

# **THE ALCHEMY OF SOUND: CREATING “UNBELIEVABLE BELIEVABILITY” THROUGH AUDIOVISUAL FUSION**

by  
Rhyno Jansen

*Thesis presented in partial fulfilment of the requirements for the degree  
of Magister Philosophiae of Music Technology in the Faculty of Arts at  
the University of Stellenbosch*



Supervisor: Mr, Theo Herbst  
Co-supervisor: Prof., J. Vermeulen  
Faculty of Arts  
Department of Music

## Declaration

---

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2011

Copyright © 2011 University of Stellenbosch

All rights reserved

## Abstract

---

This thesis represents an attempt to explore the use of metadiegetic film sound and its connection with qualities displayed by human physiology. Metadiegesis is chosen as the focus of this study, in order to point out the potential of film sound to be representative and not to slavishly imitate its visual counterpart. Therefore, because metadiegetic sound is deictic in nature, the first hurdle to clear was to navigate through a terminological minefield, allowing a clear glimpse of its true meaning and its connection with the filmic image.

The researcher attempts to create hypothetical scenarios in order to analyse and discuss metadiegetic examples from films, convincingly utilised and less-convincingly so. The intention was to understand, clarify and disambiguate terminological uncertainties and inaccuracies. An exploration of asynchronous metadiegetic sound follows as refinement of the first step. This is done by resorting to existing examples in the form of extracts from films for demonstration purposes.

As a result, the use of metadiegetic film sound is clearly defined and its use explained by attaching it to three concepts: a ‘story within a story’, external story space, and internal representation. It is argued that through these concepts, sound can be amalgamated with image to create a different realm, where sound and image tell more than an audio-visual story.

## Opsomming

---

In hierdie tesis poog die navorser om die gebruik van “metadiegetic” film-klank, en die verband wat dit met die menslike fisiologie hou, te verken. “Metadiegesis” vorm dus die fokuspunt van die navorsingstuk. Hierdie fokus beklemtoon film-klank se verteenwoordigende potensiaal, en die feit dat dit nie net ’n slaafse nabootsing van die visuele aspek van film is nie. Aangesien “metadiegetic” klank dus deikties van aard is, is die eerste uitdaging om ’n weg te vind deur ’n terminologiese mynfeld. Dit word gedoen in ’n poging op ’n duidelike begrip van die ware betekenis van “metadiegesis” te ontwikkel, en om die verband wat film-klank met die visuele aspek van film het, duidelik te maak.

Die navorser poog in die eerste plek om hipotetiese scenario’s te skep om sodoende “metadiegetic” voorbeelde in films te analiseer en te bespreek. Die doel is om terminologiese onsekerhede en dubbelsinnighede te verstaan en te verklaar. Dan volg die verkenning van asinchroniese “metadiegetic” klank as verfyning van die eerste stap. Dit word gedoen deur gebruik te maak van film-uittreksels ter verduideliking van die bogenoemde stappe.

Na aanleiding hiervan, is die gebruik van metadiegetiese film-klank duidelik gedefinieer, en die toepassing daarvan gekoppel aan drie konsepte: ’n “storie binne-in ’n storie”, “eksterne storie ruimte” en “interne verteenwoordiging”. Deur middel van hierdie konsepte kan klank saamsmelt met beeld om ’n ander dimensie te betree, waar klank en beeld meer vertel as ’n oudiovisuele storie.

# Table of Content

---

<b>DECLARATION</b> .....	<b>I</b>
<b>ABSTRACT</b> .....	<b>II</b>
<b>OPSOMMING</b> .....	<b>III</b>
<b>TABLE OF CONTENT</b> .....	<b>IV</b>
<b>LIST OF TABLES</b> .....	<b>VI</b>
<b>LIST OF FIGURES</b> .....	<b>VII</b>
<b>CHAPTER 1: INTRODUCTION</b> .....	<b>1</b>
1.1 BACKGROUND/RATIONALE .....	1
1.2 MOTIVATION.....	4
1.3 HYPOTHESIS/AIMS.....	6
1.4 RESEARCH PROBLEM AND OBJECTIVES .....	6
1.5 METHODOLOGY AND STRUCTURE.....	8
<b>CHAPTER 2: THEORETICAL OUTLINE</b> .....	<b>10</b>
2.1 NARRATOLOGY.....	10
2.2 THE NARRATIVE WORLD IN FILM.....	11
2.2.1 <i>Diegesis</i> .....	11
2.2.1.1 Aesthetics.....	12
2.2.2 <i>Non-diegesis</i> .....	16
2.2.3 <i>Introduction to Metadiegesis</i> .....	17
2.2.3.1 Semiotics: The Semiosis of sound.....	18
2.2.3.2 Reality.....	24
2.3 AMBIGUITY .....	27
2.4 CONCLUSION .....	28
<b>CHAPTER 3: THE SOUND OF MEMORY</b> .....	<b>30</b>
3.1 FROM HEARING TO THINKING.....	30
3.2 A WILLING SUSPENSION OF DISBELIEF AND ILLUSION.....	34
3.3 CONCLUSION .....	46
<b>CHAPTER 4: THE LISTENER</b> .....	<b>49</b>
4.1 MODES OF LISTENING.....	50
4.1.1 <i>Causal Listening</i> .....	51
4.1.2 <i>Semantic Listening</i> .....	54

4.1.3 <i>Reduced Listening</i> .....	55
4.2 CONCLUSION .....	58
<b>CHAPTER 5: METADIEGETIC SOUND IN FILM .....</b>	<b>60</b>
5.1 ‘A STORY WITHIN A STORY’ METADIEGESIS .....	63
5.2 THE ‘EXTERNAL STORY SPACE’ .....	66
5.3 INTERNAL METADIEGETIC SPACE .....	71
5.4 CONCLUSION .....	74
<b>CONCLUSION .....</b>	<b>76</b>
<b>BIBLIOGRAPHY .....</b>	<b>79</b>
FILMOGROPHY .....	83

## List of Tables

---

TABLE 1 CHAR ILLUSTRATED BY RASKIN (1992: 2).....	23
TABLE 2: SONNENSCHNEIN (2001: 55) .....	65

## List of Figures

---

FIGURE 1: IVANA BAQUERO AND DOUG JONES IN PAN’S LABYRINTH (2001).....	13
FIGURE 2: DIEGETICAL WORLD FROM PETER JACKSON’S LORD OF THE RINGS.....	15
FIGURE 3: THE NON-DIEGETICAL WORLD OF THE VIEWERS FROM PETER JACKSON’S LORD OF THE RINGS.....	17
FIGURE 4: THE METADIEGETICAL WORLD OF THE CHARACTER FROM PETER JACKSON’S LORD OF THE RINGS.....	18
FIGURE 5: NAVY NCIS: NAVAL CRIMINAL INVESTIGATIVE SERVICE (2003 - 2008) .....	20
FIGURE 6: NARRATIVE PATH .....	23
FIGURE 7: THE COMBINED EFFORTS OF THE NARRATIVE FROM PETER JACKSON’S LORD OF THE RINGS.....	24
FIGURE 8: CROSSOVER BETWEEN NARRATIVE SOUND ELEMENTS .....	28
FIGURE 9: PERFUME (2006), DIRECTED BY TOM TYKWER .....	38
FIGURE 10: PERFUME (2006), DIRECTED BY TOM TYKWER .....	39
FIGURE 11: IN BRUGES (2008), DIRECTED BY MARTIN McDONAG.....	39
FIGURE 12: THE DEVIL’S BACKBONE (2001), DIRECTED BY GUILLERMO DEL TORO .....	42
FIGURE 13: DELICATESSEN (1991), DIRECTED BY JEAN-PIERRE JEUNET AND MARC CARO .....	45
FIGURE 14: DELICATESSEN (1991), DIRECTED BY JEAN-PIERRE JEUNET AND MARC CARO. ....	53
FIGURE 15: THE TRUMAN SHOW (1998), DIRECTED BY PETER WEIR. ....	54
FIGURE 16: LOST IN TRANSLATION (2003), DIRECTED BY SOFIA COPPOLA .....	55
FIGURE 17: HARRY POTTER AND THE ORDER OF THE PHOENIX (2007), DIRECTED BY DAVID YATES. ....	56
FIGURE 18: TERMINATOR SALVATION (2009), DIRECTED BY MCG .....	58
FIGURE 19: THE IMPLICIT CIRCLE FROM BRANIGAN (1992: 90) .....	67
FIGURE 20: PAN’S LABYRINTH (2001), DIRECTED BY GUILLERMO DEL TORO .....	68
FIGURE 21: MEMENTO (2000), DIRECTED BY CHRISTOPHER NOLAN .....	70
FIGURE 22: HARRY POTTER AND THE PRISONER OF AZKABAN (2004), DIRECTED BY ALFONSO CUARÓN.....	72
FIGURE 23: ETERNAL SUNSHINE OF THE SPOTLESS MIND (2004) DIRECTED BY MICHEL GONDRY .....	73
FIGURE 24: LORD OF THE RINGS: THE FELLOWSHIP OF THE RING (2001) DIRECTED BY PETER JACKSON .....	74



# Chapter 1: Introduction

---

## 1.1 Background/Rationale

In his book, *Film Technique and Film Acting* Pudovkin<sup>1</sup>(1974: 183) questioned “what new content can be brought into the cinema by the use of sound?”

Although this question was posed in the early 1900s, it continues to hold potential and can undoubtedly be revisited. One of the reasons for this is that Pudovkin argued that, technologically speaking, film had almost been perfected. It has to be remembered that by the time that the English edition of his book was published in 1929,<sup>2</sup> the conceptualisation of *film sound* was in its infancy. The audio engineering challenges, including the choice of appropriate technical equipment and microphone placement, had been met in terms of the requirements set by the radio broadcasting fraternity. It was therefore to be expected that radio stations' technicians stepped in to address technical challenges posed by the use of sound in film.

Despite this reliance on professional technical input in order to co-ordinate the visual and auditory or sound<sup>3</sup> elements, the *synchronisation* between the two constituted uncharted territory. Most importantly, the expressive potential harboured here was not immediately recognised. According to Mott (1990: 30), film sound received very little attention in the early stages of the industry. Delivery of the most convincing visual end product was all-important and undoubtedly took precedence. Therefore, most of the funding was allocated to improving the *visual* qualities of the final product with the result that film studios' sound departments were left little room for experimentation or improvement.

---

<sup>1</sup> Vsevolod Illarionovich Pudovkin (\*1893 †1953), a contemporary of Eisenstein and perhaps best known for his theories concerning *montages*, i.e. cut up images, stuck together to form a single filmic sequence.

<sup>2</sup> The first English edition appeared in 1929, this research endeavour relies on the 1974 English edition.

<sup>3</sup> *Sound*, a problematic term, will for the purpose of this research study imply *not music* in film, i.e. dialogue, Foley, ambience, sound effects, etc.

Actually, *talking* films encountered quite serious opposition. Mott (1990), referring to the work of the iconic Chaplin<sup>4</sup> argued that:

Charlie Chaplin, the great silent screen pantomimist, perhaps fearing what the addition of dialogue would do to his beloved *Little Tramp* character, had this to say about sound: “They are spoiling the oldest art in the world – the art of pantomime. They are ruining the great beauty of silence. They are defeating the meaning of the screen”(p.35).

Pudovkin stated (1974: 183) rather enigmatically that, regarding the utilisation of sound for film in the form of a soundtrack, the only element that still needed to be perfected was the aspect of sound’s “development as a means of expression”. In all probability, what he meant by this was, that film-sound’s “expressive achievements” were still lagging behind its “technical potential”. Put differently, he objected to the fact that the only sound that was audible, was the sound associated with objects and elements that appeared on screen. Therefore, “expressive” elements of film were limited to *two dimensions*. What film lacked, was an additional or *third dimension* where *sound tells the story where visuals fail*.

Sonnenschein (2001: xix) agreed and defined this *third dimension* by stating: “Giving meaning to noise, sound becomes communication”. He continued by arguing that, in reality, it was the *sound* of a spear piercing the side of a deer that lead to its death; the cascading of a waterfall that induced the earth to *tremble*; the *blaring* trumpet announcing battle *rallied* the soldiers.

Because of its complexity, the study of the expressive qualities of film sound presents an array of challenges. A particular definition of aesthetic judgment can be used as an example. According to the Oxford English dictionary, aesthetics is defined as “the critique and perception of what is beautiful”.<sup>5</sup> Here, judgement is heavily dependent on the observer's frame of reference. Gracyk (2003), for example argued that:

Good taste presupposes an active imagination. Suppose one wakes in the morning and smells the distinctive aroma of coffee, and the experience is pleasurable. This

---

<sup>4</sup> Charles Spencer Chaplin (\*1889 †1977).

<sup>5</sup> “aesthetic”, a. and n. The Oxford English Dictionary. 2nd ed. 1989. OED Online. Oxford University Press. 18 May. 2007 <<http://dictionary.oed.com.ez.sun.ac.za/cgi/entry/50003631>>.

appreciation depends on a learned, imaginative association: the smell brings to mind its cause, the brewing coffee, and its purpose, the consumption of the coffee. The agreeable sentiment is a response to this complex association of impressions and ideas, not to the smell alone.

Gracyk resorted to Hume's (1826: 224) position that aesthetic judgement "considers not the object simply, as it is in itself, but surveys it with all the circumstances, which attend it" to explore the "context" in which an "evaluation" takes place. Critical evaluation is therefore contextual.

Furthermore, aesthetic judgement relating to the perception of *sound in film* ideally relies on the simplification of the complex nature of the human mind. In this regard, it is helpful to distinguish between *synchronous* and *asynchronous* sounds. As the term indicates, synchronous sounds are synchronized with the filmstrip, for example, the imagery, while asynchronous sounds are all the sounds that cannot directly be associated or do not occur in synchronization with the imagery. This research endeavour will focus on the sounds that occur asynchronously to the film image.

An example of asynchronous sounds include remnants of the whispers of yesterday's heated conversation still burning and screaming in the back of one's mind; the sounds one thinks one heard while alone in the dark and the echoing memories manifesting themselves in the conscious and subconscious mind. This is the domain of *metadiegesis*, which will be discussed later on in this research endeavour.

Etymologically, the term is derived from the word *diegesis*, which is understood by Gorbman (1980: 197) to apply to "music that (apparently) issues from a source within the narrative". Although Chion (1994: 109) includes music in his discussion of the term diegesis, to him it also applies to "...voices..." and "...noise..." i.e. sound as defined above in footnote three.

Within the third dimension alluded to above, a field warranting further study, would be Pudovkin's finding (1974: 184) that, "the role which sound is to play in film is much more significant than a slavish imitation of naturalism...". To elaborate, film-sound does not just imitate existing realities, but can even imply and suggest silent, mute or unseen objects and things. This perspective dissolves the borders within which sound has traditionally been caged. Biró (1982: 5) states that "the growing

interest in film language and film structure may be traced back to this need for independence, film's rejection of the tyranny of naturalism as well as of the approach and achievements of its sister arts". She calls this the "second revolution of the film".

The following could serve as an example of the latter: a film character is awakened abruptly by a loud noise. He sits up, straining to identify its source. The audience hears the ticking of a wall mounted clock, the creaking of the house around the character as structurally it settles after a hot summer's day, but has no clue as to the origin(s) of the sound. Suddenly, the character realises that he had been dreaming about sauntering at a carnival, paying casual attention to surrounding conversations when a warning sounds: "LOOK OUT!" At that moment, an unidentifiable object crashes down on the stand next to him.

This sequence of events, as recalled by the character, was illusionary and all associated sounds were inaudible. Still, he *listened* to different conversations and was startled by a loud noise in close proximity. It is the richness and diversity of the dimensions associated with this example and the potential for fleshing it out while creating a doorway into the character's state of mind, that the researcher intends to explore. To use the above terminology, this would mean that the endeavour revolves around the investigation of metadiegetic sound through asynchronism to investigate the *third dimension*.

The aims and reasons of this study are as follows:

## 1.2 Motivation

The realisation of the potential harboured by the problem that psychoacoustics deals with the perception of sound. Therefore, the memory of the perceived sound is not sound itself, but is drawn from our *perceived library of sounds in our own frame of reference*. For example, when hearing a crashing noise emanating from a street while no visual reference as to its origin exists, one would assume, better yet, *know*, that an automobile accident had occurred. This knowledge would obviously presuppose the idea that one must have heard similar sounds, before allowing the brain to reconstruct what has transpired. Morris (2001: 368) writes:

The sound of memory can be a tangible “recording” of how an event is remembered acoustically, while the memory of sound presupposes a relationship to the sound that once was and is now lost.

This is also true for sounds one is not familiar with. Forrester (n.d.) in describing this writes:

But consider, it is on hearing the [scratching] noise that we then imagine that the sound is the kind of noise a rodent might make when scraping or scratching around for food. Our knowledge of such sounds has come from the cultural repertoire of all those available imaginable sounds, i.e., we don't in reality have to have seen a rat or mouse making such a sound, a great deal of our knowledge comes from the available cultural discourses about sounds and their causes.

The conclusion can be drawn that memory is not sound, it is a type of *representation* of that which is real. And, because memory is not sound, there is no limit to the ways in which it (memory) can be expressed in order to convince an audience of that which they observe visually, even when it is *phantasmagorical*.<sup>6</sup>

The researcher decided to focus on metadiegetic sound in film for the reason that its use is undervalued. The very idea that someone can be manipulated into temporarily believing something that is not real, to exist. Furthermore, realising the error in the argument that memory of sound is not sound itself and therefore can be used to the sound designer's advantage. By focussing on sound elements that were once heard, but now lost, the sound designer can create believable sounds for phantasmagorical scenes, from sounds the audience has previously heard, or can relate to.

Thus far, narrative film theory mainly addressed issues regarding actors' dialogue, cinematographical construction, the script (story), and even music, but narrative never referred to sound and sound events within film other than the dialogue. The idea that sound design can be used to create an outer-space battle scene (Star wars), magical

---

<sup>6</sup> Phantasmagoria - “A shifting and changing scene consisting of many elements, *esp.* one that is startling or extraordinary, or resembling or reminiscent of a dream, hallucination, etc.” (“Phantasmagoria, *n.*” OED Online. June 2007. Oxford University Press. 10 Dec. 2004 <<http://dictionary.oed.com/cgi/entry/50177120>>)

forests where trees come alive (Lord of the rings), or giving feeling to inanimate objects (Toy story), can be realised by opening a door into the subconscious mind of the protagonist.

The physiological and psychological link between sound and memory has long been open to debate. Examples of these relations would be, by creating sudden loud sounds, causes someone to flinch; through different noises being of equal loudness, will imbue the audience with panic or uneasiness; using soft sounds with long stretches of utter silence, induces a foreboding atmosphere. This will form the focus of this research endeavour.

### 1.3 Hypothesis/Aims

Starting as a concept describing *different dimensions* in narrative, *diegesis* is now widely used to describe the *placing of sound* in film. This term *sound*, refers mostly to music, e.g. Gorbman (1980: 197) referring to music in a metadiegetic way. However, as discussed above, it is clear that Chion (1994: 109) understands diegetic sounds as consisting of not only music, but also “...voices...” and “...noise...”.

The essence of the research project revolves around metadiegesis, which have been identified in the broader field of film sound. The aim is to argue, explore and demonstrate the potential that the asynchronic use of sound within metadiegesis hold. Further aims will include the exploration of the gargantuan task of unravelling often conflicting terminological and methodological issues, matters of a psychological and perceptual nature. For example, concerning the identification and understanding of *memory* and *conditioning*, as well as, how to audibly express the inaudible and the creation and description of hypothetical instances expressing and explaining the implementation of metadiegetic sound, etc.<sup>7</sup>

### 1.4 Research problem and objectives

Recognition of the heterogeneous nature of the cinema experience not only opens the field to consideration of a broad spectrum of objects, processes, and activities, but has an especially direct impact on the study of sound (Altman, 1992: 6)

---

<sup>7</sup> see the Research Problem and Objectives sections for further information.

From this perspective, one hears more than one sees. To put it differently, there exists greater potential to auditively evoke strong emotions and reactions, perhaps greater than is the case with the visual sense. It follows that the sound dimension to film ought to receive considerable attention. In film, visual imagery is limited to what one sees on screen whereas with film-sound there are no limits to the imaginary worlds that can be created and implied. Skilful sound design allows for the realisation of even the improbable, the only limitations being those technologically imposed or set by the sound designer. Sonnenschein (2001: 64) explains:

unlike light, sound can travel through solid materials like walls... Sound also bends around corners more easily than light, giving less privacy than our visual world. This property lends itself to telling much more of the film story than is seen on screen.

The research problem associated with this endeavour involves issues relating to manners in which metadiegetic sound inform the perception of that which is not audible.

An investigation in this field is a daunting undertaking. The first hurdle would be to negotiate the minefield of terminological ambiguities, inconsistencies and inaccuracies.

With this in mind, the research problem boils down to the following:

There exists an ambiguous relationship between metadiegetic and extradiegetic sounds. They are congruent when not attached to the visuals. In using asynchronous metadiegetic sound, the audience can confuse the sound's placement as extradiegetic. One of the few differences between the two is that metadiegetic sounds ought to be attached to a sign or symbol from the real world. This attachment is abstract in itself because one uses aesthetic and semiotic theories to attempt a telling of the story where the visuals fail.

At this point, the question as to how the inaudible is clearly and unambiguously made audible, becomes relevant. Sonnenschein (2001: 61) writes:

No matter how strange or out of context a sound may be, our minds tend to look for something to recognise. This happens when we are looking at a natural rock wall and end up finding the form of a face. These kinds of archetypal templates can work in our favour

when we discover or create non-human sounds that appear to have vocal characteristics with an 'emotional envelope'.

This state of affairs enables the formulation of the research problem in the form of a hypothesis in which it is argued that:

the alchemy of metadiegetic sounds draws power from triggers present in most human beings:

- **past experience and associations** (*memory, incidences of life*)
- **primal instincts** (*e.g. human behaviour, instinctive fears*)
- **tapping into the dreamscape of the mind** (*the conscious and unconscious imagination*)
- **the audiences' expectations and emotional impact**

The research question is therefore formulated as follows:

How would one amalgamate these triggers with aesthetic signs and symbols found in film to give a voice to the *third dimension*?

The objectives of this research endeavour will include the following:

Firstly, the researcher intends to create hypothetical scenarios, and analyse and discuss examples from films where metadiegetics are in use, convincingly utilised and less-convincingly so. Secondly, the intention is to understand, clarify and disambiguate terminological uncertainties and inaccuracies. In the third place stands the exploration of asynchronous metadiegetic sound as refinement of the first step. This will be done by resorting to existing examples in the form of extracts from films and also by creating original short examples and excerpts for demonstration purposes.

## 1.5 Methodology and Structure

In terms of the type of study, the researcher will be taking a non-empirical approach. This study draws on philosophical and conceptual analysis and theory building. The researcher intends to use philosophical and psychological terminology and conceptual frameworks in order to define and describe, in other words, clarify metadiegetic sound usage in film.



Metadiegetic sound is deictic in nature. At the beginning of the thesis, therefore, it will be necessary to identify and define specific and relevant proven, quasi-philosophical, and psychological concepts. These include, in no particular order, aesthetics, semiotics, willing suspension of disbelief, phantasmagoria and psychological theories concerning memory. The next step will be to explore the unfathomed potential of metadiegetic sound as outcome to the study.

In searching for ways for tapping into an audience's dreamscape, the researcher will rely on the description, design and limited implementation of experimental scenarios aimed at soliciting emotive reaction from a member of a film audience. This will be done by stating hypothetical instances demonstrating and highlighting aspects of applicable theories, by investigating examples from films where the third dimension was used and through the limited practical demonstration of the researcher's own examples.

Mouton (2001: 175) finds that the strength of a non-empirical approach lies in the fact that "conceptual analysis brings conceptual clarity". If the conceptual analysis is well structured, the true meaning can become clear for all to use and interpret. On the other hand, inadequate conceptual analysis leads to "theoretical ambiguities and fallacious reasoning". It is felt that this research endeavour will contribute to the former by relying on examples in order to arrive at a clearer understanding of what remains a less-investigated field of research.

## Chapter 2: Theoretical Outline

---

In this chapter, the focus will fall on the explanation of crucial terms and arguments as a doorway into the audio-visual creation that is our mind. The researcher will start by describing narrative and its relationship to film, which is the basis whereupon film is built.

### 2.1 Narratology

“Narrative is enormously complicated, even when we set aside its exchange value as a manufactured object in a community and concentrate on its use value as a psychological object for a perceiver” (Branigan, 1992: 118).

The concept, narrative, has long been under a lot of scrutiny, being defined and redefined. A typical understanding of the term narrative, as described by Branigan (1992: 36), would be that “*Narrative* is a way of comprehending space, time, and causality”. He explains that there are at least two ways of understanding the above. For him, the idea that narrative is the means to express the information from the “frame of the screen” into a diegetical world.

For Genette (1980: 25), narrative is an “ambiguous” term, and the “difficulties of narratology” arising from it “are perhaps due to this confusion”. He explains that to disambiguate this confusion, one must first understand that narrative has three notions.

Firstly, the most commonly used, “narrative refer[s] to the *narrative statement*, the oral or written discourse that undertakes to tell of an event or a series of events”.

Secondly, the notion mostly used by “theoreticians of narrative content”, uses narrative as a “succession of events, real or fictitious, that are the subjects of this discourse, and to their several relations of linking, opposition, repetition, etc.”.

Thirdly, this notion of narrative, according to Genette, is probably the oldest of the three and refers again to an event. He explains that it does not refer to the event that is recounted, but rather to the event in which someone is *recounting something*.

## 2.2 The narrative world in film

The narrative world can be divided into a number of divisions. The most important divisions are the diegetic space and time of the characters of the filmic world, the non-diegetic world of the viewers and the metadiegetic world consisting of a character's perception in the filmic world (that is to say, the narrative within the narrative). These narrative concepts were also extended to describe the narrative of sound and music. E.g., when diegetic sound is mentioned, it would be describing all the sound that exists in the space and time of the filmic world.

### 2.2.1 Diegesis

As the first level in the narrative ladder, diegesis in theory describes everything that is tangible in the world of the character. This would include all sounds heard and all items and things seen by a character, or characters in the narrative world. Some of the principal writers on diegesis whose names occur repeatedly, include the French filmologue Etienne Souriau (1953), the literary theorist Gerard Genette (1982), Christian Metz (1976), Claudia Gorbman (1980), Rick Altman (1992) and Michel Chion (1994). Souriau was responsible for defining diegesis:

Diegesis, diegetic: all that belongs, "by inference," to the narrated story, to the world supposed or proposed by the film's fiction. Ex: (a) Two sequences projected consecutively can represent two scenes separated in the diegesis by a long interval (several hours or years of diegetic time). (b) Two adjoining studio sets can represent locations supposedly hundreds of feet apart in diegetical space. (c) Sometimes there are two actors (e.g. a child and an adult, or a star and a stuntman or double) to successively depict the same diegetic character" (1953: 7).<sup>8</sup>

The other writers have extended its meaning to such an extent that they may also bask in the glory of its rays.

Another term that also warrants an introduction is aesthetics. Without an understanding of aesthetics, it is difficult to understand the subtle art of submerging viewers into an alternate reality. A short introduction into aesthetic judgment will follow.

---

<sup>8</sup> Translated by Claudia Gorbman (1980)

### 2.2.1.1 Aesthetics

To understand aesthetics better, one can refer back to the Greek meaning of the words *aisthanomai*, which mean, “I perceive” and *aisthetike*, which mean “sense perception” (Zettl, 1990: 2).

Everyone utilises aesthetic judgement. Every time one decides what to wear, where things should go when one tidies a room or as Zettl (1990: 1) puts it, “I know what I like”, one is in the process of applying aesthetic judgment. Zettl proposes that a new type of aesthetic judgment should be utilized when it comes to mass communication, he calls it, applied media aesthetics.

For him “applied media aesthetics differs from the traditional concepts of aesthetics in three major ways” (Zettl, 1990: 2). Firstly, he explains that aesthetics is no longer limited to the traditional philosophical concept dealing with the perceiving and the appreciation of beauty or the knowledge of what is beautiful. He also feels that aesthetics is not limited to “theory of art and art’s quest for truth”. He explains that instead of separating art from life, applied media aesthetics is concerned with the interdependency of art *and* life.

Secondly, he explicates that media, in today’s understanding, would refer to television and film, and “no longer” to a “neutral means of simple message distribution”. Therefore, these mediums have become an essential element “in the aesthetic communication process”.

Thirdly, the traditional use of aesthetics is to analyse and examine why art is art as well as to examine “why art is more beautiful than and closer to the truth than the ordinary things in life”. Applied media aesthetics concerns itself with giving us dependable “criteria for the analysis of existing television shows and films”.

If one considers the Greek meaning of aesthetics as the true meaning of the word, then it does not necessarily mean to know what is beautiful, but to know how to analyse what one perceives. This might include how to perceive beauty, ugliness, what is frightful, what is wretched or even what is tasty. To have a scene in a film appear beautiful or scary to the audience, one must first have the knowledge of what makes something beautiful or scary. Take for instance the film *Pan’s Labyrinth* (2001) (also known as: *El Laberinto del fauno*) (below), directed by Guillermo del Toro. This film

plays out in rural Spain of 1944. The young stepdaughter (Ofelia) of a sadistic army officer enters an eerie fantasy world, as a means to escape the harsh Spanish war.



Figure 1: Ivana Baquero and Doug Jones in Pan's Labyrinth (2001)

The story progresses to the point where Ofelia meets an old faun who gives her three dangerous tasks to complete to prove that she is the daughter of the king of this fantasy world. This story would have been absurd if one took it literally and would not have been scary at all. Therefore, one would sacrifice one's understanding and beliefs to accommodate the absurdity of the diegetic world and accept it as true, so that, as Easton (1954: 51) puts it, "an aesthetic attitude may predominate". One would then argue that for the sake of the diegetic world's aesthetic character, one would imagine, *what if something like that really happened?* This in turn would create a relationship with the lead character Ofelia, played by Ivana Baquero, to try to complete her tasks in this dark fairytale land.

This type of accommodation is called a *willing suspension of disbelief* and according to Easton (1954: 51), "is a phase of the attitude known as aesthetic, and means that one is willing to accept a lack of truth or logic or morals so that an aesthetic attitude may predominate".

Even though a willing suspension of disbelief might hold true for some, many theoreticians still criticise this idea of thinking about surrealism. Some people will not watch a science fiction film for the simple reason that it seems implausible. Nevertheless, a suspension of disbelief does not only apply to science fiction, it has to be applied to all film genres. The reason for this is that anything presented on film, is to some extent an imitation of the reality of someone's imagination. When watching a film, one is completely aware that it is not real, but as soon as one concentrates on the story unfolding in front of one's eyes, one *must sacrifice* the idea that the people are

actors and start believing they are the characters they are supposed to portray. Biro (1982: 74) reasons, “Myth can reach universality, however, only by lifting man out of the confinement of his selfishness”.

Another example, expressed by Chion (1994: 109) was about a nature lover who wrote to the *Télérama* (a French TV guide). He explained that he had seen Bertrand Blier’s *Trop belle pour toi* (1989) and was not happy that in one scene of the film, there were sounds of birds in the background that were not supposed to sing during that time of year the story took place. They were not even native to that part of the world. This film was spoiled for the nature lover; the soundtrack prevented him *believing* the scene.

...the same spectator who’s finicky about sounds might be indifferent to aberrant light (incoherent lighting setups given the light source depicted) that might disturb the photography specialist. In other words, every film is premised on the the (*sic*) acceptance of rules of the game – not the least of which is agreeing to see flat images in depth! (Chion, 1994: 109).

This understanding of aesthetics helps the audience to create a diegetical space and time that they can *argumentatively believe* exists, and serves as the basis of this study.

The previous example described diegetical sound. Thus, the question can be asked, what is meant by diegetical sound? Branigan (1992: 35) explains, “A sound in a film, for example, is diegetic if the spectator judges that it has been, or could have been, heard by a character”. This not only refers to elements on-screen, but to elements off-screen as well. The camera can argumentatively be represented as a large keyhole that the audience looks through. One can only see a limited view of the world, thus, there might be sounds present which are not in the visual shot that still counts as diegetic sounds. The only criterion being, that more than one character in the film must be able to hear the sound. The reason is that metadiegetic sound can only be heard by one character. This statement might pose as a problem that will be described later in this chapter (See 2.3 Ambiguity on page 27).

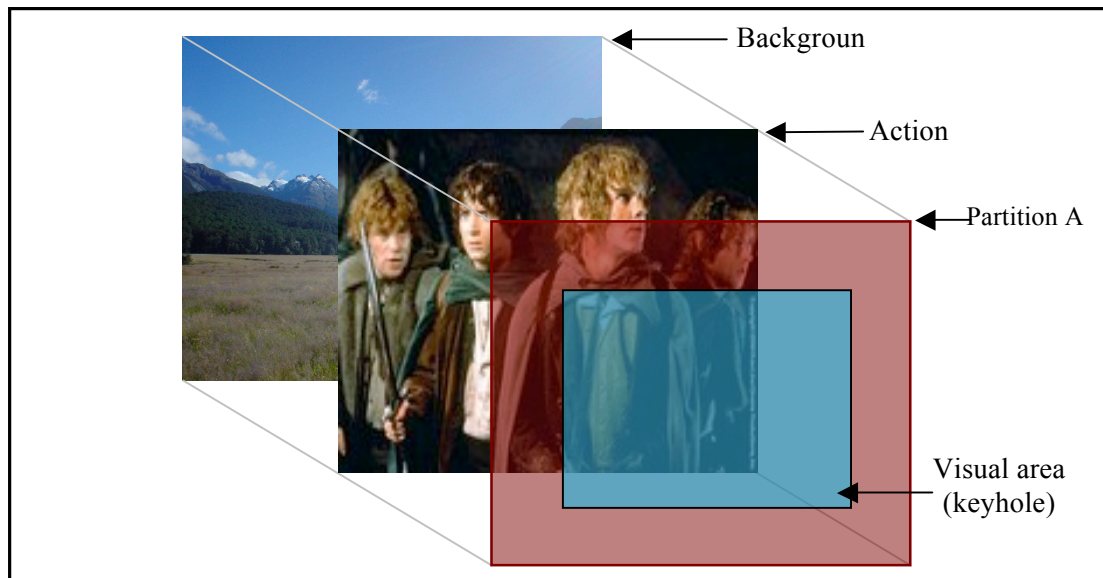


Figure 2: Diegetic world from Peter Jackson's Lord of the Rings

The diegetic level of narrative sound can be divided into two dimensions. Firstly, the action dimension, where all the main actions take place, e.g. the main story dialogue, the foreground sound effects (the actor(s) in focus'<sup>9</sup> grunting, punches, crashes, kisses, footsteps etc.). Secondly, the background sounds consisting of ambient noises, group Walla<sup>10</sup>, and all sounds not directly associated with the main plot of the story, and only serves to help place the scenes.

As indicated previously, the diegetic level consists of on-screen sounds and off-screen sounds (see Figure 2: Partition A). The blue square (indicating the on-screen area), or keyhole as explained earlier, indicates what the audience see, while the red area serves as the off-screen world that can only be perceived through sound or previously visible elements and/or characters.

In contrast to this, nondiegesis can only be experienced by the audience and will be described next.

---

<sup>9</sup> The actor in focus refers to the actors or elements being focused on by the camera. These elements or actors' sounds would be louder than all the other sounds in the film to emphasize their significance.

<sup>10</sup> *Walla* refers to all the human voices in the background that is muffled and cannot be understood (Yewdall, 2003:290-291).

### 2.2.2 Non-diegesis

“There might be something inherently paradoxical about the presence of music in films, even as our experience as spectators seems to affirm that music quite “naturally” belongs in movies” (Gorbman, 1980: 185).

There exists a lot of research material on this subject and it is not the researcher’s intention to conduct an in-depth study of film music. The idea is to provide the reader with enough information about nondiegesis as possible so that the focus of this study can egress clearly.

Nondiegetic music is one of the oldest levels in film sound, and for many years until the present, it served as background to enhance the mood of film. Branigan (1992: 35) explains:

...some on-screen elements (e.g., “mood” music) are *nondiegetic* and addressed only to the spectators. These elements are *about* the diegetic world of a character and are meant to aid the spectators in organising and interpreting that world and its events. Nondiegetic elements are not accessible to any of the characters.

Claudia Gorbman (1980: 188), one of the many researchers of film music, explains that film music possesses “its own internal logic”; it always tells the audience something about the film it accompanies. She (1980: 189) criticises scholars’ explanation of the “possible film/music relationship...” arguing that it “seems curiously primitive”. It is limited to “*parallelism* and *counterpoint*”; either resembling or it “contradicts the actions or mood of what happens on the screen”.

The stereotypical idea exist that a scary scene is accompanied by sinister music or a heroic scene is accompanied by majestic triumphant music. This statement is true to some extent, but is much more complicated. Film music is similar to poetry. It has a purpose. Adorno and Eisler (1994: 70) explains that the relationship between the film and the music is “one of question and answer, affirmation and negation, appearance and essence”, rather than similarity. This causes Gorbman (1980: 189) to ask the question: “isn’t *any* music usually sufficient to accompany a segment of film?” Her answer to this question was that all music would “*do something*”, but all scenes would be expressed differently according to the music’s “dynamics and structure” accompanying it.



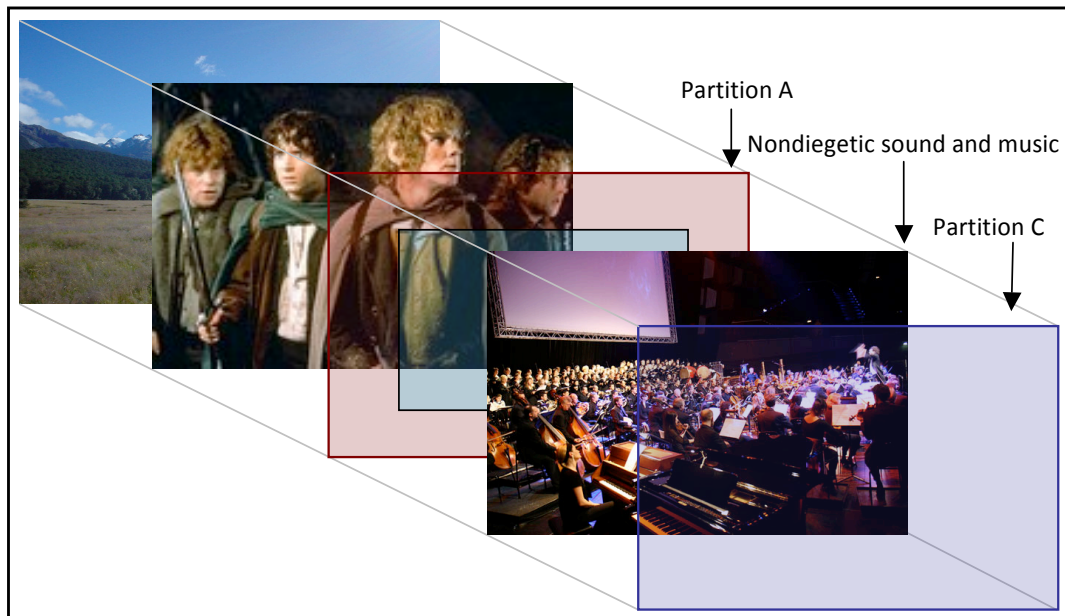


Figure 3: The non-diegetic world of the viewers from Peter Jackson's Lord of the Rings

Figure 3 demonstrates the position of the non-diegetic level in a film (placed between partition A and partition C).

The last level is known as *metadiegesis*, and is sometimes confused with diegetic and non-diegetic sounds when not used correctly. This error or ambiguous use of metadiegesis will be explained under the heading 2.3 Ambiguity on page 27.

### 2.2.3 Introduction to Metadiegesis

The objective of this section is to serve as introduction, however, the complete dissection of this term will continue in chapter 5 on page 60.

Metadiegesis can better be expressed as internal sounds - these sounds describe what is happening inside one character's mind (for example a flashback of something that the character remembers). It is also important to note that metadiegetic sounds, can only be heard by the audience and the character that is thinking or creating the internal sounds. Chion (1994: 76) audits the term internal sound and explicates that it can be further subdivided into *objective-internal* and *subjective-internal* sounds. Objective-internal sounds are physiological in origin, consisting of breathing and heartbeats. Subjective-internal sounds, on the other hand, concern mental representation and memory.

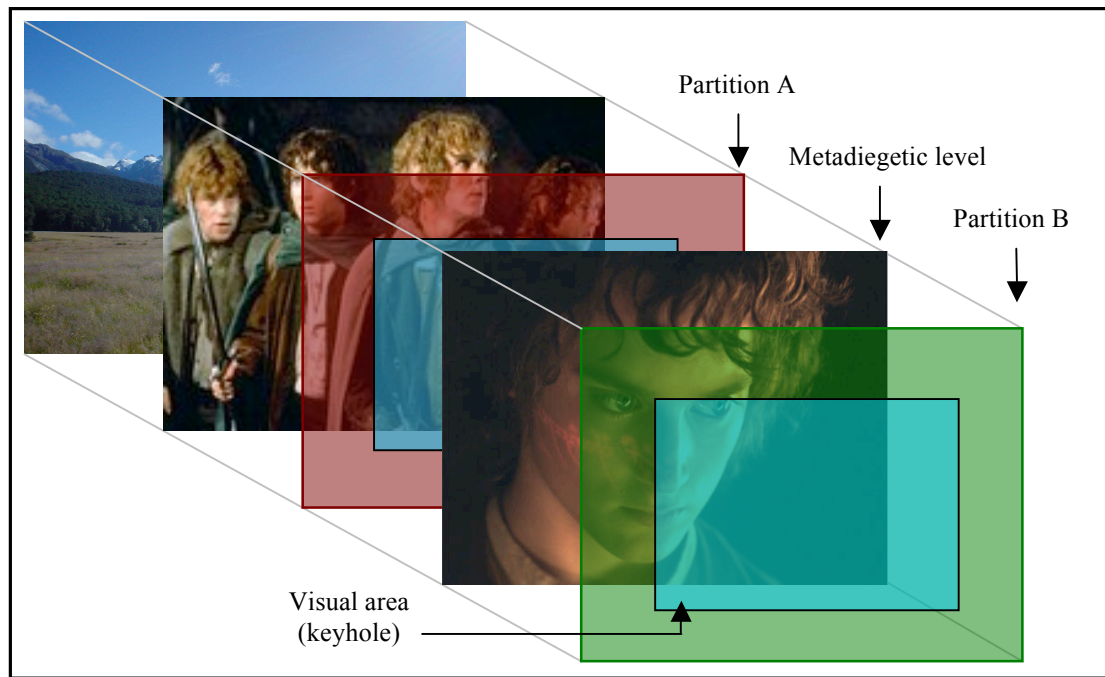


Figure 4: The metadiegetic world of the character from Peter Jackson’s *Lord of the Rings*

Figure 4 describes the metadiegetic level in the filmic experience. As seen above, Frodo in Peter Jackson’s *Lord of the rings (2001-2003)*, played by Elijah Wood, illustrates the internal susurrations of Sauron’s voice. The ring only spoke to the one bearing the ring and sometimes taunting someone around it. This level is placed between partitions A and B, which is still part of the filmic world (unlike the non-diegetic level).

### 2.2.3.1 Semiotics: *The Semiosis of sound*

A deeper understanding of metadiegetic sound flows from relying on philosophical and psychological concepts. Both are essentially *deictic*<sup>11</sup> expressions, the utilisation of which requires degrees of insight into a particular context in which they are used. Reversing the audience’s hearing process eliminates confusion when expressing an on-screen character’s memory in sound, ensuring that the latter is not obscured. In order to effectively do so, the memory must be reattached to imagery associated with real-life sounds, implying reciprocity between metadiegetics and semiotics. The fact

---

<sup>11</sup> “*Linguistics* Of or relating to a word, the determination of whose referent is dependent on the context in which it is said or written” (“Deictic”, *The American Heritage® Dictionary of the English Language, Fourth Edition*. Retrieved June 13, 2007, from Dictionary.com website: <http://dictionary.reference.com/browse/deictic>).

being, as explained earlier, that metadiegetic sounds are the sounds that are internal and therefore inaudible except to the character from which they originate.

At this point, it is also helpful to remember that, broadly defined, semiotics concerns the philosophical study of signs and symbols<sup>12</sup> and, as explained by Pierce (1931-1958: 335), it is “the doctrine of the essential nature and fundamental varieties of possible semiosis”. Therefore, it is abstract in its very nature. However, it is perhaps unavoidable that semiotic imagery is relied on as an aid in describing audience’s dreamscape<sup>13</sup> manipulation. As Gee (2003: 16) puts it:

”Semiotic” here is just a fancy way of saying we want to talk about all sorts of different things that can take on meaning, such as images, sounds, gestures, movements, graphs, diagrams, equations, objects...All these things are signs (symbols, representations, whatever term you want to use) that “stand for” (take on) different meanings in different situations, contexts, practices, cultures, and historical periods.

He refers to a “semiotic domain”, which is “any set of practices that recruits one or more modalities...to communicate distinctive types of meanings” (2003: 18).

In this context, Pierce (1931-1958: 332) refers to it as semiosis<sup>14</sup>. He explains:

It is important to understand what I mean by *semiosis*. All dynamical action, or action of brute force, physical or psychical, either takes place between two subjects [whether they react equally upon each other, or one is agent and the other patient, entirely or partially] or at any rate is a resultant of such actions between pairs. But by “semiosis” I mean, on the contrary, an action, or influence, which is, or involves, a cooperation of *three* subjects, such as a sign, its object, and its interpretant, this tri-relative influence not being in any way resolvable into actions between pairs.

A typical sentence containing semiosis would be: “*When indexing a webpage, the Meta tag in the heading of the HTML code, does not really do anything when a Googlebot crawls the page*”.

---

<sup>12</sup> “semiotic, *a.* (and *n.*)” The Oxford English Dictionary. 2nd ed. 1989. OED Online. Oxford University Press. 18 May. 2007 <<http://dictionary.oed.com.ez.sun.ac.za/cgi/entry/50219524>>.

<sup>13</sup> A literary description of a dream. This is an attempt to measure out boundaries in how far one can push abstract sounds (target audience).

<sup>14</sup> Thibault (2004:6) also refers to semiosis as ‘meaning-making’

When referring to webpage design, a certain language is used and not all people will understand this jargon. Each word is a symbol or a sign relating to something (e.g. *Meta tags* refers to keywords inserted in the webpage, so that when typing those words in Google's search bar, the website would be displayed. *HTML code* is the language in which the webpage is programmed. *Googlebot* refers to the programme created by Google, whose sole purpose is to extract information from websites across the globe. *Crawling a web page* refers to the action when a Googlebot, or other bots, are busy reading a page with the aim to gather information from it). If one does not understand the signs and symbols, then the process of indexing a webpage in Google cannot be understood. Signs and symbols are one with the context in which it is used.

A good example of semiotic codes within the narrative would be *Navy NCIS (2003 - 2008)*, directed by P. Belisario. Within this series, a lot of military jargon is used without it being explained. For example "local leo's", "scuttlebutt", "SOP", "UA", etc.



Figure 5: Navy NCIS: Naval Criminal Investigative Service (2003 - 2008)

The audience is left to figure out the meanings of the jargon for themselves. Some words might be introduced early in the first series and never explained again throughout the rest of the four seasons. This is the way sound in film is intended to work.

When creating a film with sounds loaded with signs and symbols, how would one introduce unknown signs and symbols of metadiegetic sound, so that it holds meaning to the viewer? One of the problems that may arise, is the fact that memory is not sound, it is a type of representation of what is real and is now gone. Therefore, when trying to audibly express the inaudible (metadiegesis), the listener should reconnect the inaudible sound with a symbol from the real world. The consequence for not

connecting metadiegetic sounds to an image sign or symbol is that it will then be confused with non-diegetic sounds, which will be described later in this chapter on page 27.

To describe other sound elements within film, the researcher will refer to a chart illustrated by Raskin (1992: 2):

<p><b>DIEGETIC SOUND</b></p> <p>Sound whose apparent source is known [<i>sic</i>] or presumed by the viewer to be located in the fictional situation at hand</p> <p>or</p> <p>in the mind of a character</p>	<p><b>ACTUAL SOUND</b></p> <p>Sound which, according to the [<i>sic</i>] premises of the fiction, anyone present in the situation at hand would be able to hear</p>	<p><b>ON-SCREEN</b></p> <p>Sound whose apparent source is presumed to be visible on screen at the time it is heard</p>
		<p><b>OFF-SCREEN</b></p> <p>Sound stemming from a source not visible at the moment, but understood to be a part of the situation at hand.</p>
		<p><b>TRANSITIONAL</b></p> <p>Sound bridging two scenes or two registers within a sequence, and which can be subdivided into: prolonged [<i>sic</i>], anticipatory, intercut and pivotal forms of overlapping.</p>

		<p><b>DIFFERENTIALLY AUDIBLE</b></p> <p>Sound which is objectively present, but presented as it can only be heard by one character in the situation at hand because of some special circumstance (not involving subjective distortion)</p>
	<p><b>SUBJECTIVE SOUND</b></p> <p>Sound occurring in the mind of a designated character, and which other [<i>sic</i>] characters present would not normally be able to hear</p>	<p><b>INNER VOICE</b></p> <p><b>REMEMBERED SOUND</b></p> <p><b>IMAGINED SOUND</b></p> <p><b>DISTORTED SOUND</b></p> <p><b>SPOKEN WRITING</b></p>
	<p><b>NARRATIVE</b></p>	<p><b>PERSONAL NARRATIVE</b></p> <p>A retrospective narration spoken by a voice the viewer recognizes or will come to know, as belonging to one of the characters in the fiction, who--if on camera when this voice is heard--is understood to be located in a different time/space than</p>

		that of the narrative voice
<b>NON-DIEGETIC</b>	<b>VOICE</b>	<b>IMPERSONAL NARRATER</b> Narration spoken by a voice the viewer does not perceive as belonging to a character in the fiction, or to any narrator built into the fiction,
	<b>NON-DIEGETIC MUSIC</b>	sound [ <i>sic</i> ] inaudible to the characters in the fiction, and with no apparent or implied source within the story space music understood by the viewer as not stemming from any apparent source within the story space and which no character can hear

Table 1 Char illustrated by Raskin (1992: 2)

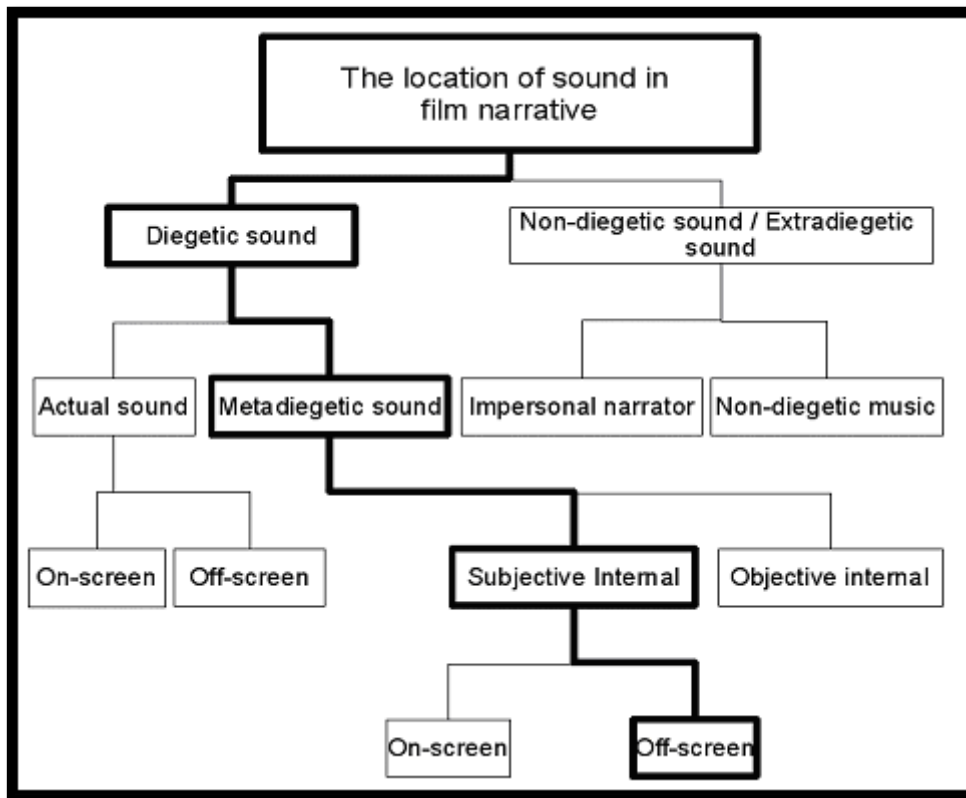


Figure 6: Narrative path

Figure 6 demonstrates the hierarchy of narrative sound in film thus far. The complete combination of all narrative sound levels is illustrated below in figure 7.

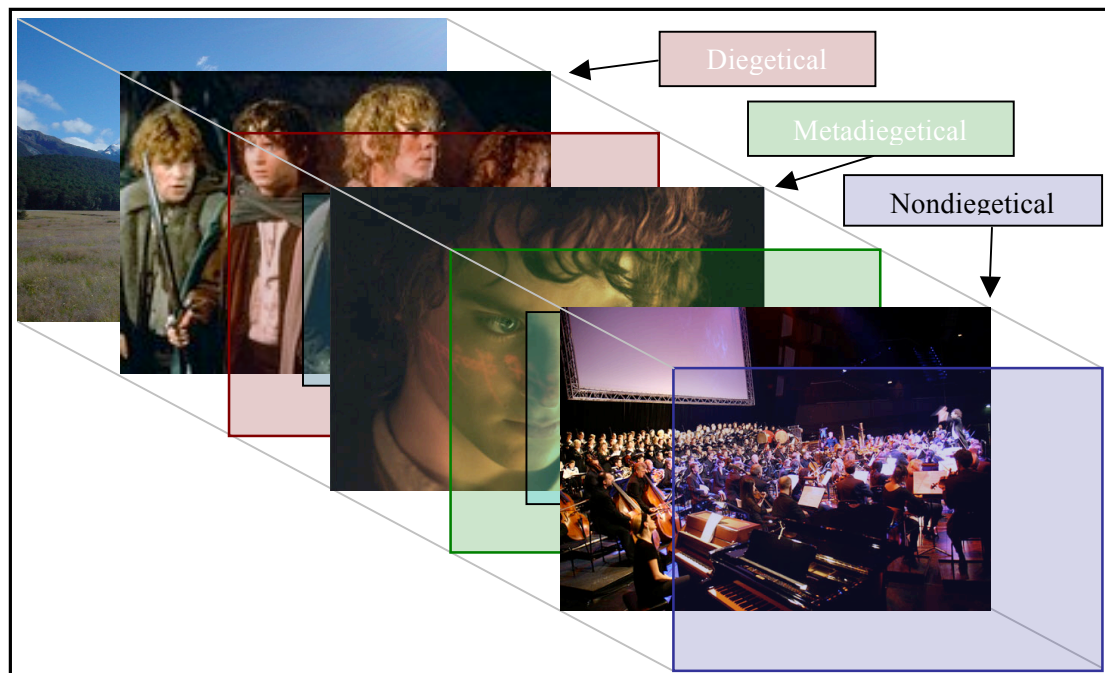


Figure 7: The combined efforts of the narrative from Peter Jackson's Lord of the Rings

### 2.2.3.2 Reality

... sound that rings true for the spectators and sound that *is* true are two very different things... If we are watching a war film or a storm at sea, what idea did most of us actually have of sounds of war or the high seas before hearing the sounds in the films? (Chion, 1994: 107).

It can be proposed that *personal reality* (things that can be touched, observed and felt) can never be *true reality*. The reason being, that when something is observed, it becomes a subjective observation and interpretation, and therefore it is only true to the individual.

The idea that, when sound is created in post-production, the sound designer does not necessarily need to use the real sound associated with the object, substitutes for the said object can be used. For example, when creating footsteps on snow in the studio, one would not drag in real snow just to record a few steps. Foley<sup>15</sup> artists use starch

<sup>15</sup> Foley is art of creating reaction sound effects real time in a studio. These sounds are created with props whilst looking at the footage of the actions (e.g. footsteps, door slams, clinks, squeaks, etc. )



instead. The sounds are similar but the question then arise, are the sounds real or are they fake?

As sound artist, one must know how to create sounds with confidence. In other words, when creating an underwater battle scene, anything can be used, because the average audience does not know what an underwater battle sounds like. The researcher proposes that when creating *metadiegetic sound*, which is not sound, almost anything can go as long as the sound image<sup>16</sup> can be recognised as metadiegetic. The reality in the metadiegetical sphere is verisimilar at best, for it originates from only one person and refers to a certain time, space and context of the observer in which it once existed or was fabricated. Therefore, it cannot be tested unless the memory took place where everyone (that is, the viewer) observed the real event. In the event, where a memory cannot be tested, the viewers have to judge for themselves if they want to believe the events unfolding in front of their eyes. Again, this refers to a willing suspension of disbelief.

A more technical way of thinking about reality in film is to ask, where should the line be drawn. Take for instance sound's frequency range - bats, dogs and other mammals can hear much higher frequencies than humans. Those sounds crossing the 20 kHz range are still real, but humans cannot hear it. It would be pointless to place sounds in a film soundtrack that no one can hear. Therefore, when creating realism for film, there are certain things that can be excluded from the soundtrack, even when they exist in reality. It can be argued that the soundtrack is the representation of the ideal selective subjective reality - just enough selective realism to tell the story in a way that does not interfere with our understanding of reality. This does not imply that one is not allowed to include sounds that are not part of the filmic world. An example of this would be a character walking through a desolated neighbourhood at night; one would hear crickets, maybe a distant highway, or even an owl. If the sound designer placed a car crash sound in the distance (as part of the background ambience), the audience would probably hear it and start incorporating it into the filmic world, even though it has nothing to do with the story. This can be useful, but might also be confusing at times.

---

<sup>16</sup> Sound image will refer to a sign or symbol of sound.

At this point, it is unavoidable to introduce the difference between *hearing* and *listening*. Unlike the eyes, the human ear can never be closed. It is always hearing something. However, to understand something one is hearing, one has to focus on (listen to) it. Sonnenschein (2001: 77) Writes:

“Hearing is passive and listening is active. While hearing involves receiving auditory information through the ears, listening relies on the capacity to filter, selectively focus, remember, and respond to sound”.

For example, when one is walking on a beach with someone, deep in conversation, one would probably not listen to the waves or the seagull flying over one’s head; one might hear the waves or the seagull, but at that moment, the person is selectively listening to the conversation unfolding in front of him. In other words, listening to something, is highly selective while hearing is a passive ability.

With this in mind, applying *hearing* and *listening* to film, is a very complicated procedure. Using the example of the character walking in the street... Forcing the audience to adopt the selective listening of the character walking in the street, the sound designer has to limit the amount of passive and irrelevant sounds in the background, so that, when a sound suddenly appears, the audience automatically *listens* to the sound instead of only hearing it.

To further expand this interpretation of filmic reality, it might be appropriate to make use of the concepts *ground sounds* and *figure sounds* (Sonnenschein 2001: 80). Ground sounds or noise would be all the sounds that would not force someone to want to listen to a sound or sounds that is only meant to be heard. Figure sounds would then refer to the contrary, any sound forcing someone to focus on a specific sound or sounds.

Incidentally, ground sounds do not necessarily refer to everyday sounds. When a sound designer has to create ground sounds for a science fiction film, the ground sounds are not necessarily something one is used to, or know that it exists, so the sound designer has to create the ground sounds in such a way that it would not interfere with the story’s *figure sounds*. It goes without saying that these sounds depend on the audience’s willingness to let go of their everyday beliefs and understanding so that the sound designer’s aesthetic creation can be believed. The

willing suspension of disbelief and all of its uses as well as its interpretations, will be discussed in depth within Chapter 3 on page 34.

According to Metz & Guzzetti (1976: 76):

The spectator lets himself be carried away-perhaps deceived, for the space of a second-by the anagogic powers belonging to a diegetic film, and he begins to act; but it is precisely this action that awakens him, pulls him back from his brief lapse into a kind of sleep, where the action had its root, and ends up by restoring the distance between the film and him. It accomplishes this to the extent that it develops into a behavior of approval: approval of the spectacle as such and not necessarily of its quality, still less of all its diegetic features; approval brought from without to an imaginary tale by a person performing real actions to this purpose.

To conclude this section, because reality is subjective, believing that something exists would automatically be true and real for the individual. Therefore, if the audience believed (through a willing suspension of disbelief) that Luke Skywalker really was a Jedi, and that the technology existed when channelling energy through a handle will produce a laser sword, then, within that *frame of reference*, any sound can be real, if the audience can connect it to something in the scene. Sonnenschein (2001: 61) writes:

No matter how strange or out of context a sound may be, our minds tend to look for something to recognise. This happens when we are looking at a natural rock wall and end up finding the form of a face.

### 2.3 Ambiguity

As demonstrated previously (chapter 2, 2.2.1.1, chapter 2.2.2, and chapter 2.2.3.1) there exist a crossover between diegesis, nondiegesis and metadiegesis. This crossover exists when the boundaries of one of the levels are not accurately defined. Gorbman (1980: 197) gave an example of a certain scene:

Early in a film we witness the great romance of protagonist X, which ends tragically during the war. Years later, while X and his best friend Y sit in a bleak café discussing their irretrievable joys, Y brings up the name of X's lost love. This strikes a chord: a change comes over X's face, and music swells onto the soundtrack, the melody that had played early in the film on the night X had met her.

Here there exists a problem - her example of metadiegesis can also be named *extradiegetic* or *non-diegetic* because this music is not coming from any orchestra in the café. In other words, the music does not stem from any instrument that exists in the time and space of the café where they are talking. This problem creates confusion between the two definitions. Metadiegetic sound refers to the sound that happens internally (audible only to the character and the audience), whilst extradiegetic or non-diegetic sound cannot be heard from any character within the narrative realm. Only the audience can hear the music. Therefore, there exists a crossover where the audience struggles to understand the origins of the music. If they know what the character is thinking, it is *metadiegetic*; and if they do not know, the music would be *extradiegetic* or *non-diegetic*.

This problem is not an isolated one, it also applies to diegetic and metadiegetic sound. Creating a sound that is not placeable in relation to any object or element in the film can have a myriad uses and holds much potential especially because of its isomorphic qualities. However, it can also have quite disastrous consequences if used incorrectly or unconvincingly.

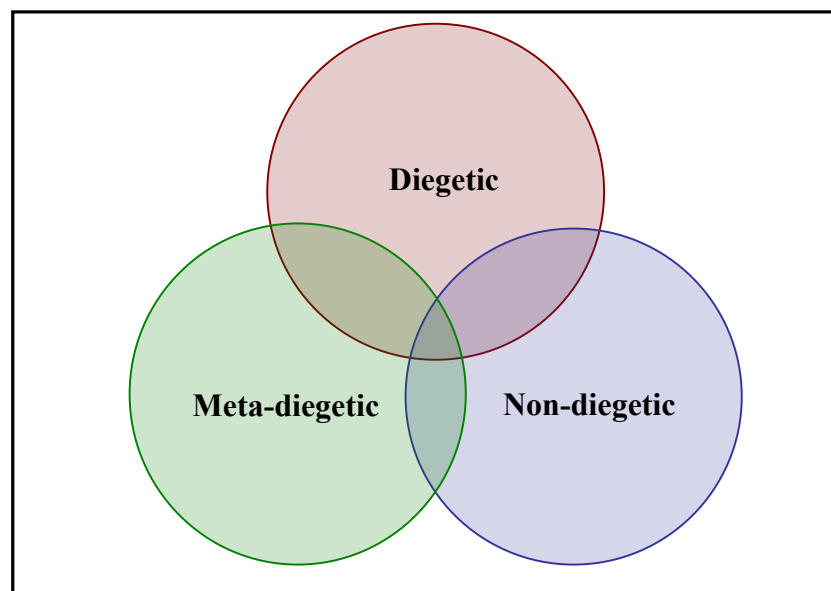


Figure 8: Crossover between narrative sound elements

## 2.4 Conclusion

As explained in the beginning of this chapter, it is the researcher's objective to introduce specific concepts and ideas, so that the reader would understand the basis of this thesis whereupon everything else will be built.

The researcher started with a brief explanation of narrative and focussed more specifically on the narrative of film sound. The main filmic sound elements that were focussed on were diegesis, non-diegesis and metadiegesis. Within these elements, it became necessary to elaborate on certain concepts to better explain the deictic nature of the three narrative sound definitions. This forced the researcher to introduce a few highly debated philosophical concepts - aesthetics, reality and semiotic theory.

The next chapter will make use of the concepts described previously, to further delve into the subconscious state of the filmic character's mind. To do so, the researcher will discuss how sounds are processed from hearing to memory and then used in relation to film.

## Chapter 3: The Sound of Memory

---

I heard the carts going past the garden fence, and sometimes I could see them too, through the shifting gaps between the leaves. How the wooden spokes and axles creaked in that hot summer! Labourers came home laughing from the fields – laughing scandalously. (Kafka, 2007: 3)

Everyday the human ear is bombarded with a myriad of complex sounds that the mind in turn has to decipher, so that knowledge can be gathered from it. This process is rendered without knowing it, in fact, one do not even hear a sound with a high pitch, resonating against a hard surface, one hears a wagon driving on a cobblestone street. The mind creates meaningful connections to match sounds on and the sound becomes the thing it represents.

Within this chapter, the researcher intends to answer the following questions: Firstly, what is the relationship between sound and memory? Secondly, how does one learn new sounds? Finally, how does one handle sounds that are not familiar?

To answer these questions, it will be appropriate to start from the beginning of the hearing process.

### 3.1 From Hearing to Thinking

It might be appropriate to first explain how the ear converts sound waves into imprints the brain remembers. According to Lindsay & Norman (1973: 237), the answer is a disappointing one. They state that “very little is know[n] of the nature of auditory neural processing”. They explain that researchers’ lack of information might be because of the “lack of understanding of the analysis of auditory patterns”. Suffice it to say, when listening to human speech, it is clear that there are “specialized sound detectors for dealing with it”. Even though researchers do not know what happens between hearing the sound and its assimilation to memory (Rauschecker, 2005: 125), some things are quite clear. Anomalies such as 40% of the neurons in the auditory cortex will not respond to pure tones and will only respond to complex sounds. Another anomaly that research has shown is that, of the 60% of the neurons that fire when a pure tone is presented, they do so with different conditions. Some will fire

only when the tone is turned on, and others will only fire when the tone is turned off (Lindsay & Norman, 1973: 237, 239).

There are many theories ranging from pitch perception to perceiving loudness. All of these elements are needed to interpret sounds. Nevertheless, what happens after the sound is stored in memory and it is needed for something? This question has no easy answer. Lindsey & Norman (1973: 376) poses the question “what was Beethoven’s telephone number?” The answer to this question is obvious, Beethoven died before the telephone was invented. To answer this question, one had to draw on different information that had to be stored in the memory before it could be recalled in order for it to be answered. To compare this type of question to film, a few problems might arise. One of the objectives of a director or writer is to tell a story. He or she cannot tell the story if no one knows what they are talking about. One must first have the knowledge of who Beethoven was, and know when he died.

With this example, the researcher can give a glimpse of the *problem solving* the mind has to go through to perceive its surroundings. Everybody knows that when a glass falls, it creates a crashing sound of complex, hi-pitched frequencies. Everybody knows if that glass had water in it, there would be a sloshing sound mixed in the accident. Depending on the sequence of events, the story can be told just by listening to the actions. Therefore, past experiences have to be drawn on to create new sounds. Because sound in film is contextual, every sound is based in something. Truax (2001: 20) used an example from the Vancouver soundscape project:

You can even tell the difference between the various delivery carts just by listening to the horses. The iceman had a couple of very heavy cobs drawing his cart, and the coalman had a pair of substantial Percherons which always walked. But Drysdale’s, a drygoods store... had a light cart and a couple of beautiful lightweight horses. They would prance along at a much gayer rate. The Chinese vegetable men had funny old carts with very lazy horses, which would just clob, clob, clob along. (World Soundscape Project, 1978:18, cited in Truax, 2001:19-20)

The world, as it is known, is full of sound. It is sometimes very hard to put sounds into words, describing something one has heard. Most of the time, one does not even hear sounds, one would only hear events. In other words, people will not hear a dull

sound resonating against a dull surface, scraping and grinding, they will hear a cart's wheel on a dusty road. The sound is connected to an event from memory.

The example given by Truax is completely a contextual citation. He argues that all of the sounds described in the excerpt were associated with sounds that would have been heard by people living in that context. He states that a person waiting for a delivery would listen for one of these sounds and distinguish between the different sounds the carts made, until they heard the one they were listening for. Taking the perceiver from the scene into the audience, it is a very different matter. The question, *will the perceiver still know the difference between the types of delivery carts if he was not previously introduced to it?* The answer is, not necessarily. The perceiver must sometimes be introduced or taught certain things before sound can be used to tell a story. This is only the case where the sound might not be familiar to the perceiver. In other instances, such as the portrayal of everyday life and in the context of the perceiver, everything is already processed. To put it in another way, listening to sounds, one will hear the sounds as a meaningful whole. In a room full of people talking, one can focus on a specific conversation whereby the "...speech is heard as meaningful words, not as a jumble of sounds that must somehow then be translated to make sense" (Lindsay & Norman, 1973: 1).

To describe how new sound is learned, it may be appropriate to introduce the perception of sound through two different theories. Gaver (1988: 14) explains that there are two theories of perception, namely cognitive perception and ecological perception. He argues that cognitive perception entails that, when a sound enters the ear, the mind starts breaking the sound down to manageable clumps of data. It is then attached to meaningful units of previously perceived data. The meaningful data is then reassembled into a new meaningful unit that represents the sound as it was heard ("what I heard"). This new meaningful unit is now added in the memory to later recall when needed or used to recognise another unknown sound.

Ecological perception on the other hand, entails that, when a sound is heard, the source must also be in the area. Therefore, something must exist that is creating the sound. The other senses are called upon to find the source (and information about the source) so that it can be identified.



Therefore, everything that is known, is stored as a meaningful whole that can be recalled when it is needed Truax (2001: 19) states:

A theoretical understanding of listening is greatly complemented by examining what listeners have to say, particularly about experiences in the past... The way in which sounds are stored in memory, not separately, but in association with their original context, betrays a fundamental aspect of the listening process.

The sound of memory is clearly a deluded statement. However, when it comes to film, the sound of memory has a lot to do with what one subjectively *knows*, is real. In other words, recalling a memory of wind blowing through dry leaves on a tree is not sound itself; it comes from an experience that was previously perceived. Even things described by someone else, will still have to be interpreted by one's own mind. As a result, all reality and truth is subjective in nature. The researcher is aware that reality and truth are subjects still open to debate by many philosophical theorists. It is not the researcher's intension to state that this is the final truth of the subject, however, for the purpose of this research endeavour, the researcher has to use the current state of the philosophical debate as the starting point for an explanation into the dreamscape of the mind.

The opportune part of subjective reality is when the source is not known, and one only hears the sound. As previously described by Gaver's (1988: 44) explanation of cognitive perception theory, the mind tends to look for patterns to match something previously experienced. That is why one sometimes hear something and realise it is something entirely different (for example, hearing shots fired in the distance, later realising that it was only fireworks). The question at this stage is, is the sound still the same as the memory of the previously heard sound, or is it something new?

Recreating reality in film with the use of sound is somewhat a flawed statement in itself, for the simple reason that all the sound in a film has been enriched or created to some extent. Even the sound recorded during production, has been enriched by using dampening and different angles during the production. These techniques are used to submerge the audience into the filmic world using *willing suspension of disbelief*.

As explained in Chapter 2, the willing suspension of disbelief is inexorable when it comes to science fiction films. Using willing suspension of disbelief to justify futuristic, unexplainable or unheard sounds, it still begs the question why one is

willing to believe these sounds and further more, whether there is a reason that some of these *created* sounds, sound *fake*. This statement is a semantically null sentence because how can a sound that does not exist, sound fake? What does it mean that the spaceship sounds real when no one has ever heard a spaceship fly or crash?

As previously indicated in Chapter 1, it is not necessary for the listener to have heard a sound previously, before knowing what it is. If one understands this statement, the walls are torn down that is encaging the limits of filmic sound. One of the problems that might arise from this statement would be the audience's past experience with film sounds. The researcher stated that the mind tends to look for patterns in the new sound to match it to previous experiences. *What if the previous experiences, were something heard in a film?* The result would be that the memory of the remembered sound is erred. Every person watching film is in a process of being conditioned. Most people would know that a shotgun sounds very different from a normal handgun without having heard them being fired in real life. Might one argue that through years of conditioning by watching science fiction films, one has come to terms that science fiction space battles must sound a certain way? Is there a way to change this forced limit to sound?

All of these questions might inherently bring up more questions and not many answers. Suffice it to say that the concept of willing suspension of disbelief is probably as old as fiction itself and many theorists such as Jean Mitry, Jean-Louis Baudry, Christian Metz, Laura Mulvey, Noël Burch and Mary Ann Doane, have written and discussed the subject at length so that some theories will help some of the above questions to dissolve. However, there will probably never be a final answer to the age-old problem of *willing suspension of disbelief*, and why one is willing to submerge oneself into a fake reality. It is not the researcher's intension to enter in an in-depth discussion of the timeline of *illusion* and *willing suspension of disbelief*, but rather describe the main threads of the concepts and attempt to shine some light on the subject as it is revealed.

### **3.2 A willing suspension of disbelief and illusion**

The concept *willing suspension of disbelief* is generally associated with Coleridge's idea of illusion, which he calls the *dream*. Smith (1995: 113) stated in his article *Film spectatorship and the institution of fiction* in referring to Coleridge, that "in effecting

the “willing suspension of disbelief,” the “poet does not require us to be awake and believe; he solicits us only to yield ourselves to a dream””.

Coleridge (1956-71: 641) clarifies:

...Images and thoughts possess a power in and of themselves, independent of that act of the judgement or understanding by which we affirm or deny the existence of a reality correspondent to them... The forms and thoughts act merely by their own inherent power: and the strong feelings at times apparently connected with them are in point of fact bodily sensations, which are the cause or occasions of the images, not (as when we are awake) the effects of them.

Coleridge’s statement of images and thought is important in the sense that he argues that it has a “power” of “themselves”. In other words, the perceiver knows when something is not real, but because the judgement is absent in thought, it does not matter if something is real or not. The feelings one has of certain images or experiences cannot be changed. To clarify Coleridge’s statement further, Vallins (2000: 44) stressed the fact that, to say that one “*believe* in such a causative relation, however, is to Coleridge inaccurate in two aspects”. Firstly, he argues, that whilst dreaming, the will is suspended and therefore not able to judge whether something is real or not. Therefore, without the ability to compare if something is real or not, it is impossible to affirm or deny its truth or reality. Secondly, for Coleridge, “there is no separation between feelings and thoughts or images” in dreams. It is different to “waking consciousness” for the simple reason that one is able to “judge the reality of our circumstances and feel certain emotions as a result of our judgment”.

Because, in dreams, one does not have the ability to judge if something is real or not, is it not better to say that when watching a film, it is as if one is going to sleep? The observer is entering a dream-like experience where he is no longer permitted to have the ability to judge what is real or not.

Brinker (1977: 193) writes

Everything takes place as if the spectator is aware that there can be no interference on his part, and as if he had consented to be satisfied with an emotional reaction. From a spectator’s reaction it is impossible to infer either belief or disbelief in the reality of what happens on the stage.

Even though the above-mentioned argument of Coleridge is valid to an extent, Price & Braithwaite (1964: 154, 155) reasons that for some, a willing suspension of disbelief is not enough. They state that

For them it is only the first stage, and they pass through it into a more positive attitude: not quite an attitude of belief, but something near it. In their case, there is indeed a voluntary suspension of disbelief when they open a novel... But afterwards, when they have become absorbed, as we say, in the novel or play, the state they are in is not merely an absence of disbelief, but something more and something more positive

Price & Braithwaite calls this kind of state *not quite belief, half-belief*. When someone watches a film, they enter into Coleridge's dream state and know that they are not able to change the flow of the narrative in any way as Brinker described. However, after the movie ends, they would still accept that the experience is make-belief and even though, as observer, they *half-believed* it for the time being, they would not have to sit there and except everything as shown. They are allowed to *wakeup* from this dream.

With all this in mind, Vallins (2000: 44) explained that Coleridge expanded on his original idea of the dream. He added further differentiation between "feelings and images of waking consciousness from those which occur in dreams". He explains that when having a nightmare, it is not only terrifying because one "cannot exert influence over the terrors" but rather because the feeling of terror is not true terror that might indicate the presence of danger, but to some degree the pain of a tooth-ache (Coleridge, 1957). It is not the pain itself but more the sensation arising from it. Therefore, one might argue that when dreaming about something terrifying, the mind tends to find a similar sensation and connects it to the image. This might also be one of the reasons why film spectators submerge themselves and apply their own sensations to the given images. They know it is not real, but during this *dream*, they also *know* what it must feel like to be that character or to be in a similar situation (without ever being in such a situation).

Memory and thought have always eluded people. Most of the theories of thought and the representation of memory in film are strongly connected to emotion and feelings Chion (1994: 109) explains:

The film spectator recognises sounds to be truthful, effective, and fitting not so much if they reproduce what would be heard in the same situation in reality, but if they render (convey, express) the feelings associated with the situation.

His argument is that the audience does not directly compare sounds to an event in real life, but rather to the feeling connected with it. A good example of such an emotional connection would be *Perfume (2006)*, directed by Tom Tykwer. During the scene depicting Jean-Baptiste's (Ben Whishaw) birth, his mother gives birth to him under her fish stand. She cuts the umbilical cord with the same knife she guts fish and pushes the newborn baby with her foot aside, gets up and go back to work. Disgusting images of fish guts, larvae, dogs gnawing on animal carcasses etc., also enforced this barbaric action of his mother. All the while, the sound of the shown images is also bombarding the viewer. The images are flashed one after the other creating a sensory overload with gruesome images. The effect is that one feels repulsed and sickened by the action. The feeling when one observes the inside of a rotting animal or rotting fish guts, not to mention the smell, would be overwhelming and the viewer would assimilate these feelings with the mother's act of rejecting her baby.





Figure 9: Perfume (2006), directed by Tom Tykwer

Another situation, later on, illustrates the way a perfume John Baptist made, smelled. The perfumer Baldini (Dustin Hoffman) opened the lid of the jar containing the perfume standing in his shop's basement. The moment he smelled the sweet scent, he was immediately transported to a beautiful green garden with pillars full of flowers and children laughing in the background. A sensation of peace and tranquillity is created by the sound. To cap it off, a beautiful woman enters the frame and tells him that she loves him and kisses him on the cheek. The illustration of what this perfume smelled like was described by using common feelings most of the audience had. Images of a beautiful garden with colourful flowers, accompanied by sounds of birds, a light breeze and children's laughter in the background created this feeling of peace and tranquillity. All this with a woman (associated with a lover) telling him that she loves him and kissing him on the cheek, creates that special feeling one gets when one is in love.

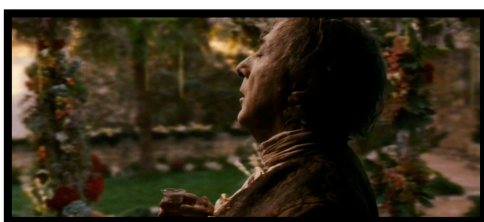




Figure 10: *Perfume* (2006), directed by Tom Tykwer

It is clear that *feeling* in this example was not just stressed by sound and image alone. This is why Coleridge’s concept of the image-feeling connection is much more complex than just recalling a feeling associated with a certain image. Chion (1994: 112) states, “...in reality rendering involves perceptions that belong to no sensory channel in particular”. He uses an example written by Leonardo da Vinci in one of his notebooks: “If a man jumps on the points of his feet, his weight does not make any sound”<sup>17</sup>. Chion continued, “Leonardo da Vinci marvelled that sound does not render the fall of a human body, he was thinking not only about the body’s weight but also its mass as well as the sensation of falling, the jolt it causes to the person falling, and so forth”. He argues that da Vinci was not experiencing something that can “be reduced to one simple sensory message”. In film, this is why most of the action sequences are over exaggerated, ensuing a body fall to sound heavy, maybe even emphasise a few cracking or dislocating bones to indicate pain. This is obviously contradictory to real-life events. Take *In Bruges* (2008), directed by Martin McDonag, as an example. The character Ken (Brendan Gleeson) stood on top of a tower and jumped off. His body crashed on the cobblestones of the street.

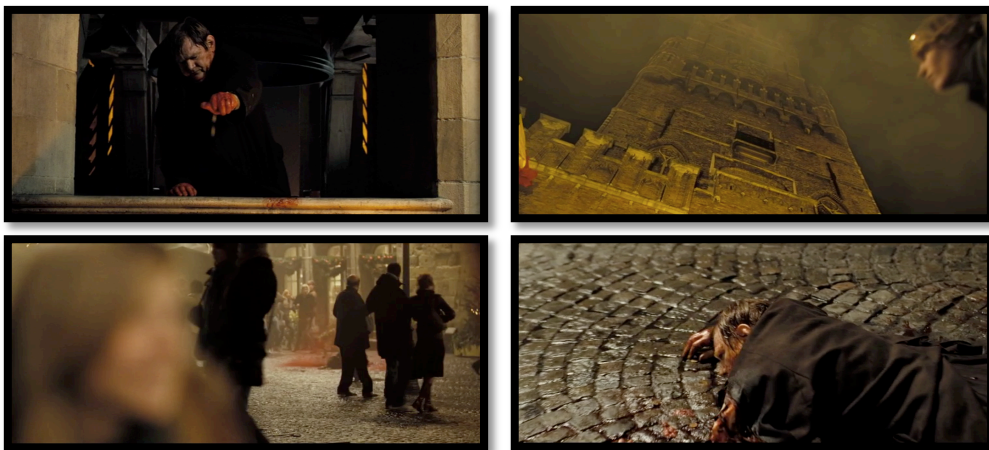


Figure 11: *In Bruges* (2008), directed by Martin McDonag.

---

<sup>17</sup> Notebooks of Leonardo Da Vinci, sec. 10, 1:280.

This painful and clearly disturbing sequence was even more exaggerated by showing the body, where the audience realised that he was not dead, though badly disfigured. The painful experience was clearly emphasised by combining different blood-freezing sound effects such as cracking bones, and something bursting (probably illustrating the body bursting from the fall).

All of these combined sounds create “clumps of agglomerated sensations” as Chion (1994: 112) puts it. Therefore, the audience might experience a jolt in their stomach as he leaps from the tower and a cringing sickening feeling when he dropped on the cobblestone street. It is clear that situations exist where the viewer has to choose their stance towards a subject. In other words, one has to choose how one is going to feel about a certain situation or occurrence. Did the falling of the man disturb the viewer or made the viewer feel angry or sad? Therefore, recalling memories of feelings is not as simple as just sitting in a chair and watching a film, sometimes the viewer is called upon, or might one even say, forced to participate. This statement clearly renders the notion that when watching a film, it is as if one goes to sleep and enters a dream, incomplete. In fact, Biro (1982: 14) ascribes this problem to the fact that the participant clearly has a *will* even though the participant is in a dream state. She refers to a Hungarian psychologist Ferenc Méri, who gave rise to the concept *allusion* and stated: “Allusion is a shortcut to giving memories a new life, relying on remnants of common sensations”. Her argument is that allusion “not only... resuscitates memories, but also that the forces of emotion and will are not separated from it”. She clarifies:

This results in an exciting play that presupposes complicity on the part of the participants, since identification and experience of the subject happen not only to the initiator of the communication but to its recipient as well.

To clarify Biro’s statement, the storyteller tells a story, but knows that he can tell it any way he wants, but it does not mean that the viewer or recipient is going to except what is told. Everything told by a storyteller, whether it is true or not, has to be interpreted by one’s own subjective reality.

In fact Burch (1982: 18), in referring to Metz, also did not agree that the audience’s filmic experience was the same as dreaming. He posits:



...belief in the cinematic images as an analogue of real phenomena, if it ever was an hallucination (such as might be induced by drugs or psychosis, for example), has long ceased to be one; it is, indeed, a *willing* suspension of disbelief, an emotional involvement which may certainly attain great depths of anguish or compassion, but which is always grounded in the awareness that the subject is 'only watching a film'. It is in this respect, Metz further suggests, that the filmic experience resembles that of the phantasy rather than the dream.

There is also an alternative - Link (2004: 77) argues "...taking a seat in a movie theatre is to take an alternate subjectivity... Cinematic spectatorship becomes a modulation of identity. By being its audience the perceiver can paradoxically cease to be its audience and become its subject". What he is implying is that the film metaphorically becomes the mind and the audience takes on the subjective reality, as if already justified by the processes of the metaphorical *mind* through certain mechanics applied in such a way that the viewer "...ceases to be its audience and becomes its subject". He reasons that if the film is the mind, then the "cinematic body" cannot be ignored. Link referred to Sobchack (1995: 37) who described the reason for this metaphor:

...more than any other medium of human communication, the moving picture makes itself sensuously and sensibly manifest as the expression of experience by experience. A film is an act of seeing that makes itself seen, an act of hearing that makes itself heard, an act of physical and reflective movement that makes itself reflexively felt and understood. Objectively projected, visibly and audibly expressed before us, the film's activity of seeing, hearing, and moving signifies in a pervasive, primary, and embodied language that precedes and provides the grounds for the secondary significations of a more discrete, systematic, less "wild" communication.

This statement clearly materialises a concept of a body seeing and hearing things to some extent just as a normal person would, but with the extra omnipresence added. The point to this statement is to remind the reader that creating something to evoke certain feelings or sensations is incomplete - it is much more than that. The idea that feelings and sensations can be disconnected from the audiovisual body, that is to say, all the mechanics involved, to help tell a story in the way the writer or director intended the subjective reality had to be perceived. This will further be discussed at length in chapter 5.2 The 'External Story Space on page 66.

There are also theorists that justify the topic of willing suspension of disbelief through other relevant theories. One of these theorists, Mulvey (2006: 33), argues:

Illusions of the supernatural typically brings into play peculiarities of both human vision and the human mind. A mind bewildered by optical and other kinds of illusions, doubting the reality of what it sees with its own eyes, is more prepared to be credulous when exposed to the emanations of the supernatural. An otherwise confident and competent relation to the world is suddenly faced by a sense of uncertainty. Whatever ideology or commercial enterprise might fuel these phenomena, they share an aesthetic of deception, an appeal to the human mind's pleasure in illusion and its constant readiness to be fooled. The machines of deception necessarily had a complex relation to the different ideologies within which they functioned, veering between complicity with irrational belief and its debunking.

A good example of this would be *The Devils Backbone*, (2001)<sup>18</sup> directed by Guillermo del Toro. Applying Mulvey's statement to the film, one is willing to believe the story unfolding in front of one's eyes because the irrational fear or subconscious belief in some, that there might be an afterlife, and that there might be spirits moving around them. Because this reasoning cannot be proven or disproven, some are willing to let go of their doubts of the existence of the supernatural, whereupon the viewer might start feeling scared or uncomfortable.



Figure 12: The devil's backbone (2001), directed by Guillermo del Toro

---

<sup>18</sup> This story is about an old orphanage in the time of the Spanish war. In the middle of the orphanage's courtyard was an unexploded bomb, dropped there by a plane during an air raid. The film genre was established when the lead character, Carlos, placed his ear on the unexploded shell and hearing strange ethereal ticking, voice-like chant sounds. One started to realise that there might be something strange about this place. The story progresses to the point one realise that the orphanage is haunted.

In other words, when one is bombarded by illusions, one is likely to be naïve towards the given illusions. This, however, only pertains to the visual side of human perception. According to Lindsay & Norman (1973: 33-34), “the image on the eye is in constant motion”. This is not because the environment change while walking around or watching a film, it is because there is a constant tremor in the eye. This tremor is called “physiological nystagmus” and it is this movement that forces the human mind to see lines and shapes. They explain that when all movement in the eye is stopped, the eye will no longer see the image in front of it. As long as the eye moves (through the process of physiological nystagmus amongst others) the complex feature detectors will continue to respond. This anomaly however, does not apply to the auditory processes. They ask the question: “Does a constant auditory tone become inaudible after a while?” (Lindsay & Norman, 1973: 36) The answer actually lies in the physics of the sound wave. Sound is in constant motion, it is a continuous change in air pressure over time, and therefore “a constant auditory signal does not exist”.

The argument here lies with the combination of audio and visuals where the true illusion lies. When the eye is erred, the mind is in doubt and the sound can build on this error. This is not to say that all films are optical illusions, but rather that the film medium draws upon different psychological and physiological processes present in all human beings to help create a *fake reality*.

This in turn helps to submerge the audience of a supernatural occurrence, to perceive it as a credulous situation. Mulvey (2006: 34) states:

Whatever ideology or commercial enterprise might fuel these phenomena, they share an aesthetic of deception, an appeal to the human mind’s pleasure in illusion and its constant readiness to be fooled. The machines of deception within which they functioned, veering between complicity with irrational belief and its debunking... Freud’s great contribution to modernity was to recognize that the irrational was intrinsic to human reason, ‘housed’ in the unconscious.

The last part of this statement may hold the key why one sometimes feels uneasy when watching a horror film depicting the supernatural. Freud (2003: 148) argued that religion feeds the instance of the afterlife and in turn, the return of the dead. Might one argue that if something supernatural exists, one might feel helpless against such an apparition? This question, however, does not have an answer, for the simple reason that everyone’s belief systems are different. Nevertheless, can one argue that filmic

history introduced and exploited the human's opinion of the afterlife? Mulvey (2006: 34) argues that the cinema has gathered all the illusions of its prehistory and formed it by means of what she calls "natural magic". This natural magic is what gives films of modern day, the edge in playing out the subconscious and the imaginary.

This argues the fact that, as Freud puts it, "one fear could summon up the other, perpetuating the power of the irrational over the rational"

All of these theories make use of the subjective reality of the observer or audience. It draws upon the mind's list of previously experienced sensations and fears, or might one rather say, conditioned sensations and fears. Every time one watches a film, the mind learns new connections between feeling and memory images. Therefore, at a later stage, the mind calls upon sensation-images that were experienced within a film. It does not matter if the image or sensation was real or not. The fact is that each viewer is in the process of being conditioned. Chion (1994: 108), in his explanation of reality, states:

...when the spectator hears a so-called realistic sound, he is not in the position to compare it with the real sound he might hear if he were standing in that actual place. Rather, in order to judge its "truth," the spectator refers to his memory of this type of sound, a memory resynthesized from data that are not solely acoustical, and that is itself influenced by film.

The point of this abstract is not to discuss his views on reality but rather the fact that he states that "memory resynthesized from data that are not solely acoustical... is itself influenced by film". In other words, Chion describes a type of new culture with a catalogue of conditioned sound and images that help them interpret a certain scene. Take for example the film *Delicatessen* (1991), Directed by Jean-Pierre Jeunet. There is a character (Aurore) within the film that wishes to kill herself because she is hearing voices. Later in the film, the audience realizes that it is a man talking to her through the air vents.



Figure 13: *Delicatessen* (1991), directed by Jean-Pierre Jeunet and Marc Caro

This concept in the film presupposes that the audience knew that the voice the woman heard, was not from anywhere within the diegetical space and time, and must be in her head, even though the voice sounded as if it was in the same diegetical space as the character. The only way the audience would have made that connection would be that they had to have heard voices themselves or to have seen such an occurrence or any of the combination thereof. It also goes without saying that by the same token, it is also true for the opposite, where some would have figured out that the voice was another character creating this voice and not a ghost or the concept of a mental illness, without having seen the revelation later in the film.

The sound serves as cues or hints to what a story's outcome will be. This in turn helps the audience to interpret the film through the given elements. Carroll (1988: 173) explains:

When following a narrative film, I want to say, a spectator internalizes the whole structure of interests depicted in the drama, and this structure includes alternative outcomes to various lines of action which the spectator must keep track of in some sense before one alternative is actualized in order for the film to be received as intelligible. I postulate that the spectator does this by tacitly projecting the range of outcomes as subconscious expectations which we can represent as questions.

One can argue that even though the audience thinks that they are interpreting the film as it unfolds, which is true to some extent. However, the interpretation is mostly the story line of a belief system that the director and writer want to adopt. Take the film *Hero* (2002) (aka *Ying Xiong*) directed by Zhang Yimou as an example: the film starts with the main character, played by Jet Li, about to meet the emperor of China. One starts to realize that he did something heroic. A few minutes into the film, one realizes that the hero killed three assassins that tried to assassinate the emperor in the

past. The reward for killing each of the assassins was gold and to be able to sit closer to the emperor.

The interesting part of the film is that the emperor wants to know how this hero killed such powerful adversaries and asks the hero to tell him the tale that led to their deaths. With each tale, the audience has the chance to either believe the story or not. The great majority of the audience will believe this tale because there is no reason to doubt the hero's tale. However, as the story progresses to the point where the emperor starts to question his erroneous tales, one realizes that he was lying and that there must be an alternative reason for lying to the emperor. With each assassin's death, the hero is allowed to sit closer to the emperor, whereupon the audience might start to realize that the hero's intention is to get close to the emperor. The audience's suspicions are confirmed when the emperor states that the stories are false and that the hero's intention was only to get close to the emperor to kill him.

Even though the audience might feel that they are in control of their own decision-making capabilities, they are still being fed certain details so that they are only focusing on specific elements. One can argue that the director forces the audience to adopt a certain strain of belief systems or a specific subjective reality that was already processed. In other words, the audience will believe what they are told.

### 3.3 Conclusion

To conclude this chapter, it is important to know why all of these theories were mentioned at all. When watching a film, the audience seldom believe it to be the truth. As Coleridge stated, within a dream one is not in a position to judge the truth of anything. He stated that within the dream, the will is suspended and without the ability to compare whether something is true or not, it is impossible to affirm or deny its truth. This statement gave birth to the notion that when watching a film, it is as if one goes to sleep. One is in no position to judge the truth of the film, and thus accepts it as true. This, however, is not the whole truth of the matter; allusion gave rise to the concept that the viewers still had a will to make their own assumptions.

It cannot be denied that, when someone watches a film, it still has to be interpreted by the audience. An example that will later be illustrated is *Memento (2000)* (see chapter 5.2 the "external story space"). This film is told in reverse, and one starts to realize that the lead character suffers from short-term memory loss. This is why the story is

told in short, reversed segments to illustrate this disability and the audience has to sit and figure out the storyline along with the lead character.

The other perspective discussed, was the film-as-mind. It does not matter what background the viewer comes from, the film has already processed all the data in such a way that the viewer just has to watch it in the way everything is revealed. This perspective is almost as if one is looking through a pair of binoculars that is attached to the ground and listening through earphones which is attached to a microphone that is focused on the area where the binoculars are pointing at. One cannot move the binoculars in any direction or focus on something else. The viewer just has to keep on looking at the scene and cannot interfere with the total composition. The only thing the viewer can do, is walk away or try to figure out what is happening by observing and listening.

The relationship between the audience and the film is very important when it comes to creating sound for film. The sound designer has to know that, when telling a story through sound, the question arises; would the audience interpret it the way he imagined, or did he create something that confused the viewer? It is clear that sound in film has a large role to play. It has the potential to sway emotions, to create a presence of a location or character. It also helps by telling the story where the visuals fail. It is also clear that when creating sound for a film, the audience cannot be disregarded from interpreting the film. Different ways of listening has different ways of interpreting data.

In addition, through these theories it can be concluded that the film draws on agglomerated sensations the audience has previously experienced or learned about through audio-visual triggers. The researcher can also conclude that, through these learned agglomerated sensations, the audience can be given abstract or surreal scenarios, and they in turn, will interpret what is observed, with the same or close to the same sensations experienced by the characters in the film. In other words, through the triggers of the audience's past experiences and the dreamscape of their minds, the sensations that cannot be connected to any sense in particular, can be drawn upon to create a third dimension of experience where one would feel or experience what the character(s) goes through.

Even though the sensations do not come from any sense in particular, the trigger (see chapter 1 on page 8) must still be set off in a way. Within the next chapter, the researcher will introduce different ways of listening that help or compel the audience to learn about the filmic environment that is presented in front of them.



## Chapter 4: The Listener

---

“Both in theory and practice, listening is the crucial interface between the individual and an environment” (Truax, 2001: 15).

In film theory, the audience is usually referred to as the viewer. This might imply that the film is foremost a visual medium. This however, was the case in the past where silent films dominated the cinema. Nevertheless, today it is an audio-visual collaboration. It is a *give and take* between the narrative that is shown by the image and the narrative that is told by the sound. As previously described in Chapter 2: Theoretical Outline, on p. 10, film can be divided into different narrative divisions, and even though they are in different categories, combined, they serve a valued purpose.

This chapter will focus on the listener and his responsibilities to the film as a viewer *and* a listener. Truax’s (2001: 18) view on one’s surroundings can be compared to film. He states that the human’s primary listening model is foremost a value of survival. This survival instinct causes someone to gather information about his or her surroundings that might threaten the body’s survival. Sudden loud noises that might point to something that can hurt the body, or a sudden sound within a certain environment, will cause someone to flinch. An example of this might be, while walking in a field and suddenly hearing a hissing sound or a rustling sound in the bushes, one would automatically retreat. It goes without saying that if one feels comfortable with one’s surroundings, these sounds will not necessarily be considered as threats, because the person might expect or anticipate sounds that are usually within a certain condition or setting.

Further in this research document, the researcher will give an in-depth explanation on how the mind distinguishes between different sounds by using the information that is gathered by the hearing mechanism. Suffice it to say, that everybody has different ways of listening to the world around them. Truax (2001: 17-18) referred to Lusseyran (1963: 22, 32) to indicate how sounds are projected and reverberated from objects around the listener. He states:

After I went blind, I could never make a motion without starting an avalanche of noise... Whenever I took a step, the floor cried or sang – I could hear it making both these sounds – and its song was passed along from one board to the next, all the way to the window, to give me the measure of the room. If I spoke out suddenly, the windowpanes, which seemed so solid in their putty frames, began to shake, very lightly of course but distinctly... I could hear the smallest recession in the wall from a distance, for it changed the whole room. Because this nook, that alcove were [sic] there, the wardrobe sang a hollower song. ... As I walked along a country road bordered by trees, I could point to each one of the trees by the road, even if they were not spaced at regular intervals. I knew whether the trees were straight and tall, carrying their branches as a body carries its head, or gathered in to thickets and partly covering the ground around them.

It is this hyperreality that helps the human body learn about its surroundings when the visuals fail to do so. At this point, it might be appropriate to introduce the different theoretical modes of listening to interpret one's surroundings.

#### 4.1 Modes of Listening

As previously indicated, Sonnenschein (2001: 77) differentiated between hearing and listening, a differentiation that cannot be ignored when discussing the use of sound in film and means of creating the correct sound to ensnare the listener's attention.

Gaver (1993: 1- 2) describes in his article "*What in the world do we hear?: An ecological approach to auditory event perception*" two types of listening and ways to understand what one hears. He differentiates between *musical listening* and *everyday listening*.

He created a hypothetical scenario by asking the reader to imagine a situation where one is walking along a road at night, when suddenly a sound is heard. He continues by stating that on the one hand, one might pay attention to the pitch and loudness and how it changes over time. He also specifies that one "might attend to the sound's timbre, whether it is rough or smooth, bright or dull; or even notice that it masks other sounds, rendering them inaudible". He calls this *musical listening*.

Furthermore, he explains that

...on the other hand, as you stand there in the road, it is likely that you will not listen to the sound itself at all. Instead, you are likely to notice that the sound is made by an automobile with a large and powerful engine. Your attention is likely to be drawn to the

fact that it is approaching quickly from behind, and you might even attend to the environment, hearing that the road you are on is actually a narrow ally, with echoing walls on each side (Gaver, 1993: 2).

This type of instance he coins *everyday listening*.

For Chion (1994: 25) however, three modes of listening exist. He writes:

When we ask someone to speak what they have heard, their answers are striking for the heterogeneity of levels of hearing to which they refer. This is because there are at least three modes of listening, each of which addresses different objects. We shall call them *causal listening*, *semantic listening*, and *reduced listening*.

#### 4.1.1 Causal Listening

Causal listening, as described by Chion (1994: 26), is any sound that is listened to with the purpose of gathering information about its source. He explains: “When the cause is visible, sound can provide supplementary information about it; for example, the sound produced by an enclosed container when you tap it indicates how full it is”. Another example of causal listening can be a mechanic listening to a car’s engine to figure out what is wrong with the car. The main purpose of causal listening is to use the sound as a tool to help determine what causes the sound. This can also include recognizing a friend’s voice amongst a crowd of people, in other words, as Chion puts it, “the sound produced by a particular unique object”. The emphasis for Chion is on the adjective *unique*, because one can learn or draw information from the sound. When a group of friends converse amongst other people, one would probably hear them and know that they are there without seeing them.

He also argues that this type of listening mode is depended upon the existing knowledge of what one knows. In other words, when a certain sound is not known, the mind will attempt to match it up to known objects or causes that is in the area (see ecological perception, chapter 3.1 from hearing to thinking). He states that humans cannot discern between different dogs within the same breed, but when one hears a dog barking in the back yard, one might associate it with the barking of one’s own dog.

There are examples of this type of listening mode in most films. However, to clearly illustrate what causal listening means, the researcher will refer to a certain scene from

*The Assassination of Jesse James by the Coward Robert Ford (2007)* that was directed by Andrew Dominik. The scene starts where a character, called Wood (Jeremy Renner), enters the house and tells everyone that Dick (Paul Schneider), who is staying at their house, philandered with his father's wife. The scene cuts to Bob (Casey Affleck), lying in his bed, awakened by Wood talking downstairs. At this point, Bob knows that Wood wants to kill Dick and starts to wake him. Both of them wait with their guns at the ready for Wood to come up the stairs and enter the room. What makes this scene interesting is that the audience knows that, even though there are at least three people downstairs, they know that it is Wood coming up the stairs. This reasoning comes down to the causal listening capability, hearing a sound and matching it to the most likely cause. In this case, Wood's voice is connected to the likelihood that he might be coming up the stairs.

It becomes apparent that causal listening is a conditioned sound that was once heard and can be recalled for future reference. This means that, when one hears a sound, the mind can file through a collection of images (or sound-images<sup>19</sup>) to identify its source<sup>20</sup>. This enables people to say, "it sounds like rain" when they hear rain while being in-doors and cannot see outside. Therefore, causal listening is contextual.

Back to the scene where a man is walking in the street when he hears a sound getting closer, realising it is a car, he jumps out of the way. Almost everyone must have seen a car at least once in their lifetime and therefore knows what a car sounds like. Therefore, it would not be too hard for the brain to assume that this kind of sound is the same as the sound a car would make (even though different types of cars have different sounds). Now, what if the sound that is heard, is not familiar to someone? Say for instance that one hears a loud creaking sound... There are a myriad of objects that can produce a creaking sound, but when one hears a sound in one's immediate

---

<sup>19</sup> Sound-image – Any sound or combination of sounds that when it is heard, can be recognized and interpreted as something previously observed. This is forged from previous experienced synchresis (*see Chion (1994:5)*).

<sup>20</sup> Most sounds comprise of more than one individual sound, and this sound in combination with image, through a process Chion (1994:5) coined synchresis, in turn creates a sound-image in ones memory that can later be recalled.

surroundings, one would automatically calculate what could possibly be making the sound.

Another example would be a certain scene from *Delicatessen* (1991). Half an hour into the film, Louison (Dominique Pinon) is in the process of fixing the butcher's bed that is squeaking. To find out which of the springs are making the noise, he and the butcher's wife sit on the bed whilst bouncing up and down.



Figure 14: *Delicatessen* (1991), directed by Jean-Pierre Jeunet and Marc Caro.

In this scene, this mode of listening is used to figure out which spring to fix. Another example would be *The Truman show* (1998), directed by Peter Weir. During a scene where Truman (Jim Carry) finds out that there is a conspiracy going on and he pretends to clear up the basement of his house, in order for him to feign falling asleep. What he actually did was, he recorded some snoring sounds on a tape-recorder, placed a blown-up clown under a blanket and escaped. The producer of the Truman show (Ed Harris) finds Truman's actions a bit strange, and tries to hear if there actually is someone in the basement, whereupon he hears the snoring. The director of the Truman show, tries to listen in order to gather information whether someone actually is in the basement.

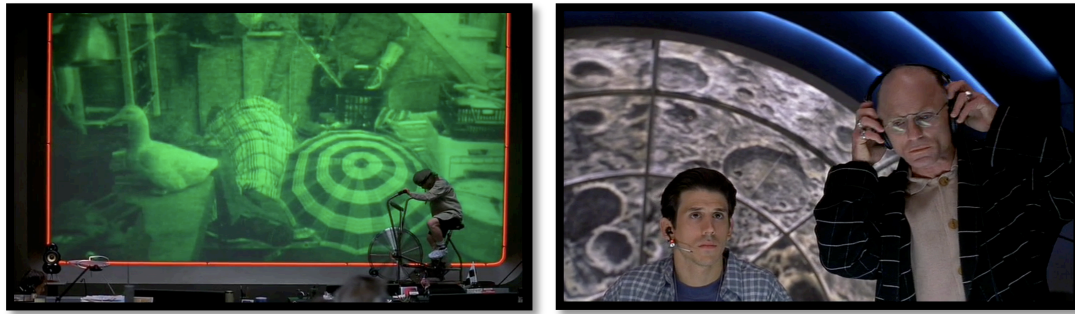


Figure 15: The Truman show (1998), directed by Peter Weir.

It is clear that this listening mode uses sound as a representation of something that one knows. It is either present in one's frame of reference or introduced during the film. Either way, this is one of the listening modes that are mostly used in film today.

#### 4.1.2 Semantic Listening

For Chion (1994: 28), one of the most widely studied and extremely complex listening modes is semantic listening. This listening mode is characterised by interpretation such as an understanding of a language, or Morse code. He explains: "One crucial finding is that it is purely differential. A Phoneme is listened to not strictly for its acoustical properties but as part of an entire system of oppositions and differences".

Applying this concept to metadiegetical understanding of sound in film, the process of interpreting and understanding different techniques of sound narrative that shapes syncretism, creates a doorway into the surreal world created by the film creators.

He stresses "semantic listening often ignores considerable differences in pronunciation (hence in sound) if they are not pertinent differences in the language in question". He continues that one can also employ both semantic and causal listening modes at the same time, listening to what is said as well as how it was said. For example, when listening to a very soft or a badly distorted voice, one is not listening to what the voice sounds like, but rather to try to figure out what is said. An example of this would be *Lost in translation* (2003), directed by Sofia Coppola. At the end of the film, Bob (Bill Murray) whispers something into Charlotte's (Scarlett Johansson) ear. The interesting part about this shot is that one can only hear a whisper, but cannot understand anything. As the perceiver, one is intrigued by this and would really like to know what is said, and automatically attempts to listen harder to hear what is said.



Figure 16: Lost in Translation (2003), directed by Sofia Coppola

One will notice that the viewer is not focussing on the whisper, but actually trying to interpret the whisper.

### 4.1.3 Reduced Listening

Reduced listening, as described by Chion, are all the sounds listened to for the sound itself and not listened to for what it refers to or represents. In other words, the sound is not semantic as in causal listening. Chion (1994: 31) explains:

... it would seem that film and television use sounds solely for their figurative, semantic, or evocatory value, in reference to real or suggested causes, or to texts – but only rarely as formal raw materials in themselves. ...The emotional, physical, and aesthetic value of a sound is linked not only to the causal explanation we attribute to it but also to its own qualities of timbre and texture, to its own personal vibration.

Within reduced listening, Chion refers to a concept called, an acousmatic situation. This concept describes all sounds which sources or origins are hidden or obscured. In other words, a sound that one hears, not knowing what the source looks like. Figuring out what is creating the sound, one might listen to frequency, pitch and loudness to interpret the source of the sound, which in turn makes use of reduced listening to interpret the sound by listening to the characteristics of the sounds produced.

In all fairness, it may be deduced that reduced listening would always be used in conjunction with another listening mode. The problem with reduced listening lies in its very nature – that it represents nothing but the sound itself. In other words, the nature of the sound is not connected to anything. This concept is very difficult to grasp, for the simple reason that, every human being hears sounds connected to something. When describing sound, one might be lost for words, because in describing a sound, it is already connected to something, at the very least, to the word

itself. The deduction made about reduced listening concerning film, might refer to any sound that enhances an abstract concept.

Take a certain scene from *Harry Potter and the Order of the Phoenix* (2007) directed by David Yates. Harry had dreamt about being a huge snake that has just attacked his best friend's father. The next scene shows Harry being led to Dumbledore (Michael Gambon), who is the headmaster of the school, in order to explain his dream. During the film, it becomes clear that Harry and Voldemort<sup>21</sup> have an unexplainable mental connection. Harry tries to ask Dumbledore what is happening to him, but he will not answer, almost as if Harry is not there. If one listens carefully, one can hear a sound, almost like a voice, but not a voice, that clearly does not fit into the space and time of the scene. It starts very subtle and suddenly peaks, whereby Harry shouts at Dumbledore "look at me!" This sound, prior to the shout, is not grounded in anything and might refer to the unnatural connection between Harry and Voldemort.

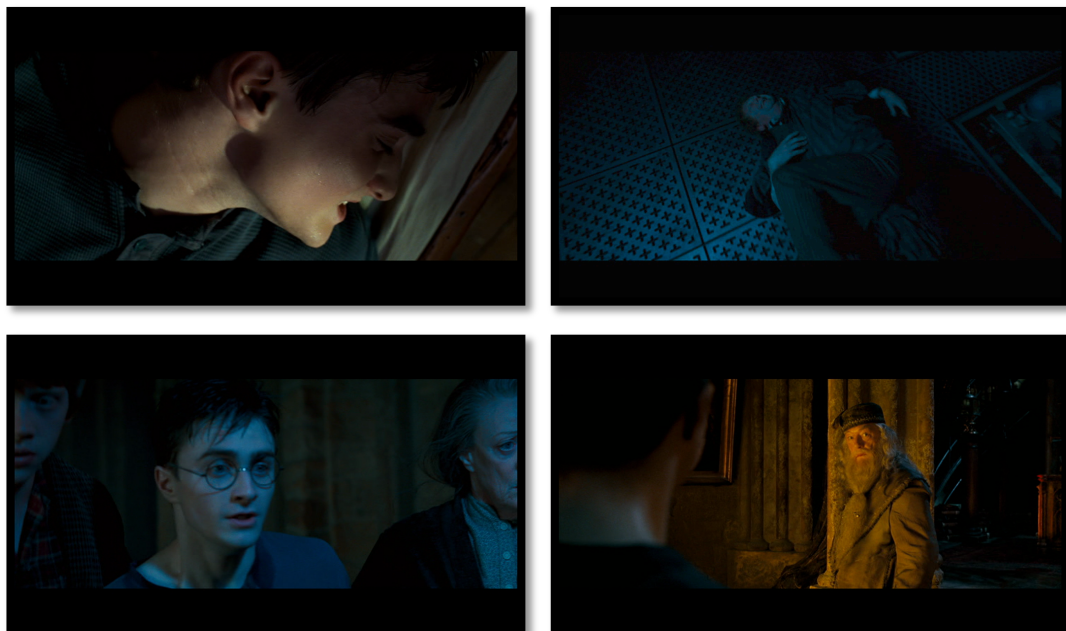


Figure 17: Harry Potter and the Order of the Phoenix (2007), directed by David Yates.

---

<sup>21</sup> Voldemort is the dark wizard of the magical world of the Harry Potter films, who tried to kill Harry when he was a baby. He tried to cast a killing curse on Harry, but in doing so, he killed himself. Voldemort did not die completely, because he created a way that he can live on after being killed and existed as a spirit. With the attempt to kill Harry, Voldemort was now connected to him in some unexplainable way. This connection entitled Harry to feel what Voldemort felt and in some instances see what Voldemort saw. Within the fourth film, Voldemort was brought back to life where he and Harry still shared this connection.



This listening mode differs from causal listening in the sense that the audience will attempt to connect the sound to something within the film's frame of reference and also to their own past experience, whereupon they might realise that the sound's origin is not known and might only be speculated. To clarify this concept, Chion confirms that the characteristics of reduced listening are harboured in the idea that the source is not known. This might also refer to dream sequences where most of the sounds are jumbled up noises that may not even refer to, or indicate something.

It is important to realize that when listening to sound everyday, one will not differentiate between the different listening modes consciously. All sounds heard are in context to one's surroundings. As previously indicated, one will get out of the road when a car is approaching from behind, and not focus on the fact that the combination of sounds are the process of causal, semantic or reduced listening. The main reason the listening modes are mentioned is for the sole purpose of creating awareness to anyone creating sound for film. One must be aware when creating too much realism, the focus on the important facts might be disturbed, not to mention that some *real* sounds would sound *fake* to some viewers, even when it is the real sounds being used.

The argument is that all sound is half code and half music, and there are different ways of interpreting it. According to Murch (2005), "every language is basically a code with its own particular set of rules". If one does not understand the rules, one is not in a position to understand its meaning. For him, sound in film has two extremes, namely encoded sound, which he describes as speech, and embodied sound, which he compares to music. He explains that all humans decode meaning from the language code without realizing it. He states:

Just because we usually do this automatically, without realizing it, doesn't mean it isn't happening... The meaning of what I am saying is 'encoded' in the words I am using. Sound, in this case, is acting simply as a vehicle with which to deliver the code.

He argues that music is completely different, "it is sound experienced directly, without any code intervening between you and it". For him, this is the reason why music is considered as a "universal language".

Therefore, the question arises, what lies between these outer limits? He explicates that in the same way that film sound falls between the 20Hz-20KHz spectrum, all sound falls somewhere between encoded sound and embodied sound.

He reasons that most sounds fall halfway between these extremes, being half language and half music. This midway creation that is sound effects, is balanced between the universally understood music and code-carrying speech. Take for example the sound of the Terminator collector-droid in the *Terminator salvation* (2009) directed by McG.

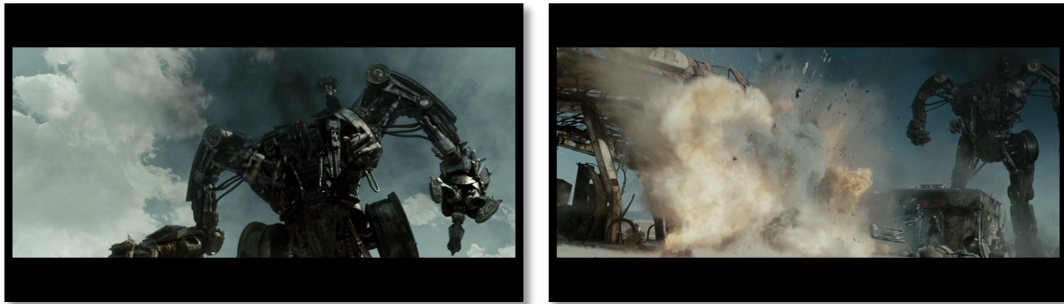


Figure 18: Terminator salvation (2009), directed by McG

The sound emanating from this robot is both menacing and frightening. This *presence* comes directly from the sound the robot radiates. One can hear servos turning and pistons extending as this robot moves about, but one can also hear a different electronic sound that creates a language with which this robot is associated. This musical language, as it were, is precisely what Murch was referring to. Murch also used the example of R2-D2, which is a character in the *Star Wars film series* (1977-2005). This character had a set of beeps of different time frequencies and pitches with which it communicated. Murch explained that one could interpret the emotion of R2-D2's voice just by listening to this tone.

It is clear that sound has different ways of being presented and interpreted. It all comes down to making meaning of the sounds. For example, a knock on the door does not only insinuate that it is made from wood or metal, or that it is a loud sound etc, but more specifically that someone is at the door that wishes to enter.

## 4.2 Conclusion

Within this chapter, the researcher discussed the different ways of learning about one's environment through listening. These listening modes give the audience a chance to realistically interpret different sounds, while skilfully connecting them to different sensations.

It also became clear that when using different modes of listening, the audience is given a vehicle to predict what is going to happen next. This helps to heighten the

emotional impact when it happens as they predicted or turned out completely different from what they expected. These listening modes are also necessary for interpreting different narrative levels of sound in film, which in turn, helps minimise ambiguous crossovers between the narrative levels

Within the next chapter, the researcher will discuss and make use of these listening modes to help the audience interpret the story within a story space.

## Chapter 5: Metadiegetic sound in film

---

The dreamer does not know that he is dreaming; the film spectator knows that he is at the movies: this is the first and principal difference between the filmic and oneiric<sup>22</sup> situations. (Metz, 1976: 75)

When recreating a *memory* on screen for someone to relive, the audience has to accept the memory representation as is. The reason for this is, if they knew that it is a memory of a character, then they cannot criticize the character's way of observation of reality (because it is the character's subjective reality). It is this mental representation that gives way for a metadiegetical understanding of film sound. According to Raskin (1992: 10), metadiegetic sound, which he calls subjective sound "...is undoubtedly the most neglected class of film sound which, hierarchically speaking, deserves to be considered on a par with actual sound, rather than as a subclass of off-screen sound or narrative voice".

Metadiegesis is foremost a narrative concept, nevertheless, it is also a concept with which film sound cannot go without. It must be stated that, because metadiegesis is a narrative concept, it can be assumed that metadiegetic sound as described above, falls under the third notion of narrative theory according to Genette<sup>23</sup>. The reason for this is as Seager (1991: 1) explains it: "metadiegetic narrative is a narrative within another narrative, a story within a story". This may imply that the story being told within the spatiotemporal sphere can be anything that the character believes it is. In other words, the metadiegetic narrative does not have to fit the spatiotemporal sphere. Genette (1980: 26) posits: "it is ... evident that the narrative discourse ... depends absolutely on that action of telling..." He argues that if the character were narrating the story falsely, he would in fact be expanding the importance of the narrative, "for on it depends not only the existence of the discourse but also the fiction of the existence of the actions that it "relates"".

---

<sup>22</sup> *Oneiric* (OED Online (June 2003)) – Of, characteristic of, or relating to dreams; dreamlike.

<sup>23</sup> See page 10 the third notion

In all likelihood what he meant by this is, that when a character's *story* is told, it illustrates much more than the story that is presented. It will also give the audience a means to understand the mindset of a character within the *new* filmic world to which they are introduced.

At this point, the question probably is how metadiegesis evolved from 'a story in a story' to internal text<sup>24</sup>. Genette (1980: 232) described metadiegesis as pertaining to narration by a secondary narrator. Because of the metadiegetic narrative concept, a few people tried to apply it to film in the form of sound. For example, Gorbman (1980: 197) posed the question: "may we speak also of metadiegetic film music?" As previously indicated, she created a hypothetical example in order to demonstrate her idea of metadiegesis:

Early in a film we witness the great romance of protagonist X, which ends tragically during the war. Years later, while X and his best friend Y sit in a bleak café discussing their irretrievable joys, Y brings up the name of X's lost love. This strikes a chord: a change comes over X's face, and music swells onto the soundtrack, the melody that had played early in the film on the night X had met her (Gorbman, 1980: 197).

This music is not diegetic at all, because it is not playing in the temporal space and time of the protagonist, so it must be something else. She names this type of music, *metadiegetic sound*. Gorbman (1980: 196) explains:

..."metadiegetic" images – those supposedly narrated or "imagined" by a character in the film ... In addition to dreams, visions, fantasies, and the like, a whole flashback introduced by a character (who thus becomes a secondary narrator) is a common element of film discourse.

Therefore, according to Gorbman, metadiegesis refers to *internal text* of a character, whereby she states, becomes a *secondary narrator*. This starts the evolution of metadiegesis from 'a story within a story' to internal text. To further delve into the evolution of metadiegesis, Chion (1994: 76) divided metadiegesis into subjective internal and objective internal, however, he never indicated or revealed that he is

---

<sup>24</sup> "text" according to Branigan (1992:87) is "a certain *collection of descriptions of an artifact* where the artifact must be one that materializes a symbol system, and the descriptions that are offered of it must be sanctioned by a society"

referring to metadiegesis, neither did any of the other internal sound theorists. For instance, Raskin (1992: 9) called it subjective sound and explains that other theorists have their own name for it. Spottiswoode (1935: 185) named it non-realistic, Branigan (1984: 68) called it intra-diegetic and Bordwell & Thompson (1980: 201) described it as internal diegetic. This might hold true for some, but Spottiswoode, Branigan and Bordwell & Thompson never said that their descriptions are also called metadiegesis, whereby it still begs the question what the real definition of metadiegesis is.

There existed a transition where metadiegetic sound no longer stood for its original definition, and was now only referring to the *internal events* of a character. Because of this, it might be safe to say that metadiegetic sound theory, at this point, refers to internal (psychological and physiological) sound events<sup>25</sup>. It is also clear that the term may be limited by its evolved boundaries.

Before one can use metadiegesis as a narrative sound element of film, it might be appropriate to back track, so that the original meaning of the word can be used and not further complicate things. An interesting starting point into the investigation of metadiegetic sound narrative is Genette's explanation of the term. His description of metadiegesis is different from what one would imagine it should mean. According to Bal and Tavor (1981: 41) "...in the logico-linguistic tradition the prefix *meta-* indicates an activity having for its object an activity of the same class..." Where metanarrative according to this concept would state: "a narrative *on* the narrative" Genette's explanation of the addition of the prefix, has it refer to: "narrative *in* the narrative" (Bal & Tavor, 1981: 42). This shift helps to explain the use and application of the term metadiegesis.

With this in mind, the researcher proposes that its application should be expanded, to not only include internal sound elements, but ultimately to all sounds that point to a story within a story. This further changes, and maybe generalises the concept, to include everything as long as the action of telling involves additional narrated, uttered or referential text.

The researcher insists on the idea that, even though metadiegetic sounds generally refer to internal narrative, it should always be grounded in its basic form, to tell a

---

<sup>25</sup> See Raskin (1992:10)

story within a story. The question arises, how does one define and apply ‘a story within a story’ metadiegesis to film. In addition, does ‘a story within a story’ metadiegesis only take place within the filmic sphere? Lastly, how does one define internal metadiegesis?

### 5.1 ‘A story within a story’ Metadiegesis

Because metadiegesis is a type of narrative function in this regard, its application will first be explained on its narrative level and then the researcher will attempt to apply it to sound in film. Seager (1991: 3) explains that Genette’s idea of metadiegesis, may have had “three functions in relation to the primary narrative within which it appears...” Firstly, he describes that it serves an explicative function. He reasons that it narrates the reason that led to a certain circumstance or situation. Therefore, to narrate the cause of a current situation, the narrative has to recall something that was previously narrated or something that previously happened, that was not yet narrated about. It can be surmised that the reality in this narrative has to be gained from knowledge that was in the past, and is now needed to help the narrative progress. This piece of past narration will help the perceiver to understand the current situation without losing the main thread of the narrative.

Biro (1982: 116) explains, “...irreversibility is in the nature of things”. In other words, real life situations cannot be changed. If someone did something, one cannot go back and change it, but with narrative, the past can be recalled and a new perception can be relied on to interpret a current situation. She explains that the mind tends to convert things into “reversible forms”. She states:

Transition from raw empiricism to mental recognition begins at the point where we are able to traverse events in both directions: to start at beginnings and reach consequences, and vice versa. This movement from parts to whole, from cause to effect, from operating forces to final results, is never straight; it includes “back up” maneuvers, detours, and interactions of its various factors.

One of the applications of an explicative function in film is called the *flashback*. Deleuze (1994: 48) explains that the flashback is a narrative tool that brings images or even sounds, from the past into the present. The workings are as follows: The audience will be looking at something within the present diegetical space and time, whereupon something from the character’s past is triggered and the audience is flung

back into the past of the character. After the past narrative has been told, the audience is again returned to the present spatiotemporal realm of the character. This function is useful when information or knowledge is needed from a character's past, to tell the story within the current time and space. He states: "...it is a multiplicity of circuits each of which goes through a zone of recollections and returns to an even deeper, ever more inexorable, state of the present situation".

He also adds that certain conditions must be met before a flashback can be used. Conditions such as the narrative cannot be told in the present. Something else, outside of the current space and time is needed to move the narrative forward. This may also refer to narrative that was told earlier in the film, but is now needed for the story to progress. An example of this would be *Hot Fuzz* (2007), directed by Edgar Wright. The film is about a London police officer that is transferred to the country. Here he discovers that something is horribly wrong with the way the local inhabitants handle criminal offences. Every time someone was killed, everyone agrees that it must have been an accident. This frustrates the police officer, and he starts to investigate these killings on his own. Whilst eating a Cornetto ice cream, he has a brainwave. The words of the woman in the shop they have just visited, keeps repeating in his head: "no luck catching them killers then?" This gives him the idea that when she referred to *killers* and not *killer*, he had his answer. The narrative was allowed to resume because he realised within the flashback that there has to be more than one killer.

This narrative mechanism is widely used in films today. It may range from total representation of something that happened in the past, to remnants of a word being remembered.

Secondly, Seager (1991: 3) explains that "it may have a thematic function, i.e., establish relations of contrast and similarity between diegesis and metadiegesis". What he essentially is saying is, that the focus falls on the analogy of the current narrative. This analogy would be perceived by the *audience* and applied to the current narrative. An example of this can be found in *O' brother where art thou* (2000), directed by the Coen brothers. This film is very loosely based on Homer's *Odyssey*. Characters such as the Oracle, the Cyclops, the flood and Ulysses are all in this story, and although the characters are not mythological, the analogy still exists between them. Another example would be the illustration given in the previous chapter about Harry Potter, experiencing the feelings of Voldemort (see chapter 4, reduced listening



on page 55). Harry’s shout of “look at me!” was totally out of character. The analogy must be drawn that it was the anger of Voldemort, channelled through Harry, which led to this remark.

One might even argue that this type of metadiegetical representation is grounded in reduced listening. To be able to notice this type of metadiegetical representation, one would have to make the connection between Harry’s utterance and Voldemort’s anger.

Other uses of this type of metadiegetical representation might include what Sonnenschein calls, sound imagery language. To apply this concept to sound and film, Sonnenschein created a table of the language of sound imagery that is used to create this type of insinuations or analogies:

<b>Language of Sound Imagery</b>	
<b>Simile</b>	Acoustic similarity of two sounds (scream and siren)
<b>Hyperbole</b>	Obvious and intentional exaggeration (scream with alarm clock)
<b>Metaphor</b>	Suggest comparison of an actual sound with an idea (scream with blinking red light)
<b>Allegory</b>	Representation of abstract through concrete (scream held mysteriously until climax, e.g., in <i>The Shout</i> )
<b>Irony</b>	Contrast of least-expected opposites (scream with smile)
<b>Paradox</b>	Apparent contradiction that may express inner truth (scream from cigarette)
<b>Vivification</b>	Manifest living traits in an inanimate object (scream from doormat)

Table 2: Sonnenschein (2001: 55)

The last example of ‘a story within a story metadiegesis’, according to Seager, refers to the action of narrating something that hinders the main narrative. He (1991: 4) uses

the example of Arabian Nights to argue his theory. In the story, Scheherazade delayed her death by continuing to tell stories to the sultan. It is this act of telling that hindered the inevitable end of the main diegetic narrative.

The researcher explained many theories concerning the audience's relationship to the film, and it is this relationship that is called upon to argue the next metadiegetic understanding.

## 5.2 The 'External Story Space'

At this point, it is necessary to stress that the 'external story space' is not so much an extension to metadiegesis but rather a metadiegetic understanding of semiotics and meaning making. The reason for this is, because semiotics is a symbol being interpreted by an observer through his or her own frame of reference, in the same way, the 'external story space' refers to the viewer's capability of interpreting the given story space through their own subjective reality. This too is a type of 'story within a story', and one should be aware of this fact when creating metadiegetic sound for film. When abstract sounds are created for oneiric scenes, the sound designer must always be aware that this created *reality* still has to be interpreted by the viewer. If the created sound is not part of the viewer's repertoire, then the sound might be wrongly interpreted.

To clarify what is meant by *viewer's interpretation*, the researcher will make use of Branigan's definition of implicit and explicit narrative. He (1992: 89) states that explicit narrative refers to all the narration that cannot be confused, and which is clearly defined. Implicit narrative would then be all the narrative elements that are not so well defined and may be ambiguous, mostly forcing the audience to rely on their own intuition or frame of reference to interpret a certain scene or text. When a story is told, the narrator or the author only sketches the elements of the narrative that is important to relay the story, all the rest needs to be assumed by the reader or audience.

To further explain what is meant by implicit and explicit narrative, Branigan relies on an optical illusion.

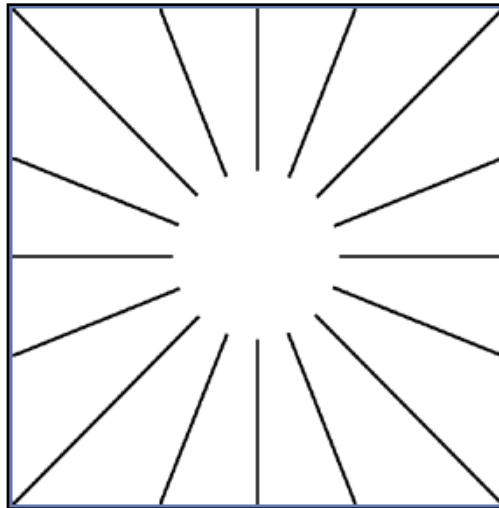


Figure 19: The implicit circle from Branigan (1992: 90)

He reasons (1992: 90) that the lines represent the explicit narrative, while the circle illustrates the implicit narrative. The fact is that there is no circle without the outside lines. He also argues that the more lines in the image, the clearer the circle would become, implying that the more explicit narrative elements presented, the better the clarity of the implicit narrative.

To expound the concept further, Sonnenschein (2001: 81) describes a concept he calls “the law of closure”:

“The law of **closure** states that the mind will tend to unite two disconnected lines lying along the same trajectory. This happens as well with melodic fragments, broken sentences, or any other interrupted sonic information”

Sonnenschein’s statement may even be further expanded to include Branigan’s description, implying that the mind tends to look for patterns, and then completes the elements that are missing. This paves the way for a new understanding of filmic metadiegesis, the *external story space* that includes the audience.

A good example of this type of insinuation of *internal* and *external* ‘story within a story’ metadiegesis would be *Pan’s Labyrinth* (2001). As described in Chapter 2, the story is about a little girl called Ofelia who travels with her mother to visit her stepfather during the Spanish war. Ofelia, who obviously does not like her stepfather, retracts into a fantasy world where fairies, giant frogs, and weird horned creatures roam. The story pans out into a full-featured fantasy story. The only condition being that only Ofelia can observe this world. As the story progresses, she meets an old faun

who tells her that she is a princess of the underworld, and she has to complete a series of tasks to prove that she is worthy to enter the underworld kingdom.

As viewer, one is submerged into this fantasy world, where all the mythical creatures and occurrences are real, and one starts to believe the phantasmagorical creations. On the other hand, one is also aware that only Ofelia experiences this fantasy world and one cannot help to think that everything is in her mind. The truth of the reality of the fantasy world is never confirmed in the film, and the viewers have to decide for themselves if the fantasy world was real or just a figment of Ofelia's mind. The key element with this film is that the viewer replays the film in his or her head to reconfigure the occurrences within the film to reconstruct the first impression of reality.

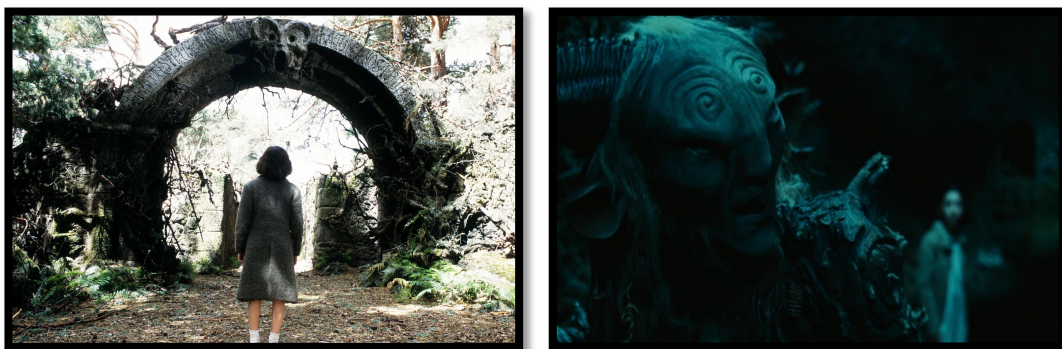


Figure 20: Pan's Labyrinth (2001), directed by Guillermo del Toro

It might be appropriate to add that, learning about this type of metadiegetic representation, one is likely to use causal listening to gather information from the film and process it within the mind to conceptualise a truth. The researcher may even refer back to the examples used to describe causal listing – In *Delicatessen (1991)*, the audience might wonder, why Louison is bouncing on the bed when the butcher's wife asked him to fix it. They might have even figured out that he is listening to find the broken spring.

It might become apparent that there exists a boundary between cognition and perception. When does perception change into cognition? In other words, when watching a film and certain filmic elements are given during the film, to what extent will the viewer perceive the elements, before they have to recall past experiences to interpret what they perceive, or will the given information be enough so that they don't have to delve into their subconscious to retrieve stored data? This isomorphic

boundary might appear to be a problem, but the researcher argues that it might be a useful one.

Referring back to *Delicatessen (1991)*: In the beginning of the film, a character is seen sharpening a cleaver and during this sharpening, the audience is lead along a pipe, all the way following the sound of the sharpening of the cleaver. In the next shot, another man is shown listening to the sound of the cleaver being sharpened. One can see that he is panicky and the audience watches him wrap himself in brown paper. The next scene shows that a garbage truck has come to collect the garbage, and that the man who wrapped himself in the brown paper was now in one of the dustbins. The next shot shows only the man's eyes within the dustbin, and the audience can hear the truck outside while the dustbin is being picked up. The interesting part is that the sound of the truck is getting softer, almost as if the dustbin is being carried away from the garbage truck. He is finally put down and the lid opens, where he realises that the butcher has found him. The sounds that can be heard during this sequence of events, is the sound of the truck that is getting softer, a bell ringing (that might indicate a shop door opening and closing), a knife being picked up and lastly the dustbin that is opened.

Within this last sequence, one is fed different sounds whereupon the audience is called upon to draw their own connections to the sounds heard. In other words, they must draw their own conclusion. With this example, the balance between cognition and perception was relied on so that the story can progress. This example might also infer to the emotion of the character. One might feel sorry for the wrapped up character because he is clearly distressed by the sequence of events.

Another example would be *Memento (2000)*, directed by Christopher Nolan where the story is told in reverse. The film is about a character that was attacked, causing partial damage to his brain, the part that converts short-term memory into the long-term memory part of the brain. He uses a few systems to help him remember things that are important. The only thing he knows for certain is the memories prior to his accident. He figures out that someone raped and murdered his wife, and now he is plotting revenge. One is painfully aware of his mental condition, because, like him, the audience is also left in the dark as he unravels the mysteries of his past. He would tattoo answers to questions he had all over his body, so that he would not need to figure it out again.



Figure 21: Memento (2000), directed by Christopher Nolan

The film starts out with him holding a photo of a dead person. One quickly realises that he must have killed this man. The next scene shows him in his hotel room, clearly not aware of what just happened. The audience might realise that it must be prior to the killing. They have to reinterpret the movie as it unfolds. In the beginning, everything is confusing, with a story that feels as if one has entered late into a conversation. The aim for the audience is to start connecting the dots and retelling the story for themselves as it progresses; figuring out who the characters are, after each segment, re-evaluating their motives and actions. Bit by bit the gaps are filled in until the whole story of the mystery is revealed and the audience knows exactly what everything represents.

As a viewer, the audience would have no input as to how the writer and director will have the film progress and end. They will have to watch it with the knowledge that they have no power to change the course of the events unfolding in front of them. With this in mind, there is one thing that they have a power over - how the narrative is going to be interpreted. The writer and director's job is to audio-visually tell a story. Therefore, when a story is told, not all elements of a certain occurrence can be created or held back to help tell the story or keep it interesting, therefore some of the things in the filmic world have to be assumed.

The result of above mentioned 'story within a story' metadiegesis statement, should not stop there. If one is to bring together the first understanding of metadiegesis (according to Genette) with the evolution of metadiegesis into a musical metadiegetic statement (according to Gorbman), as well as pushing the boundaries to incorporate filmic sounds, it should not only incorporate 'the story within a story', but also 'internal' elements.

### 5.3 Internal Metadiegetic Space

According to Milicevic (n.d.), the journey into the mind of a character relies on two concepts, metadiegesis and oneiric. As previously indicated, oneiric refers to something that has a characteristic or relating to a dream. Milicevic argues that metadiegesis refers to a character's subjective perspective of reality, whilst oneiric refers to the "character's total or partial departure from reality". Both of these concepts point to the subjective reality of a character in film. He argues that oneiric situations in turn, consist of two theories namely "*hypnagogic*" and "*hypnapompic*". Hypnagogic refers to the process of going to sleep or departing from the 'real', whilst hypnapompic in turn indicates the process of waking up or returning to 'reality'.

One of the questions that might come to mind would be, how does one convey thought into a meaningful unit, enabling the audience to interpret the memory as if they are the ones thinking the thoughts. Another question hindering the simplicity of the concept, would be when creating a metadiegetic scene, what element would indicate that the scene is oneiric and not the normal (diegetic) space and time of the character?

Morris (2001: 368) refers to Taylor (2000: 34) and Chion (2000: 204) to explain the relationship between sound and memory. According to Taylor (2000: 34), of the five human senses, sound needs a medium through which it is conveyed. With this in mind, Chion (2000: 204) explicates: "if sounds are easily projected by the spectator onto the film image, it is because the image is circumscribed by a frame that can be located in space, whereas sound lacks a frame". Morris (2001: 368) concluded that because of the lack of frame for the sound, an exploration into the relationship between sound and memory would be a speculative one.

Take for instance a certain scene out of *Harry Potter: The prisoner of Azkaban* (2004) directed by Alfonso Cuarón. Harry (Daniel Radcliffe), Ron (Rupert Grint) and Hermione (Emma Watson), were sitting on the Hogwarts train when it suddenly stopped. All the lights in the train went out and ice formed on the windows indicating that something unnatural was happening. A hooded figure opens the door to the section where Harry, Ron and Hermione were sitting. This figure starts to, what appears to be, sucking the life out of Harry (*see figure 22*). The scene goes black and the only sound that is heard is a woman screaming.



Figure 22: Harry Potter and the Prisoner of Azkaban (2004), directed by Alfonso Cuarón.

This scream was clearly in Harry's head, even though there were no visuals to corroborate the notion. Because of the abstract nature of memory in film, the audience became used to certain cues that indicate mental scenes or memory sequences. Cues such as sounds with a long decay<sup>26</sup> (almost ethereal) will indicate that it is not from this space and time, because the reverberation of the sound does not match the current setting of the characters. According to Sonnenschein (2001: 83), space is defined by visual and auditory cues. Therefore, when hearing a sound with a long decay while seeing a small room, one's mind tends to wobble a bit. Luckily, with years of filmic conditioning, the mind has learned that one must not panic when such irregularities are seen and heard. It is because of the troubleshooting capabilities of the mind, that one can figure out that the sound must be internal and not just erroneous.

Concerning the scene where Harry is attacked by a Dementor (the hooded figure), it is a sequence of cues that hints towards an internal voice. It starts with the camera focussing on Harry's face and zooming into the pupil of his eye when the scene fades to black and the voice echoing (with a long decay) in the darkness (hypnagogic). The audience might feel that they have just entered Harry's head. This notion is then confirmed when Harry comes to his senses and asks Ron and Hermione if they also heard the voice of a woman screaming, where they replied that they had not (hypnapompic)<sup>27</sup>. If the visual cues were not part of this sequence, there could be a chance that the audience would interpret the scream as diegetical rather than metadiegetic, not to mention that the voice could have been in anyone of the other characters' minds sitting in the same cart as Harry.

---

<sup>26</sup> Special sounds are defined by visual and auditory cues

<sup>27</sup> see figure 22



It is interesting to note that this sequence held even more information than just being internal metadiegetic sound. By using reduced listening capabilities, one might realise that the woman screaming in Harry's head was not just any woman, it was his mother. At first glance, the likelihood of anyone realising this at that moment, might not refer to the majority of the audience. It is only later in the film that Harry realises that it was in fact his mother's last verbal utterance before Voldemort killed her.

It must also be stated that cues are highly dependant on film genres. For instance, it might be easier to convey metadiegetical scenes in a film dealing with everyday situations and events, than it would be creating metadiegetical scenes for a bizarre science-fiction film.

Cues hinting towards oneiric scenes, in a natural setting, would be bizarre images that accompany the sound. Take *Eternal sunshine of the spotless mind* (2004) directed by Michel Gondry as an example. Most of the time, when Joel Barish (Jim Carrey) relives his past, something unnatural happens in the picture. In a certain scene, Joel enters the doctor's office that is organizing the memory erasing procedure. The doctor turns around to face him, but the doctor has no face (see figure 23). This is an unnatural image in a normal setting; therefore, the audience would automatically assume that it must be oneiric.



Figure 23: *Eternal Sunshine of the Spotless Mind* (2004) Directed by Michel Gondry

Concerning science fiction oneiric scenes, where monsters, magic and future technology roams, other cues may be relied upon in order not to confuse the audience. Cues such as image corroboration, or as described earlier, an unnatural room-tone to the sound that does not fit the setting. An example of both these cues would be *Lord of the rings: Fellowship of the ring* (2001) directed by Peter Jackson. When Galadriel (Cate Blanchett) speaks to Frodo in his mind, the two above-mentioned cues are used (see figure 24). Firstly, Galadriel's face was shown and the audience could see that

her lips were not moving, even though she was clearly talking. It goes without saying that Galadriel's voice first had to be introduced before it could be used in an oneiric scene. The other cue would be that her voice did not fit into the temporality hinting towards internal sound, because when the other characters spoke, their voices did not sound as if spoken in a church.



Figure 24: Lord of the Rings: The Fellowship of the Ring (2001) Directed by Peter Jackson

Other mechanisms that can hint towards an oneiric situation might be a sudden loudness in the sound, or a total silence of all sounds but the sound that the character is concentrating on. Sometimes the sound will slow down and other times a hyper-reality where the articulation of the sound is crisp and clear. All of these cues will only stand out because they do not fit into the spatiotemporal sphere of the character.

#### 5.4 Conclusion

To conclude this chapter, the researcher has discussed the intricate third level of film sound narrative. The researcher attempted to unravel the true meaning of metadiegetic sound, and discovered that it might have three different mechanisms. The true meaning of metadiegesis had it refer to elements that referred to a 'story within a story'. This true meaning is still widely used and even though its use evolved to a somewhat different meaning, it still refers to a *story within a story*.

The second metadiegetical understanding the researcher referred to was external story space. This understanding of metadiegesis was only mentioned so that it can be stressed that the audience cannot be disregarded when creating sound for a story-within-a-story metadiegesis. The audience's expectations and interpretations can be used to create a disconnected and abstract understanding of certain sounds.

The last metadiegetical understanding was the internal story space. As previously indicated, the internal story space is the most widely used mechanism. It gives the

audience a glimpse of the internal workings of a character. The internal workings may refer to a memory of a character or even to something as abstract as a feeling. This mechanism helps the audience to understand a story that is not linear but rather as Biro (1982: 116) explained, a back-and-forth movement within a linear storyline.

## Conclusion

---

“What new content can be brought into the cinema by the use of sound?” (Pudovkin, 1974: 183)

In the beginning of this research endeavour, the researcher referred to Pudovkin (1974: 183) who stated that “the role which sound is to play in film is much more significant than a slavish imitation of naturalism...”. Pudovkin explained that something more should be done with the potential held by film sound. For him, this *something* is the evolution of sound and image, materialising to the point where they meander forth independently. This argument may be construed to indicate that sound is not supposed to enhance the image and for the same reason, that image should not enhance sound. Both sound and image can progress separately, in fact, Bresson (1977: 27-28) theorises that there are eight *rules* governing the interplay between sound and image:

- To know thoroughly what business that sound (or that image) has there.
- What is for the eye must not duplicate what is for the ear.
- If the eye is entirely won, give nothing or almost nothing to the ear.<sup>28</sup> One cannot be at the same time all eye and all ear.
- When a sound can replace an image, cut the image or neutralize it. The ear goes more toward the within, the eye toward the outer.
- A sound must never come to the help of an image, nor an image to the help of a sound.
- If sound is the obligatory complement of an image, give preponderance either to the sound or to the image. If equal, they damage or kill each other, as we say of colors.
- Image and sound must not support each other, but must work each in turn through a sort of relay.
- The eye solicited alone makes the ear impatient, the ear solicited alone makes the eye impatient. Use these impatiences. Power of the cinematographer who appeals to the two senses in a governable way. Against the tactics of speed, of noise, set tactics of slowness, of silence.

---

<sup>28</sup> “And vice versa, if the ear is entirely won, give nothing to the eye” (Bresson).

It is clear that Bresson's theorised rules points to an asynchronous movement between the sound and the image. The argument can be made that, when looking at the filmic image, one is observing something that was captured at a particular time and place. Even though the sound has been recorded at a different place at a different time, it is more likely that the audience will be convinced that it is the true sound of a depicted location. In other words, the process engages in recreating reality through sound. Totaro (2004) states: "What this implies is that the ear is less likely to discern or be ontologically bothered by a technologically mediated difference. The ear accepts reproductions more willingly than the eye". One might argue that, because the eye sees only parts of the real world at any time, the ear will have to consider, enhance and explain that which the eyes are not seeing. In other words, the ear will have to *convince* the body that what is perceived, actually exists without visual clarification.

This brings the researcher back to the example Gaver (Gaver, 1993: 1-2) gave concerning everyday listening (see chapter 4.1 on page 50). One hears a car, and knows that one should not keep walking in the road. This happens without visual corroboration. This 360° hearing range creates a field through which the audience will be transported.

The researcher is attempting to convey exactly this, that film sound's full potential is reached when it is connected through an asynchronic audiovisual fusion, in other words, moving at its own rhythm and pace, not mimicking its visual counterpart. Only when sound is disconnected from the image, can it represent something other than the stereotypical concepts it was usually attached to. Sinclair (2003: 20) argues that:

...sound has a greater scope, that it can "do more" perceptually to the experiencer than the visuals can. One way it does this is by accessing a realm that the visuals may claim to invoke but that only sound can tap: the experiencer's imagination.

The result is the creation of another dimension within the already conceptualised three-dimensional filmic space. Sounds depicting smells, such as *Perfume*; sounds creating the presence of the supernatural, such as *The devil's backbone*; sounds creating the presence of loneliness, such as *Eternal sunshine of the spotless mind*, sounds creating a presence of rage such as *Harry Potter and the Order of the Phoenix*.

The point of this example is to stress the fact that film sound should not only be used to represent sound-making elements, but can also be used to represent abstract

concepts that do not necessarily refer to sound. To put it differently, the audio-visual fusion induces an exploding building to feel catastrophic, a man standing on top of a building, prepared to jump, creates a jolt in the audience's stomach, or a body falling on rocks, accompanied by blood-freezing cracking sounds induces the observer to cringe.

It is this alchemy of the audio-visual collaboration that draws from the physiological triggers present in humans that give the sound designer a starting point for creating a third dimension where the sound can tell more than slavishly imitating actions on screen.

At this point, Pudovkin's request remains challenging to realise. However, one thing is clear, namely that, film sound has come a long way since the 1900s. In addition, even though there is always room for creative improvement, sound, as a means of cinematographic expression, is finally catching up to its visual counterpart.

## Bibliography

---

- Altman, R., 1992. *Sound Theory Sound Practice*. New York: Routledge.
- Bal, M. & Tavor, E., 1981. Notes on Narrative Embedding. *Poetics Today*, 2(2), pp.41-59. Narratology III: Narrative and Perspective in Fiction.
- Biró, Y., 1982. *Profane Mythology - The Savage Mind of the Cinema*. Bloomington: Indiana University Press.
- Bordwell, D. & Thompson, K., 1980. *Film Art: An Introduction*. 2nd ed. Massachusetts: Addison-Wesley Publishing Company.
- Branigan, E., 1984. *Point of View in the Cinema: A Theory of Narration and Subjectivity in Classical Film*. Berlin: Mouton.
- Branigan, E., 1992. *Narrative Comprehension and Film*. London: Routledge.
- Bresson, R., 1977. *Notes on Cinematography*. New York: Urizen Books.
- Brinker, M., 1977. Aesthetic Illusion. *The Journal of Aesthetics and Art Criticism*, 36(2), pp.191- 196.
- Burch, N., 1982. Narrative/Diegesis - Thresholds, Limits. *Screen*, 23, pp.16-33.
- Carroll, N., 1988. *Mystifying Movies: Fads & Fallacies in Contemporary Film Theory*. New York: Columbia University Press.
- Chion, M., 1994. *Audio-Vision: Sound on Screen*. New York: Columbia University Press.
- Chion, M., 2000. Audio-Vision and Sound. In P. Kruth & H. Stobart, eds. *Sound*. New York: Cambridge University Press.
- Coleridge, S.T., 1956-71. *Collected Letters of Samuel Taylor Coleridge*. Oxford: Clarendon.
- Coleridge, S.T., 1957. *The Notebooks of Samuel Taylor Coleridge*. New York: Routledge.

- Da Vinci, L., 1938. *Notebooks of Leonardo Da Vinci*. Translated by E. MacCurdy. New York: Reynal and Hitchcock.
- Deleuze, G., 1994. *Cinema 2: The Time-Image*. Translated by H.&G.R. Tomlinson. London: The Athlone Press. First published in Great Britain in 1989.
- Descartes, R., 1984. *The philosophical writings of Descartes*. Translated by J. Cottingham, R. Stoothoff & D. Murdoch. Cambridge: Cambridge University Press.
- Easton, R.D., 1954. Student Question on Aesthetics. *College Art Journal*, 14(1), pp.46-51.
- Eisler, H. & Adorno, T.W., 1994. *Composing for the films*. London: Athlone Press. Originally published under Eisler's name alone: New York : Oxford University Press, 1947.
- Forrester, A.M., n.d. *Auditory Perception and Sound as Event: Theorising Sound Imagery in Psychology*. [Online] Available at: <http://www.kent.ac.uk/arts/sound-journal/forrester.txt> [Accessed 11 June 2007].
- Freud, S., 2003. *The Uncanny*. Translated by D. McLintock. London: Penguin Books Ltd.
- Gaver, W.W., 1988. *Everyday Listening and Auditory Icons*. Doctoral dissertation. San Diego: University of California.
- Gaver, W.W., 1993. What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception. *Ecological Psychology*, 5(1), pp.1-29.
- Gee, J.P., 2003. *What Video Games Have To Teach Us About Learning and Literacy*. New York: Palgrave Macmillan.
- Genette, G., 1980. *Narrative Discourse*. Translated by J.E. Lewin. Oxford: Basil Blackwell.
- Genette, G., 1982. *Figures of Literary Discourse*. Translated by A. Sheridan. New York: Columbia University Press.
- Gorbman, C., 1980. Narrative Film. *Yale French Studies*, (60), pp.183-203.



- Gracyk, T., 2003. *Hume's Aesthetics*. [Online] Available at:  
<http://plato.stanford.edu/entries/hume-aesthetics/> [Accessed 18 May 2007].  
 substantive revision Sep 23, 2008.
- Hume, D., 1826. *The Philosophical Works of David Hume: Including All The Essays, and Exhibiting the More Important Alterations and Corrections in the Successive Editions Published by the Author*. Edinburgh: Black and Tait.
- Kafka, F., 2007. *Metamorphosis and Other Stories*. Translated by M. Hofmann.  
 London: Penguin Books Ltd.
- Lindsay, P.H. & Norman, D.A., 1973. *Human Information Processing: An Introduction to Psychology*. 4th ed. New York: Academic Press Inc.
- Link, S., 2004. Not the Eye Filled with Seeing: The Sound of Vision in Film.  
*American Music*, 22(1), pp.76-90.
- Lusseyran, J., 1963. *And There Was Light*. Translated by E.R. Cameron. Boston:  
 Little, Brown.
- Metz, C., 1976. The Fictional Film and its Spectators: a Metapsychological Study.  
*New Literary History*, 8(1), pp.75-105.
- Milicevic, M., n.d. *Film Sound Beyond Reality: Subjective Sound in Narrative Cinema*. [Online] Available at:  
[http://myweb.lmu.edu/mmilicevic/NEWpers/\\_PAPERS/beyond.pdf](http://myweb.lmu.edu/mmilicevic/NEWpers/_PAPERS/beyond.pdf) [Accessed 3 June 2010].
- Morris, L., 2001. The Sound of Memory. *The German Quarterly*, 4(4), pp.368-78.
- Mott, L.R., 1990. *Sound Effects: Radio, Tv and Film*. London: Focal Press.
- Mouton, J., 2001. *How To Succeed in Your Master's & Doctoral Studies*. Pretoria:  
 Van Schaik.
- Mulvey, L., 2006. *Death 24x a Second: Stillness and the Moving Image*. London:  
 Reaktion Books Ltd.
- Murch, W., 1995. *In the Blink of an Eye: A Perspective on Film Editing*. Silman-James Press.

- Murch, W., 2005. *Dense Clarity - Clear Density*. [Online] Available at: <http://talk.transom.org/webX?7@@.eeb42cf/0> [Accessed 21 May 2010].
- Pierce, C.S.S., 1931-1958. Pragmatism and Pragmaticism. In C. Hartshorne, P. Weiss & A.W. Burks, eds. *Collected Papers*. Cambridge: Harvard University Press. pp.1-455.
- Price, H.H. & Braithwaite, R.B., 1964. Symposium: Half-Belief. *Proceeding of the Aristotelian Society, Supplementary Volumes*, 38, pp.149-74.
- Pudovkin, V.I.P., 1974. *Film Technique and Film Acting*. London: Vision Press Limited.
- Raskin, R., 1992. *Varieties of Film Sound: A New Typology*. [Online] Available at: <http://imv.au.dk/~pba/Homepagematerial/MMproduktionmateriale/Raskin%20Sound%20%20Paper.pdf> [Accessed 12 July 2007].
- Rauschecker, J.P., 2005. Neural Encoding and Retrieval of Sound Sequences. In G. Avanzini, S. Koelsch, L. Lopez & M. Majno, eds. *The Neurosciences and Music II: From Perception to Performance*. New York: The New York Academy of Sciences.
- Seager, D.L., 1991. *Stories Within Stories: An Ecosystemic Theory of Metadiegetic Narrative*. New York: Peter Lang Publishing Inc.
- Sinclair, C., 2003. Audition: Making Sense of/in the Cinema. *Project Muse: The Velvet Light Trap*, (51), pp.17-28.
- Smith, M., 1995. Film spectatorship and the Institution of Fiction. *The Journal of Aesthetics and Art Criticism*, 53(2), pp.113-27.
- Sobchack, V., 1995. *Phenomenology and the Film Experience*. New Jersey: Rutgers University Press.
- Sonnenschein, D., 2001. *Sound Design: The Expressive Power of Music, Voice, and Sound Effects in Cinema*. California: Michael Wiese Productions.
- Souriau, E., 1953. *L'Univers Filmique*. Paris: Flammarion.

- Spottiswoode, R., 1935. *A Grammar of the Film: An Analysis of Film Technique*. London: Faber & Faber.
- Taylor, C., 2000. The Physics of Sound. In P. Kruth & H. Stobart, eds. *Sound*. Cambridge: Cambridge UP.
- Totaro, D., 2004. "Notes" on *Notes on the Cinematographer*. [Online] Available at: [http://www.horschamp.qc.ca/new\\_offscreen/notes\\_cinematographer.html](http://www.horschamp.qc.ca/new_offscreen/notes_cinematographer.html) [Accessed 5 June 2010].
- Truax, B., 2001. *Acoustic Communication*. 2nd ed. London: Ablex Publishing.
- Vallins, D., 2000. *Coleridge and the Psychology of Romanticism: Feeling and Thought*. New York: Palgrave Macmillan.
- Yewdall, D.L., 2003. *Practical Art of Motion Picture Sound*. 2nd ed. USA: Focal Press.
- Zettl, H., 1990. *Sight Sound Motion: Applied Media Aesthetics*. 2nd ed. California: Wadsworth.

## Filmography

- Delicatessen*. 1991. [Film] Directed by Jean-Pierre Jeunet & Marc Caro. France: Constellation.
- Eternal Sunshine of the Spotless Mind*. 2004. [Film] Directed by Michel Gondry. USA: Anonymous Content.
- Fight Club*. 1999. [Film] Directed by David Fincher. USA: Fox 2000 Pictures.
- Harry Potter and the Order of the Phoenix*. 2007. [Film] Directed by David Yates. UK, USA: Warner Bros. Pictures.
- Harry Potter and the Prisoner of Azkaban*. 2004. [Film] Directed by Alfonso Cuaron. UK: Warner Bros. Pictures.
- Hero*. 2002. [Film] Directed by Zhang Yimou. Hong Kong: Beijing New Pictures Film co.

*Hot Fuzz*. 2007. [Film] Directed by Edgar Wright. UK, France: Big Talk Productions.

*In Bruges*. 2008. [Film] Directed by Martin McDonag. UK: Blueprint Pictures.

*Lord of the Rings: Fellowship of the Ring*. 2001. [Film] Directed by Peter Jackson.  
New Zealand, USA: New Line Cinema.

*Lord of the Rings*. 2001-2003. [Film] Directed by Peter Jackson. New Zealand, USA:  
New Line Cinema.

*Lost in Translation*. 2003. [Film] Directed by Sofia Coppola. USA, Japan: Focus  
Features.

*Memento*. 2000. [Film] Directed by Christopher Nolan. USA: Newmarket Capital  
Group.

*O' Brother, Where Art Thou*. 2000. [Film] Directed by Coen Brothers. USA:  
Touchstone Pictures.

*Pan's Labyrinth*. 2001. [Film] Directed by Guillermo del Toro. Spain: The Tequila  
Gang.

*Perfume: The Story of a Murderer*. 2006. [Film] Directed by Tom Tykwer. Germany:  
Constantin Film Production.

*Terminator Salvation*. 2009. [Film] Directed by McG. USA: The Halcyon Company.

*The Assassination of Jesse James by the Coward Robert Ford*. 2007. [Film] Directed  
by Andrew Dominik. USA: Warner Bros. Pictures.

*The Bourne Supremacy*. 2004. [Film] Directed by Paul Greengrass. USA: Universal  
Pictures.

*The Devil's Backbone*. 2001. [Film] Directed by Guillermo del Toro. Spain, Mexico:  
El Deseo S.A.

*The Truman Show*. 1998. [Film] Directed by Peter Weir. USA: Paramount Pictures.

*Trop belle pour toi*. 1989. [Film] Directed by Bertrand Blier. France: Ciné Valse.