26	9.40	29.54	13.43	5.37	0.00	16.11	6.71	2.69	0.00
	understanding				Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	StrugglesOverpopulation		Routine	Urbanism	Work	Ghosts
	of story or	All A	214.70	No SFX	6.65	0.00	4.32	1.00	4.32	198.41	37.56	43.80	33.92	14.97	2.00	33.89	28.27	4.00	0.00
	message. If	All B	195.30	With SFX	7.97	0.00	2.40	0.00	2.40	182.54	29.47	48.55	38.78	13.82	2.11	24.56	22.56	2.69	0.00
	any, from this																		
	sequence of																		
	people shots.	ALL Change:	-9%		20%	#DIV/0!	-45%	-100%	-45%	-8%	-22%	11%	14%	-8%	6%	-28%	-20%	-33%	#DIV/0!
		Theater Change:	0%		1%	#DIV/0!	1%	#DIV/0!	0%	-14%	-13%	6%	13%	35%	#DIV/0!	1%	-28%	#DIV/0!	#DIV/0!
		Mturk Change:	-15%		32%	#DIV/0!	-65%	-100%	-65%	-14%	-13%	19%	15%	-23%	6%	-53%	-17%	-100%	#DIV/0!
			Total		Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	StrugglesOverpopulation		Routine	Urbanism	Work	Ghosts
		Mturk A	77.00	No SFX	18.00	0.00	6.00	0.00	4.00	49.00	7.00	1.00	17.00	13.00	0.00	0.00	11.00	0.00	0.00
		Port (theater) A	52.96	No SFX	17.21	0.00	3.97	1.32	1.32	29.13	2.65	9.27	7.94	5.30	0.00	3.97	7.39	0.00	0.00
	Q#36 If you	Mturk B	69.71	With SFX	28.52	0.00	4.22	0.00	2.11	34.85	3.17	4.22	9.51	9.51	1.06	0.00	7.39	0.00	0.00
	remember at	Port (theater) B	59.09	With SFX	21.49	0.00	5.37	0.00	0.00	32.23	4.03	4.03	16.11	5.37	0.00	2.69	0.00	0.00	0.00
	all, please				Destruction	Environment	Money	Technology	Time	Humanity	Connectivity	Diversity	General HumanityLife	StrugglesOverpopulation		Routine	Urbanism	Work	Ghosts
	describe your	All A	129.96	No SFX	35.21	0.00	9.97	1.32	5.32	78.13	9.65	10.27	24.94	18.30	0.00	3.97	11.00	0.00	0.00
	thoughts, if any,	All B	128.79	With SFX	50.00	0.00	9.60	0.00	2.11	67.08	7.20	8.25	25.62	14.88	1.06	2.69	7.39	0.00	0.00
	during these																		
	images.																		
		ALL Change:	-1%		42%	#DIV/0!	-4%	-100%	-60%	-14%	-25%	-20%	3%	-19%	#DIV/0!	-32%	-33%	#DIV/0!	#DIV/0!
			12%		25%	#DIV/0!	35%	-100%	-47%	11%	52%	-57%	103%	1%	#DIV/0!	-32%	#DIV/0!	#DIV/0!	#DIV/0!
		Theater Change:																	
		Mturk Change:	-9%		58%	#DIV/0!	-30%	#DIV/0!	-47%	-29%	-55%	322%	-44%	-27%	#DIV/0!	#DIV/0!	-33%	#DIV/0!	#DIV/0!

(continued)

		Total		Destruction Weighted	Environment Weighted	Money Weighted	Technology Weighted	Time Weighted	Humanity Weighted	Connectivity Weighted	Diversity Weighted	General HumanityLife Weighted	StrugglesOverpopulation Weighted	Routine Weighted	Urbanism Weighted	Work Weighted	Ghosts Weighted
Q#41 Please describe your thoughts, if any, during this stock market shot.	Mturk A	84.00	No SFX	5.00	1.00	29.00	3.00	12.00	27.00	6.00	0.00	13.00	0.00	0.00	2.00	6.00	7.00
	Port (theater) A	41.04	No SFX	1.32	1.32	10.59	0.00	1.32	26.48	1.32	1.32	11.92	1.32	5.30	0.00	2.65	0.00
	Mturk B	68.65	With SFX	1.06	1.06	32.74	1.06	3.17	26.40	1.06	1.06	21.12	0.00	1.06	2.11	0.00	3.17
	Port (theater) B	41.63	With SFX	1.34	1.34	14.77	0.00	1.34	21.49	0.00	0.00	16.11	1.34	2.69	1.34	1.34	0.00
	All A	125.04	No SFX	6.32	2.32	39.59	3.00	13.32	53.48	7.32	1.32	24.92	2.65	5.30	2.00	8.65	7.00
	All B	110.28	With SFX	2.40	2.40	47.51	1.06	4.51	47.89	1.06	1.06	37.24	1.34	3.74	3.46	1.34	3.17
ALL Change:		-12%		-62%	3%	20%	-65%	-66%	-10%	-86%	-20%	49%	-49%	-100%	-29%	73%	-84%	-55%
Theater Change:		1%		1%	1%	39%	#DIV/0!	1%	-19%	-100%	#DIV/0!	35%	-49%	-100%	-49%	#DIV/0!	-49%	#DIV/0!
Mturk Change:		-18%		-79%	6%	13%	-65%	-74%	-2%	-82%	#DIV/0!	62%	#DIV/0!	#DIV/0!	#DIV/0!	6%	-100%	-55%

		Total		Destruction Weighted	Environment Weighted	Money Weighted	Technology Weighted	Time Weighted	Humanity Weighted	Connectivity Weighted	Diversity Weighted	General HumanityLife Weighted	StrugglesOverpopulation Weighted	Routine Weighted	Urbanism Weighted	Work Weighted	Ghosts Weighted
Q#14 Please describe your thoughts, if any, during this shot (Factory Man)	Mturk A	78.00	No SFX	1.00	0.00	1.00	3.00	1.00	72.00	23.00	0.00	10.00	0.00	6.00	2.00	9.00	0.00
	Port (theater) A	46.34	No SFX	0.00	2.65	0.00	0.00	1.32	42.37	10.59	0.00	3.97	0.00	2.65	1.32	13.24	0.00
	Mturk B	70.76	With SFX	0.00	0.00	2.11	3.17	2.11	63.37	19.01	0.00	7.39	0.00	3.17	3.17	8.45	0.00
	Port (theater) B	38.94	With SFX	0.00	2.69	0.00	0.00	0.00	36.26	13.43	0.00	6.71	0.00	3.17	1.34	2.69	0.00
	All A	124.34	No SFX	1.00	2.65	1.00	3.00	2.32	114.37	33.59	0.00	13.97	0.00	8.65	3.32	22.24	0.00
	All B	109.71	With SFX	0.00	2.69	2.11	3.17	2.11	99.63	32.44	0.00	14.11	0.00	8.54	4.51	11.14	0.00
ALL Change:		-12%		-100%	1%	111%	6%	-9%	-13%	-3%	#DIV/0!	1%	-11%	#DIV/0!	-1%	36%	-50%	#DIV/0!
Theater Change:		-16%		#DIV/0!	1%	#DIV/0!	#DIV/0!	-100%	-14%	27%	#DIV/0!	69%	-37%	#DIV/0!	103%	1%	-80%	#DIV/0!
Mturk Change:		-9%		-100%	#DIV/0!	111%	6%	111%	-12%	-17%	#DIV/0!	-26%	1%	#DIV/0!	-47%	58%	-6%	#DIV/0!

		Total		Destruction Weighted	Environment Weighted	Money Weighted	Technology Weighted	Time Weighted	Humanity Weighted	Connectivity Weighted	Diversity Weighted	General HumanityLife Weighted	StrugglesOverpopulation Weighted	Routine Weighted	Urbanism Weighted	Work Weighted	Ghosts Weighted
Q# 16 Please describe your thoughts, if any, during this shot (Elevator)	Mturk A	58.00	No SFX	1.00	0.00	0.00	0.00	1.00	56.00	3.00	1.00	18.00	3.00	6.00	1.00	11.00	0.00
	Port (theater) A	52.96	No SFX	0.00	0.00	1.32	0.00	5.30	46.34	3.97	0.00	14.56	1.32	5.30	0.00	11.92	0.00
	Mturk B	52.81	With SFX	1.06	0.00	0.00	0.00	1.06	50.70	2.11	1.06	20.07	1.06	5.28	2.11	8.45	0.00
	Port (theater) B	48.34	With SFX	0.00	0.00	0.00	0.00	0.00	48.34	1.34	0.00	9.40	0.00	4.03	5.37	12.09	0.00
	All A	110.96	No SFX	1.00	0.00	1.32	0.00	6.30	102.34	6.97	1.00	32.56	4.32	11.30	1.00	22.92	
	All B	101.15	With SFX	1.06	0.00	0.00	0.00	1.06	99.04	3.46	1.06	29.47	1.06	9.31	7.48	20.54	
ALL Change:		-9%		6%	#DIV/0!	-100%	#DIV/0!	-83%	-3%	-50%	6%	-10%	-76%	-18%	648%	-10%	#DIV/0!
Theater Change:		-9%		#DIV/0!	#DIV/0!	-100%	#DIV/0!	-100%	4%	-66%	#DIV/0!	-35%	-100%	-24%	#DIV/0!	1%	#DIV/0!
Mturk Change:		-9%		6%	#DIV/0!	#DIV/0!	#DIV/0!	6%	-9%	-30%	6%	11%	-65%	-12%	-19%	111%	-23%	#DIV/0!

Sentiment expressed per individual question

					Sentiment (not personal emotion)		
Mturk A	**Negative**	**Neutral**	**Positive**	**MTurk B**	**Negative**	**Neutral**	**Positive**
Q6	46	37	0	**Q6**	54	22	1
Q9	14	70	3	**Q9**	23	54	4
Q12	12	46	6	**Q12**	10	49	1
Q14	46	16	0	**Q14**	43	12	4
Q16	36	27	2	**Q16**	43	11	1
Q18	0	0	0	**Q18**	0	0	0
Q28	3	31	0	**Q28**	4	32	0
Q32	36	47	2	**Q32**	28	50	1
Q36	49	25	0	**Q36**	48	18	2
Q41	28	46	1	**Q41**	37	31	0
Q43	20	28	22	**Q43**	25	28	20
Q45	25	37	24	**Q45**	27	27	19
Q47	68	20	1	**Q47**	74	9	1
Q49	56	30	1	**Q49**	65	14	1
Port A	**Negative**	**Neutral**	**Positive**	**Port B**	**Negative**	**Neutral**	**Positive**
Q6	19	27	1	**Q6**	24	32	3
Q9	3	38	0	**Q9**	8	33	1
Q12	9	28	0	**Q12**	7	35	1
Q14	18	11	0	**Q14**	14	17	1
Q16	22	25	1	**Q16**	26	24	1
Q18	0	0	0	**Q18**	0	0	0
Q28	2	22	0	**Q28**	4	16	0
Q32	12	42	1	**Q32**	14	44	0
Q36	31	15	0	**Q36**	30	14	0
Q41	19	16	2	**Q41**	18	18	0
Q43	11	26	10	**Q43**	6	35	8
Q45	13	26	6	**Q45**	4	30	11
Q47	35	16	0	**Q47**	39	14	2
Q49	23	13	0	**Q49**	22	20	0

Overall Sentiment and Emotion Data, and relevant Individual Sentiment Data

	Overall Sentiment Expressed -- All Questions Coded						Mentions of No Understanding -- All Questions Coded		Expressed Emotion -- All Questions Coded			
	A : Negative	Weighted	B : Neutral	Weighted	C : Positive	Weighted	D : Zero Understanding	Weighted	E : Negative	Weighted	F : Positive	Weighted
Mturk A	439	439.00	460	460.00	62	62.00	247	247.00	53	53.00	3	3.00
Port A	217	287.30	305	403.80	21	27.80	354	468.68	19	25.15	0	0.00
Mturk B	481	508.02	357	377.06	55	58.09	269	284.11	54	57.03	5	5.28
Port B	216	290.06	332	445.83	28	37.60	324	435.09	13	17.46	2	2.69
All A	656	726.30	765	863.80	83	89.80	601	715.68	72	78.15	3	3.00
All B	697	798.08	689	822.88	83	95.69	593	719.20	67	74.49	7	7.97
ALL Change:	6%	10%	-10%	-5%	0%	7%	-1%	0%	-7%	-5%	133%	166%
Port Change:	0%	1%	9%	10%	33%	35%	-8%	-7%	-32%	-31%	#DIV/0!	#DIV/0!
Mturk Change:	10%	16%	-22%	-18%	-11%	-6%	9%	15%	2%	8%	67%	76%

	FACTORY Man - Q14 Sentiment					
	Negative	Weighted	Neutral	Weighted	Positive	Weighted
Mturk A	46	46.00	27	27.00	2	2.00
Port A	18	23.83	25	33.10	1	1.32
Mturk B	43	45.42	11	11.62	1	1.06
Port B	14	18.80	24	32.23	1	1.34
All A	64	70	52	60.10	3	3.32
All B	57	64	35	43.85	2	2.40
ALL Change:	-11%	-8%	-33%	-27%	-33%	-28%
Port Change:	-22%	-21%	-4%	-3%	0%	1%
Mturk Change:	-7%	-1%	-59%	-57%	-50%	-47%

	ELEVATOR - Q16 Sentiment					
	Negative	Weighted	Neutral	Weighted	Positive	Weighted
Mturk A	36	36.00	27	27.00	2	2.00
Port A	22	29.13	25	33.10	1	1.32
Mturk B	43	45.42	11	11.62	1	1.06
Port B	26	34.91	24	32.23	1	1.34
All A	58	65	52	60.10	3	3.32
All B	69	80	35	43.85	2	2.40
ALL Change:	19%	23%	-33%	-27%	-33%	-28%
Port Change:	18%	20%	-4%	-3%	0%	1%
Mturk Change:	19%	26%	-59%	-57%	-50%	-47%

	EMERGENCY - Q36 Sentiment					
		Weighted	Neutral	Weighted	Positive	Weighted
Mturk A	49	49.00	25	25.00	0	0.00
Port A	31	41.04	15	19.86	0	0.00
Mturk B	48	50.70	18	19.01	2	2.11
Port B	30	40.29	14	18.80	0	0.00
All A	80	90	40	44.86	0	0.00
All B	78	91	32	37.81	2	2.11
ALL Change:	-3%	1%	-20%	-16%	#DIV/0!	#DIV/0!
Port Change:	-3%	-2%	-7%	-5%	#DIV/0!	#DIV/0!
Mturk Change:	-2%	3%	-28%	-24%	#DIV/0!	#DIV/0!

Overall Themes and Humanity Sub-Themes

Overall Themes – All Questions Coded

	A: Destruction	Weighted	B: Environment	Weighted	C: Humanity	Weighted	D: Money	Weighted	E: Technology	Weighted	F: Time	Weighted
Mturk A	168	**168.00**	3	3.00	585	**585.00**	42	42.00	115	**115.00**	64	64.00
Port A	68	90.03	10	**13.24**	300	397.18	17	**22.51**	50	66.20	34	45.01
Mturk B	165	**174.27**	5	5.28	504	**532.31**	48	50.70	100	**105.62**	44	**46.47**
Port B	72	96.69	10	13.43	315	423.00	26	34.91	45	60.43	32	42.97
All A	236	258.03	13	16.24	885	982.18	59	64.51	165	181.20	98	109.01
All B	237	270.96	15	18.71	819	955.31	74	85.61	145	166.05	76	89.44
ALL Change:	0%	5%	15%	15%	-7%	-3%	25%	33%	-12%	-8%	-22%	-18%
Port Change:	6%	7%	0%	1%	5%	6%	53%	55%	-10%	-9%	-6%	-5%
Mturk Change:	-2%	4%	67%	76%	-14%	-9%	14%	21%	-13%	-8%	-31%	-27%

Overall Humanity Sub Categories – All Questions Coded

	A:	Weighted	B: Diversity	Weighted	C: General	Weighted	D: Life	Weighted	E:	Weighted	F: Routine	Weighted	G: Urbanism	Weighted	H: Work	Weighted
Mturk A	98	**96.00**	22	**22.00**	174	**174.00**	76	**76.00**	8	**8.00**	51	**51.00**	97	97.00	59	**59.00**
Port A	39	**51.63**	37	48.99	84	**111.21**	26	34.42	5	6.62	43	56.93	36	**47.66**	30	39.72
Mturk B	72	**76.04**	29	30.63	175	184.83	61	64.43	11	**11.62**	37	39.08	79	83.44	40	42.25
Port B	39	**52.37**	30	40.29	112	**150.40**	19	**25.51**	4	**5.37**	51	**68.49**	35	**47.00**	25	**33.57**
All A	137	149.63	59	70.99	258	285.21	102	110.42	13	14.62	94	107.93	133	144.66	89	98.72
All B	111	128.42	59	70.91	287	335.23	80	89.94	15	16.99	88	107.56	114	130.44	65	75.82
ALL Change:	-19%	-14%	0%	0%	11%	18%	-22%	-19%	15%	16%	-6%	0%	-14%	-10%	-27%	-23%
Port Change:	0%	1%	-19%	-18%	33%	35%	-27%	-26%	-20%	-19%	19%	20%	-3%	-1%	-17%	-15%
Mturk Change:	-27%	-22%	32%	39%	1%	6%	-20%	-15%	36%	45%	-27%	-23%	-19%	-14%	-32%	-28%

Additional Specific Relevant Questions

Q12 Stadium to City	Mentions of people	Weighted	Work-play	Weighted
Mturk A	12	12.00	26	26.00
Port A	4	5.30	10	13.24
Mturk B	19	20.07	23	24.29
Port B	10	13.43	10	13.43
All A	16	17.30	36	39.24
All B	29	33.50	33	37.72
ALL Change:	81%	94%	-8%	-4%
Port Change:	150%	154%	0%	1%
Mturk Change:	58%	67%	-12%	-7%

Q14 Man in Factory Room	Mentions of location	Mentions of location
Mturk A	9	9.00
Port A	2	2.65
Mturk B	5	5.28
Port B	1	1.34
All A	11	11.65
All B	6	6.62
ALL Change:	-45%	-43%
Port Change:	-50%	-49%
Mturk Change:	-44%	-41%

Q16 Thoughts During Elevator	Negative	Weighted	Neutral	Weighted	Positive	Weighted	Negative Emotional Response	Weighted	Positive Emotional Response	Weighted	Zero Understanding	Weighted
Mturk A	36	36.00	27	27.00	2	2.00	6	6.00	0	0.00	28	28.00
Port A	22	29.13	25	33.10	1	1.32	1	1.32	0	0.00	22	29.13
Mturk B	43	45.42	11	11.62	1	1.06	4	4.22	0	0.00	34	35.91
Port B	26	34.91	24	32.23	1	1.34	1	1.34	0	0.00	18	24.17
All A	58	65.13	52	60.10	3	3.32	7	7.32	0	0.00	50	57.13
All B	69	80.33	35	43.85	2	2.40	5	5.57	0	0.00	52	60.08
ALL Change:	19%	23%	-33%	-27%	-33%	-28%	-29%	-24%	0%		4%	5%
Port Change:	18%	20%	-4%	-3%	0%	1%	0%	1%			-18%	-17%
Mturk Change:	19%	26%	-59%	-57%	-50%	-47%	-33%	-30%			21%	28%

	Q 43 Rocket Takes Off				Q 45 Rocket Soars				Q 47 Rocket Explodes				Q 49 Rocket Falls			
	Humanity	Weighted	Technology	Weighted	Humanity	Weighted	Technology	Weighted	Humanity	Weighted	Technology	Weighted	Humanity	Weighted	Technology	Weighted
Mturk A	12	12.00	27	27.00	12	12.00	22	22.00	8	8.00	3	3.00	18	18.00	3	3.00
Port A	7	9.27	15	19.86	8	10.59	11	14.56	6	7.94	3	3.97	2	2.65	2	2.65
Mturk B	12	12.67	21	22.18	3	3.17	18	19.01	9	9.51	1	1.06	10	10.56	1	1.06
Port B	7	9.40	15	20.14	10	13.43	11	14.77	5	6.71	2	2.69	4	5.37	0	0.00
All A	19	21.27	42	46.86	20	22.59	33	36.56	14	15.94	6	6.97	20	20.65	5	5.65
All B	19	22.07	36	42.32	13	16.60	29	33.78	14	16.22	3	3.74	14	15.93	1	1.06
ALL Change:	0%	4%	-14%	-10%	-35%	-27%	-12%	-8%	0%	2%	-50%	-46%	-30%	-23%	-80%	-81%
Port Change:	0%	1%	0%	1%	25%	27%	0%	1%	-17%	-15%	-33%	-32%	100%	103%	-100%	-100%
Mturk Change:	0%	6%	-22%	-18%	-75%	-74%	-18%	-14%	13%	19%	-67%	-65%	-44%	-41%	-67%	-65%

Q28 Cigarette to Stretcher, what is she looking at?	Camera	Weighted	City	Weighted	Next scene	Weighted	Nothing	Weighted	Other	Weighted	People	Weighted	Previous Scene	Weighted	Rocket	Weighted	Sky-heaven	Weighted
Mturk A	12	12.00	1	1.00	2	2.00	4	4.00	6	6.00	4	4.00	0	0.00	4	4.00	0	0.00
Port A	2	2.65	1	1.32	0	0.00	0	0.00	6	7.94	5	6.62	0	0.00	3	3.97	8	10.59
Mturk B	13	13.73	1	1.06	2	2.11	2	2.11	5	5.28	5	5.28	0	0.00	3	3.17	3	3.17
Port B	1	1.34	0	0.00	0	0.00	0	0.00	5	6.71	1	1.34	1	1.34	5	6.71	4	5.37
All A	14	14.65	2	2.32	2	2.00	4	4.00	12	13.94	9	10.62	0	0.00	7	7.97	8	10.59
All B	14	15.07	1	1.06	2	2.11	2	2.11	10	12.00	6	6.62	1	1.34	8	9.88	7	8.54
ALL Change:	0%	3%	-50%	-55%	0%	6%	-50%	-47%	-17%	-14%	-33%	-38%	0%		14%	24%	-13%	-19%
Port Change:	-50%	-49%	-100%	-100%					-17%	-15%	-80%	-80%			67%	69%	-50%	-49%
Mturk Change:	8%	14%	0%	6%	0%	6%	-50%	-47%	-17%	-12%	25%	32%			-25%	-21%		