

# **Unravelling Point of Audition**

*A discussion of Point of Audition in the sound theory of  
modern film and television*

Marie Welde Tranås



**Master's thesis in musicology  
Department of Musicology  
University of Oslo  
November 2017**

*This page intentionally left blank*

# **Unravelling Point of Audition**

*A discussion of Point of Audition in the sound theory of  
modern film and television*

Marie Welde Tranås

Master's thesis in musicology

Department of Musicology  
University of Oslo  
November 2017

Copyright  
2017

Unravelling Point of Audition - *A discussion of Point of Audition in the sound theory of modern film and television*  
Marie Welde Tranås

<http://www.duo.uio.no>

Publisher: Representralen, University of Oslo

# Abstract

---

This thesis investigates the term Point of Audition within sound theory, by reviewing literature dealing with this term, and discusses the term's position in the academic field today, after Svein Høyer in 2013 suggested to refine and expand the term by adding four sub-categories. The purpose of this discussion is to compare and complement Høyer's categories with related or more or less corresponding terminology, to reveal some of their limitations. David Neumeyer put forth in 2015 a set of categories for describing auditory phenomena which range broader than Høyer's attempt to expand Point of Audition to make it cover more. Høyer and Neumeyer's categories lay the basis for this thesis' analysis, which aims at showcasing some of the grey areas not covered by their categories. It also opens up for a discussion of Robynn Stilwell's *Fantastical Gap* (2007) as a possibility to capture the auditory phenomena that are falling between categories. This thesis hopes to be a step towards answering a larger question: Whether the conceptual framework for discussing sound in film and television is fulfilling its purpose, namely aiding our understanding of how the different auditory perspectives influence the audience's experience of the visual narrative.

*This page intentionally left blank*

# Acknowledgements

---

After many years of studies, I am thrilled to finally be able to publish this thesis as a testament to my hard work. Several people have made this process easier, and I wish to thank you all.

First, I would like to thank my advisor Nanette Nielsen for guiding me through the academic landscape, and for being patient with me and giving thorough remarks.

A special thank you to Erlend for supporting me emotionally by keeping up my spirit, and helping me find structure in an everyday life with so much else going on.

I would also like to thank Marissa and Hanna for proof-reading, and an extra thank you to Hanna for commenting, suggesting and correcting at the finishing stage. This was highly valuable.

Friends, family and co-workers outside of Blindern have also been very supportive, thank you for being so understanding.

Last, but not least a big thank you to my class-mates for keeping up my spirit at Blindern during the last two years, and for making it fun and exciting to study.

Oslo, October 2017

*This page intentionally left blank*



# Table of Contents

---

<b>1.</b>	<b>Introduction</b>	<b>1</b>
<b>2.</b>	<b>Theory and Methodology</b>	<b>5</b>
	<i>2.1 The History of Point of Audition</i>	5
	2.1.1 Problematising POA	6
	<i>2.2 Høyer's categories</i>	8
	<i>2.3 Subjectivity and objectivity</i>	10
	2.3.1 Audio-visual Dissonance, Fidelity and Synchronisation	11
	2.3.2 Audio sweetening	11
	2.3.3 Differentially audible sound	12
	<i>2.4 Methodology and Thesis Structure</i>	13
<b>3.</b>	<b>Further theoretical contexts - Getting beyond Høyer</b>	<b>15</b>
	<i>3.1 Diegetic and Non-Diegetic sound</i>	15
	3.1.1 Simple and Displaced	16
	3.1.2 Empathy and Anempathy	18
	3.1.3 The metadiegetic and The Fantastical Gap	19
	<i>3.2 Neumeyer's categories</i>	20
	3.2.1 Combining narrative levels with framing	21
	<i>3.3 A comparison of Høyer and Neumeyer's categorisations</i>	24
	3.3.1 Comparing Høyer and Neumeyer's categorisations	25
	3.3.2 Neumeyer's onscreen diegetic	26
	3.3.3 Neumeyer's offscreen diegetic	26
	3.3.4 Neumeyer's onscreen non-diegetic	28
	3.3.5 Neumeyer's offscreen non-diegetic	28
	<i>3.4 Compilation of Categories</i>	29
<b>4.</b>	<b>Analysis</b>	<b>31</b>
	<i>4.1 Part one: Several different POAs at the same time</i>	31
	4.1.1 Defining the narrator's POA	32
	4.1.2 Simultaneous time zones	35
	<i>4.2 Part two: Active POA and the Fantastical Gap</i>	37

<b>4.2.1 The grey area between Active and Individual POA</b>	<b>38</b>
<b>4.2.2 The grey area between Observational and Individual POA</b>	<b>40</b>
<b>4.2.3 Grey area between diegetic and non-diegetic, and subjective and objective</b>	<b>43</b>
<b>4.3 Summary</b>	<b>46</b>
<b>5. Concluding notes: scope and ambition</b>	<b>47</b>
<b>6. References</b>	<b>51</b>
<b>6.1 Literature</b>	<b>51</b>
<b>6.2 Television-series and episodes</b>	<b>53</b>
<b>6.3 Film scenes, songs and images</b>	<b>53</b>
<b>Film scenes</b>	<b>53</b>
<b>Songs</b>	<b>54</b>
<b>Images</b>	<b>54</b>

# 1. Introduction

---

Our experience of audio-visual interplays, such as films and television series, are shaped by both the director's choices, and by our own personal experiences and values. An image can be understood in many different ways depending on the sound and/or music accompanying it, and watching a film clip twice with different sound can alter our entire perception of the film's meaning. This makes sound an essential component in films, and a worthy proprietor of a coherent theoretical framework. However, the terminology used for explaining and describing auditory phenomena in film and television is not sufficient in that it often is ambiguous, and several different terms are used for the same phenomenon. Also, as this thesis will show, modern film and television have come to demand a much more specific and detailed terminology, which current theory is not equipped to account for. This thesis will investigate strengths and weaknesses in current terminology, with the term *Point of Audition* as a main subject, and suggest refinements of established terminology in order to start a debate on whether it is expedient for use in analysis of modern film and television series.

The director's determination of which visual perspective the audience should have, can be executed by the direction of the angle and position of the camera; a very effective way to steer the audience's focus towards the desired perspective. Sound can also be used to steer the audience's perspective by deciding a specific "point" in the filmic universe, or *diegesis*, where the sound is derived from. The sound perspective is, however, much more complex than the visual perspective, and theorising it has historically proven to be rather problematic. Some theoreticians have opted to use the term Point of Audition (henceforth POA), as an auditory counterpart of the visual term *point of view* (POV). This has been heavily criticised, partly because it simplifies a complex phenomenon, and also because it is too closely tied to visual theory. Nevertheless, there are researchers who use the term in their analysis, although with slight variations in meaning, which can be confusing.

I first became aware of the problematic aspects of the term through the essay *The Relevance of Point of Audition in Television Sound: Rethinking a Problematic Term* (2013) by Svein Høyer, who, as the essay's title indicates, discusses its relevance when it comes to modern

television sound. Høyer attempts to refine the understanding of POA by expanding it to a set of subcategories in order to better cover more complex auditory phenomena. By expanding POA to four subcategories, Høyer has made the term more comprehensive and nuanced, which is essential if POA is to be adequately used in analysis of auditory perspectives in film and television. However, Høyer underlines that he wishes to avoid excess terminology, as the terminology should be easily accessible for other theoreticians, researchers and practitioners of sound theory. The question that arises is whether the extension of categories is helpful, or if it simply makes the term even more complex and subject for misinterpretations. Høyer's categories help nuance our understanding of more complex phenomena, and adequately cover the instances of sound presented in his essay. However, as he himself acknowledges, there are still grey areas left uncovered by his categorisations. Høyer's POA covers only *diegetic sound*, sound that derives from within the filmic universe, and he therefore excludes *non-diegetic sound* like background music or commentary sound, from his essay.

To better understand the different uses of terminology regarding sound in film and television, I conducted a literature review in 2016 (Tranås, 2016) through which I discovered that not much has been written about POA, and that the few theoreticians that have written about it all state that the term is tricky, problematic, ambiguous and difficult. They all offer differing definitions and delimitations, meaning that it lacks an agreed upon definition that unites scholars working with film theory, and they do not speak the same language when they use the term. This might be one of the reasons so few scholars actually do. This begs an important question: whether the term is at all useful. In my opinion the term focuses on something that is very important when theorising film, namely which auditory perspectives we are given, and what "points" we are "listening from". Sound plays a decisive role in audio-visual narration, and coming up with a suitable terminology that gives justice to the auditory phenomena I think is long-awaited.

This thesis has several purposes in mind, but its main aim is to explore the possibility of making POA being less ambiguous and problematic if its understanding of auditory perspective were broader. An overall scientific aim would be to find a terminology that best enhances our understanding, and facilitates debates surrounding sound theory, but the scope of this thesis allows merely for an attempt to form a step towards answering the larger question: Whether the conceptual framework for discussing sound in film and television is

fulfilling its purpose, namely aiding our understanding of how the different auditory perspectives influence the audience's experience of the visual narrative.

With a focus on Høyer's categorisation of POA, other related and more or less corresponding terminology will be discussed and compared, such as synchronous and asynchronous sound (Bordwell and Thompson, 1985), differentially audible sound (Raskin, 1992) and David Neumeier's (2015) categories, which include the non-diegetic aspect as well as whether the sound can be said to be *offscreen* or *onscreen*. This is done in order to nuance the term further and offer a better analytical tool when discussing the use of audio in film and television. Neumeier's categories cover much of the same phenomena as Høyer's, but uses different terminology than POA, and have a broader scope. A compilation of their two sets of categories will form the basis of my analysis, which present good examples of auditory situations which can be problematic to describe using the established terminology. The examples aim to show the gaps in the existing terminology in practice, to get a better understanding of the difficulties of applying sound perspective terminology.

This will demand further expansion of categories, such as Robynn Stilwell's (2007) *Fantastical Gap*, which broadens our understanding of auditory phenomena where the borders between what is diegetic and non-diegetic, as well as subjective and objective, are blurred.

My research question are the following:

*In his essay, "The Relevance of Point of Audition in television sound: Rethinking a problematic term" (2013), Høyer argues for an extension of the concept of Point of Audition, and does this by adding four subcategories.*

*Is the refinements Høyer is offering helpful, or is it simply making the term even more complex and subject for misinterpretations?*

*Focusing on aspects of subjectivity and objectivity, and conceptions of what is understood as diegetic and non-diegetic, this thesis investigates the extent to which a new approach to POA, informed by a discussion that includes alternative terminology, can offer insights that go beyond Høyer's relatively narrow framework.*



## 2. Theory and Methodology

---

The literature on POA is not extensive. In my 2016 literature review (Tranås, 2016), I identify Rick Altman (1992), Michel Chion (1994), Birger Langkjær (2000) and Svein Høyer (2013) to be the most relevant theoreticians, but since 2013 no theoretically exploratory contribution have been given to the subject. This does not mean that film theoreticians do not care about auditory phenomena, it has been quite the contrary during the last decades, but the terminology is not coherent, and it is used with great ambiguity. The theoretical framework for POA is to some extent limited by the lack of vast literature on the subject, but this is, however, the motivation for exploring the subject. The aim of this thesis is to fill some of the blanks that I perceive as problematic for the proper use of the term, in order to set the scene for a discussion on whether the conceptual framework is too ambiguous to be useful in the academic field. This chapter gives a background to the different ways POA has been or is understood, through an introduction to the term's history and its key theoreticians. I will then present Høyer's categories, in order to understand where the research on POA currently stands. Lastly, this chapter will discuss terminology that deals with grey areas between subjectivity and objectivity, which POA currently is unable to cover.

### 2.1 The History of Point of Audition

*Point of Audition*, abbreviated POA, is a film theoretical concept which seeks to define and explain which angle and position the spectator of a film, or a television show, is listening from. It has its origin in the term *Point of View* (POV), which is heavily used in film studies referring to the angle and position the spectator is watching from, and POA is often simply seen as its auditory counterpart. However, theoreticians working with POA try to avoid this somewhat simplistic definition of this concept, claiming that POA is a much more multi-faceted term to describe and apply than POV. What makes POA so fascinating as a theoretical concept is that the few theoreticians writing about it all state that the concept is difficult and problematic: Michel Chion looks at POA as “a tricky and ambiguous notion” (Chion, 1994, p. 89), Rick Altman says it is “a clumsy term whose only merit is to recall

unfailingly the “point of view” shot” (Altman, 1992, p. 60), Svein Høyer calls it “a complex phenomenon that is hard to describe” (Høyer, 2013), and Hibberd and Nielsen tells us that “POA is traditionally recognized as a highly problematic term” (Hibberd and Nielsen, 2015). So why is this term still relevant? And if it is so problematic, why use it? Well, despite all this, there are still theoreticians using the term in their analysis, such as Rebecca Coyle (2010) and William Whittington (2011), but they deal with the problematic by defining POA in the manner that suits their analysis.

The exact reason why this term is not elaborated by many theoreticians is uncertain, but within film studies it is known that the auditory has traditionally been overshadowed by the visual (Altman 1992). It is not until recent decades – the 1980s to be specific – that film sound has claimed its position as a theoretical research object. All film study research acknowledges of course the importance of sound and music in films, but it is striking how little attention it has received. Those film sound scholars who have written about it however, demonstrate the importance of having a language and a theory to explain and discuss interesting phenomena concerning both the visual and the auditory. Some researchers have opted for the need of a new terminology; one that is better equipped to discuss film sound on its own, and does not rely on the often visual terminology of film studies. As will become clear from the following review, the few theoreticians who do concern themselves with POA disagree amongst each other, and are heavily criticising each other’s theories. This debate will serve as the background for this thesis’ discussion of whether there is a need for a new conceptual framework when talking about film sound.

### **2.1.1 Problematising POA**

The oldest source that mentions POA is Rick Altman’s book *Sound Theory Sound Practice* (1992), and it is referenced to whenever later researchers discuss POA. Altman’s definition might therefore be seen as the foundation for all further discussions about this term, even though Altman cannot be credited for creating the term. He refers to Joseph P. Maxfield, a researcher who already in 1938 wrote about the dramatic importance of the internal auditor, which shows that POA is not a new discovery, in fact it has for a long time been necessary to find a term that covers its essence. Altman defines POA as “representing sound as it would be heard from a point within the diegesis [the filmic universe], normally by a specific



character” (id., p.60). For something to be characterised as POA for Altman, there has to be an *internal auditor* as opposed to an *external auditor*, where the latter has an objective perspective. He supports the claim that the sound is coming from a specific point, by referring to this subjective “point” as “the body of the character who hears for us” (ibid.), and hence he avoids a discussion of POA’s relation to POV. He rather wishes to distance POA from its close connection to POV, and sees POA as something solely connected to subjectivity.

In his book *Audio-Vision – Sound on Screen* (1994), Michel Chion has a broader scope and embraces the idea of POA being a parallel to POV. He accounts for POV’s ability to be either subjective or objective, where the objective POV refers to the physical position or space within the diegesis, which the camera is catching, while the subjective POV is when the camera shows the diegesis through one character’s eyes and this character, to use a phrase inspired by Altman, “sees for us”. Chion argues that POV can be directly transferred to POA, and divides POA into a subjective category on one hand, and a spatial category on the other hand, which has an objective perspective. Chion and Altman agree that the subjective POA is focused on one character’s individual experience of a sound, and the spectators are invited to share this inner auditory perspective. Chion’s additional dimension, the spatial category, refers to a specific point in the diegesis that produces the sound. However, Chion questions the comparison between the auditory and the visual by claiming that where the audience is standing has more influence over what we see than what we hear, and he therefore proposes to replace the term Point of Audition with *Zone of Audition* (id., p.91) to avoid specifying a single point. Chion also emphasises that a visual representation of the character sharing their subjective POA is necessary for us to understand that it is in fact an internal auditor.

Birger Langkjær mentions POA in his book *Den lyttende tilskuer/The listening spectator* (2000), where he is critical of both Altman and Chion’s definitions, labelling as “illogical and inexpedient”. They are illogical, according to him, because they focus too much on the audio-visual aspects, rather than seeing POA as something strictly auditory, and inexpedient because a POA decided by the visual would mean that the audience’s POA would change in accordance with the visual editing. POA is therefore defined by Langkjær as “the audible distance between listener and the sound object’s source” (id., p.135). This distance can be understood as strictly spatial and objective, and does not seem to apply to Chion and

Altman’s subjective POA. He continues to critique both researchers for their lack of relevance for sound other than the mono-sound: Their definitions only cover the distance of the sound, and not the direction. This leads him to a conclusion that the definitions of POA are outdated, and that new technology such as surround-sound calls for a refinement of the term.

In his essay *The relevance of Point of Audition in Television Sound: Rethinking a Problematic Term* (2013), Svein Høyer agrees with Langkjær that new technology and the development of more complex sources for TV and film requires a refinement or a “rethinking” of POA, but offers an even broader understanding of POA’s issues. Høyer wants to create a discussion of POA that takes into consideration the development of sound perspectives and their relationship with the narrative, which has become more relevant with the arrival of new and more ambitious TV projects like HBO and Netflix. He agrees with Chion’s division of a subjective and a spatial POA, but takes it further by adding to each two subcategories to make the terminology more comprehensive and identifiable to researchers, students and other practitioners (ibid.). His categories are presented in the next section. Seeing as Høyer’s contribution is the latest published critical assessment, we can state that this is where the POA research field stands today. Because this thesis aims to explore the possibility of making POA less ambiguous by creating a broader understanding, it is natural to leave the theory of Altman, Chion and Langkjær behind, seeing as they have a more narrow scope than Høyer.

## 2.2 Høyer’s categories

**Table 1** Høyer’s four POA subcategories

Subjective		Spatial	
Individual POA	Personal POA	Active POA	Observational POA

Svein Høyer (2013) states in the introduction of his essay that the motivation for rethinking POA is “the analytical and descriptive problems raised when writing about recent trends in television sound in drama, sport, news, documentaries and other television genres”, and he argues that one can better account for these productions, which use sound in a “flexible and creative” way, by adding four subcategories, shown in **Table 1** above. Inspired by Chion’s

division into a subjective and an objective POA, he suggests distinguishing between *Observational* and *Active POA*, and between *Personal* and *Individual POA*. The following section presents these subcategories, and aims to clarify the boundaries between them, which will be thoroughly discussed in chapter 3.

Høyer's Individual POA describes situations where the audience gets access to a specific character's experience. This is more or less the same perspective as in Altman and Chion's subjective POA. A much-used example of this kind of POA is when a character has experienced an explosion and all they can hear is a high-pitched ringing noise. This experience is shared by the audience, who perceives the event from the character's perspective, and hears as if they were this character. A more complex example is when a soldier finds himself on the battlefield when all the noises from the battle suddenly disappear, and the audience only hears the breathing of this one character, and is forced to take part in this character's physical and psychological experience. Finally, a more common example is a telephone conversation where we hear both sides of the dialogue, even though the camera only shows one of the characters talking.

Høyer's new categorisation is derived from narratology and literary studies, and is presented by referring to Edward Branigan's (1992) eight narrative states with different so-called *focalisations*. Four of these states act as visual counterparts to Høyer's new four subcategories as shown in **Table 2** above. The key distinction between Høyer's Individual and Personal POA is the emphasis on surface versus depth, which we can find in Branigan's two categories of *internal focalisation*. Chion (1994) also suggests a similar separation, distinguishing between *objective* and *subjective internal sounds*, which highlights the same distinctions as Høyer's subjective categories. This might have been an inspiration for Høyer, even though Chion did not suggest this division in relation to POA.

**Table 2** Correspondence between Høyer and Branigan's categories

Svein Høyer	Edward Branigan
Subjective: Individual POA	Internal focalisation (surface)
Subjective: Personal POA	Internal focalisation (depth)
Spatial: Observational POA	Non-focalised narration
Spatial: Active POA	External focalisation

Where Individual POA is characterised by the experienced physical or psychological situations, Personal POA indicates access to the character's psyche on a much deeper level. Personal POA describes situations where for instance a character experiences a flashback or sound memories. A Personal POA can only be experienced by one character alone, whereas the Individual POA can in certain situations be the same for several characters, although the audience generally takes one character's POA at a time.

Turning to spatial POA, Høyer describes Active POA as situations where the audience have full access to the relevant sound in the diegesis, although only external sounds, making it correspond with Branigan's *external focalisation*. In TV-sound Active POA is best exemplified by talk-shows, where the guests are wearing microphones which lets the audience at home hear everything perfectly. In situations like this, sound often stays at the same volume, even though the camera offers different visual perspectives, both near and far away. By comparison I would suggest that a kind of Active POA also is evident in film scenes where we can hear a dialogue taking place at a bar counter, even though the camera perspective makes this impossible from a point of view perspective.

Observational POA is, on the other hand, when the viewers have limited access to the sounds in the filmic universe, a perspective often used in documentaries or reality shows. In these situations, the audio corresponds with the visual, which makes it more realistic. When used in film or television series, the audience might be surprised, as it deviates from the convention where the audience is used to getting full access, an Active POA. A popular play with these conventions are the so-called mockumentaries, where the audience are shut out audibly to create a realistic effect, as if it was a documentary team without access. Both Active and Observational POAs are relevant in discussions concerning realism and unrealism, as the former suggests an unrealistic relationship between audio and visual, while the latter tends to represent more realistic situations.

## **2.3 Subjectivity and objectivity**

Høyer's four categories are not able to cover all aspects, and they are not meant to, as his aim is to make the terminology more accessible (Høyer, 2013). This means that there are some grey areas left uncovered between his subjective and objective categories. Høyer

acknowledges this gap in coverage, and gives references to other auditory phenomena like *audio-visual dissonance* and *audio sweetening*. In the following section, terminology is presented that attempts to cover some of these grey areas, and at the same time they illustrate why POA can be considered problematic.

### **2.3.1 Audio-visual Dissonance, Fidelity and Synchronisation**

For Høyer (ibid.), his four auditory categories together with Branigan's four visual categories are important when talking about what he calls *the audio-visual interplay*. The tension between the audio and the visual, the POA and the POV, often create situations where there are discrepancies between the two of them, creating so-called "audio-visual dissonance," and sometimes these situations are perceived as quite unrealistic. When the image shows the same as what we are hearing, like the sound of a cat meowing accompanied by an image of a cat meowing, there is *fidelity* between the two media. However, if the sound of a cat meowing is accompanied by an image of a dog, there is an *infidelity*, and these types of situations are so striking that they are often used for comedic effect. If the image of the dog corresponds with the sound of a barking dog, sound and image are *synchronised*, but if the barking sound was heard after the dog had closed its mouth, we would talk about *a-synchronisation*, which leads to quite unrealistic and distracting experiences for the audience (Bordwell and Thompson, 2010). These are quite exaggerated examples, and there are many instances of discrepancies between sound and image which are not as noticeable. These can be examples of sound and visual showing different time zones, or different understandings of the diegesis, which will be elaborated on later in the thesis. Audio-visual dissonance and a-synchronisation can confuse our understanding of what kind of auditory perspective we are dealing with, and blur the lines between subjectivity and objectivity. This shows us POAs complexity and why the auditory phenomena can be perceived as problematic, along with the terminology trying to explain them.

### **2.3.2 Audio sweetening**

Høyer (2013) uses televised sporting events as examples of situations where there often is audio-visual dissonance. The POV is largely Observational, in order to show the audience

the whole action, while the POA is often Active. Initially, the POA can be said to be Observational as well, but to capture the atmosphere in the audience, the volume of the cheering and shouting is often increased, giving the television audience a sense of being in the stadium while seeing the whole picture. Høyer uses the concept *audio sweetening* to explain the effect of mixing pre-recorded sounds with live recordings to create a more “close-up sound”. The audience seldom questions this dissonance in perspective as it is done to give them the best experience, and we are quite used to this phenomenon through TV-debates and game shows as well. This effect can be taken even further, as Høyer shows by referencing to televised Norwegian ski-jump competitions. In these broadcasts, the audience has an Observational POV, together with an Observational POA, where the crowd can be heard in the distance. In addition to this, however, the producers add a “swoosh” sound at the same time as the competitors take off. This sound is so distinctive that you would have to stand quite close to hear it, giving us an Active POA, but only regarding this specific sound. The sound is added to enhance the excitement, and this has an effect on the audience’s perception, even though the sound is quite unrealistic.

Buhler et al. (2010) defines sweetening as sound that is “rendered so that it seems to violate the conditions of verisimilitude” (id., p. 86), meaning what the audience perceives as realistic and not. The practice of sweetening is closely connected with what is known in film studies as *Mickey Mousing*, referring to the technique often used in animation and cartoons, where non-diegetic music matches the visual movements, though this is often associated with comedy and exaggeration. The altering of sound is more evident when matched with a close-up of what is realistic, and it highlights the point that we are being made aware of something the characters in the diegesis are not. Buhler et al. also mentions the *stinger*, or *shock chord*, which is an audible metaphor for psychological states, such as an experienced shock. They distinguish between the internal and external stinger, depending on its motivation being subjective, coming from one character, or objective, referring to a common threat for instance.

### **2.3.3 Differentially audible sound**

Richard Raskin (1992) points out the need to distinguish between subjective sounds that only one character can hear, and objective, or *actual sounds*, heard from a character's

perspective. This he calls *differentially audible sound*, and a grey area between subjective and objective sounds. He uses an example where a character is wearing a helmet, and the sound diminishes. As opposed to Raskin's subjective category, *distorted sounds*, the sound stays the same. To quote Raskin; «what he hears is the same as what a microphone would pick up if it were placed inside his helmet» (id., p. 9). When translated into Høyer's POA categories, however, his distorted sounds would be Personal POA, and the differentially audible sound would be categorised as Individual POA, which means that, for Høyer, both of Raskin's categories are subjective, and according to Raskin, Høyer's Individual POA would be categorised as more objective than Personal POA. This is an example of theoreticians using different terminology to talk about the same phenomenon, but they disagree on what to categorise as subjective and objective.

## 2.4 Methodology and Thesis Structure

My interest for POA came from reading Høyer's essay, and my reaction inspired me to investigate his contribution further by comparing it to other terminology, and to test this on analytical cases which I suspected would not be covered by his categorisation. By looking at other contributions to the debate on POA, and reading theory that puts forward different terminology for explaining film sound, it became clear to me that the field is lacking a coherent theoretical framework. My methodological choices result from my discoveries, and based on what I regard as deficiencies in current scholarship regarding POA, the purpose of my study is to challenge perceived notions of POA, elaborate on established terminology, and to employ it in selected analysis.

To this end, the study encompasses as a theoretical, an analytical and an interpretive approach. The main method is critical reflections communicated through comparison, discussion and argumentation. The 'data' of the enquiry is on the one hand literature, and on the other scenes from film and television series which is interpreted by using the theoretical framework. Combining empirical and theoretical research allows for a constructive interaction which helps forming a nuanced understanding of the phenomenon. Using interpretation as an important part of my research allows me to recognize my own connection to the subject of investigation, but it requires a caution and critical attention, seeing as part of this approach calls for subjective knowledge.

The examples used in chapter 4 are chosen because of the auditory perspectives presented, which can be problematic to describe using established terminology. I have included examples from both film and television series, as well as both the drama and comedy genre, to show that the challenging auditory perspectives presented in order to challenge the established terminology are extensive, and not narrow exceptions from the rule.

This chapter has been the theoretical backdrop which has set the scene for this thesis' following discursive investigation. In the next chapter I will introduce David Neumeuer's categories, which will be important for the discussion and the analysis to come. The choice of including Neumeuer's categories and comparing them with Høyer's is based on them both having quite elaborated theories with a certain amount of depth, and their contributions are relatively new; Høyer in 2013 and Neumeuer in 2015. Neumeuer's categories are more complex than Høyer's, and acquires some explanations and elaborations. This will be the presented in chapter 3, along with a comparison of Neumeuer and Høyer's categories. At the end of chapter 3, a new schematisation will be made, based on a compilation of their two sets of categories. This will be the foundational terminology used in the analysis in chapter 4, where different examples of challenging auditory perspective will be discussed. Finally, a summarisation of the findings and concluding notes in chapter 5, which puts this thesis in a larger perspective, hoping that its contribution can help stake out the course for further research.



## 3. Further theoretical contexts - Getting beyond Høyer

---

There has not been any contribution criticising or commenting on Høyer's new categorisation, so my opportunity to investigate further is based on the status quo, presented in chapter 2. In addition to this, a discussion has briefly been started, surrounding the grey areas between auditory perspectives perceived as either subjective or objective. These are not covered by POA, and my claim is that these phenomena are not exceptions from the rule, but fairly common techniques used by film creators, which therefore should be covered by the terminology. POA's exclusion of non-diegetic sound and music is also interesting, seeing as the non-diegetic aspect of film and series is quite prominent. The non-diegetic will therefore be one of the main focus points in chapter 3, functioning as an introduction to the presentation of Neumeyer's categories, which has been absent up until this point in the thesis, because of their treatment of auditory aspects that lie beyond the scope of POA.

Leaving Høyer and POA for a moment, the two following sections aim to present a different approach to auditory perspectives. By introducing terminology that includes an exploration of the boundaries between non-diegetic and diegetic sound, a greater complexity is revealed, showing aspects which POA, at its current state, is unable to cover. Neumeyer's (2015) theory takes this into consideration, and therefore serves as a suitable compliment to Høyer's theory. These two sets of categories will be united at the end of this chapter after a review of Neumeyer's categories and a comparison of these with Høyer's.

### 3.1 Diegetic and Non-Diegetic sound

For film music scholars, the difference between what is called diegetic and non-diegetic sound is crucial to the understanding of a film's narrative and aesthetic expression. Even so, the two terms have been subjects in most film music debates, concerning how to distinguish between them, but also whether it is aiding or restricting the analysis of film music (Neumeyer, 2009). However, some of the most influential film sound theoreticians and

sources, Claudia Gorbman (1987) and Bordwell and Thompson (1985), agree on the basic definitions of these two terms. Gorbman defines *diegesis* as “the narratively implied spatiotemporal world of the actions and characters” (Gorbman, 1987, p. 21), and *diegetic music* as “music that (apparently) issues from a source within the narrative” (Gorbman, 1987, p 22). This is described more simply by Bordwell and Thompson as sound where the source is “a character or object in the story space of the film” (Bordwell and Thompson, 1985, p. 191), which makes non-diegetic sound the opposite, or as Chion puts it: “sound whose supposed source is not only absent from the image but is also external to the story world” (Chion, 1994, p.73). Film music theoreticians agree to a certain extent on these definitions of diegetic and non-diegetic sound, where sound coming from the filmic universe is categorised as diegetic, whether it is onscreen or offscreen, and sound coming from outside the filmic universe is non-diegetic sound (Bordwell and Thompson, 1985). This understanding is fairly intuitive also from the spectator’s perspective, and it is presented quite explicitly by the film-makers. When the lines between diegetic and non-diegetic gets blurred, it is usually an effect meant to confuse, scare or amuse the audience, and this is very effective because a crossing of these borders challenge our conception of what is supposed to be realistic and not.

If focusing only on the term POA, the distinction between diegetic and non-diegetic sound is not a prominent issue, seeing as POA is described by all theoreticians using the term as something solely diegetic. Høyer excludes from his essay non-diegetic sounds, music and voice-overs along with diegetic voice-over, except for the subjective inner voice, without commenting on whether he thinks POA should have a non-diegetic aspect. Other theoreticians writing about film sound, however, use both diegetic and non-diegetic as important concepts in their research, as will be evident in the next section. The expansion of a non-diegetic dimension of POA could in fact be more fruitful to the discussion of auditory perspective, and it is worth investigating further.

### **3.1.1 Simple and Displaced**

Previously, we saw what happens when the synchronisation between audio and image is tampered with, and creates what Høyer (2013) calls audio-visual dissonance. Synchronous and asynchronous sounds are a part of what Bordwell and Thompson (1985) call the

*viewing time*, the time it takes to watch the movie. Bordwell and Thompson distinguish between viewing time and *story time*, the time that presumably passes in the film's diegesis. Whereas the viewing time of an average movie is approximately two hours, the story time can have a duration of days, weeks, months or years. The story time can also be divided into subcategories depending not on the synchronisation, but the simultaneousness of the stories told respectively by the audio and the image.

*Simultaneous sounds* take place at the same time as the story events shown by the visual, what Bordwell and Thompson (ibid.) call *simple diegetic*. These sounds can either be external, like spoken dialogue between characters or other *actual sounds* in the diegesis, or internal, presenting a character's thoughts in the given situation. Simple diegetic is the most straightforward phenomenon, but *non-simultaneous sounds* are also quite common. This they call *displaced diegetic*, and the sounds can either be placed before or after the story's events are shown visually. Starting with early placed sounds, they can also be either external or internal. Internal early sounds can be sound flashbacks, representing a character's memory of an incident, while the image stays in present time. These flashbacks can also be external, if they represent something that has happened earlier in the film which we have already seen, but they are external because there are no characters remembering them. Sounds placed after the visual representation can also be internal and presented as a character's thoughts in future time, or external, if for instance a character is telling a story about incidents in the past, while the image is showing these incidents.

So far, we have only covered diegetic sounds, but according to Bordwell and Thompson (ibid.) non-diegetic sounds can also be displaced. Their examples of non-diegetic sounds that are connected to events that take place before the visual, include a speech from Winston Churchill together with images of Britain today. Similarly, non-diegetic sounds that happens later than what the image is showing, can for instance be a non-diegetic narrator talking about events in the past, shown by the image. They emphasise that most non-diegetic sound, especially "mood music", does not have a temporal relationship with the story, and it is therefore not relevant, or possible, to define whether there is simultaneousness. However, when it comes to voice-over narration, this is highly relevant, and something I will return to.

### 3.1.2 Empathy and Anempathy

Another aspect of diegetic and non-diegetic sounds' relationship with the story concerns the attitude towards the action that the audience experiences through sound and music. Chion (1994) distinguishes between *empathetic sound* and *anempathetic sound*, which can be both diegetic and non-diegetic. The most obvious of these sounds are non-diegetic and empathetic, and we know them as "mood music" that underlines the emotion portrayed by the image, e.g. horror, romance, sadness or action to name a few. Diegetic sounds can also be empathetic towards the action, but seeing as the diegetic sounds are supposed to be realistic in that they are present in the diegesis, this can create a comedic effect, because we are not expecting the natural sounds in the environment to behave in this way. An example of diegetic empathetic music can be a random song on the radio expressing the exact feelings of the character at the given time.

According to Chion, anempathetic sounds are usually diegetic, and that is understandable giving that diegetic sounds have no ability to empathise: the objects in the diegesis are by default indifferent to the action that surrounds them. Nonetheless, this can become a powerful commentary. A much-used example of anempathetic diegetic sounds is the shower-scene in Hitchcock's *Psycho* (1960), where the shower keeps running at the end of the scene as if nothing has happened. This of course is the only realistic outcome, but before this, there has been highly empathetic music underlining both the horror of the event, and the actual stabbing motions of the killer's hands. When the music fades, we almost expect sad music, or at least a cut to the next scene, but Hitchcock leaves the audience in this scene together with the victim for a while, and the sound of the running shower accentuates the feeling of cruelty by not "caring". Non-diegetic music can also be anempathic or indifferent, but it often offers a comedic function, or at least a disturbing function. If the theme or the mood of the non-diegetic music does not match the visual events, it can have altering effects on the audience's perception.

In her article *The Fantastical Gap between Diegetic and Nondiegetic* (2007), Robynn Stilwell draws an interesting connection between subjectivity and objectivity on the one hand, and empathy and anempathy on the other. She points out that the element separating them is that they "diverge from a single point, the point of view/audition/feeling of a

character in the diegesis” (id., p. 191). Stilwell describes empathy and anempathy as the relationship between a character and the audience, and how the audience through music and sound decide whether they either recognise, are indifferent to, or reject a character's feelings. We are used to being steered towards certain emotions by non-diegetic music, and this greatly influences our opinions of the different characters. The fact that non-diegetic music generally functions this way, makes it a convention that we tend not to question.

“The alliance of empathy with the underscore and anempathy with source music is certainly prevalent in the classical Hollywood aesthetic, and it is still a dominant mode of scoring” (id., p. 190). As Stilwell points out, anempathy can indicate indifference or objectivity towards a character or the events, but it has often the ability to create a rejection of the same actions within the audience (ibid.). In the same way, as empathetic music can create strong feelings, the absence of these expected reactions, often guided by non-diegetic emotional music, often creates greater emotions because it deviates from the conventions, and it underpins more strongly the harshness of the events. When we consider situations with anempathic music or sound, the stark contrast makes the empathetic music's function more evident, and we might see reasons for distinguishing between subjective and objective non-diegetic sound and music as well.

### **3.1.3 The metadiegetic and The Fantastical Gap**

Based on the established definitions of diegetic and non-diegetic sound, we find many situations where the lines between the two are blurred. Claudia Gorbman (1987) mentions situations where a character's subjective memory triggers non-diegetic music, a phenomenon she calls *meta-diegetic*. Gorbman is borrowing the term meta-diegetic from Gerard Genette (1980) who distinguishes between three diegetic levels of narration; *diegetic*, *extra-diegetic* and *meta-diegetic*. In addition to these levels, he distinguishes broadly between two different types of narrators; the *homo-diegetic* narrator, which is a character in the diegesis, and the *hetero-diegetic* narrator, which is not a part of the diegesis. In Genette's sense of the term meta-diegetic, a character in the diegesis becomes a homo-diegetic narrator by telling a story which becomes “a story within the story.” Translated into film sound theory, this character takes control over the narration, determines and takes over the non-diegetic aspect. On one level, the sound in these situations is very much a part of

the diegesis, seeing as one of the characters is thinking about it in his mind, but this can at the same time be presented as non-diegetic sound or music to the viewer. This indicates a grey area not only between subjectivity and objectivity within the diegesis as mentioned earlier, but between subjectivity in the diegesis and the non-diegetic, which is a much more complicated phenomenon than Høyer's categories are able to cover.

Robynn Stilwell (2007) suggests a refinement of the term metadiegetic, focusing on the distinction between empathetic and subjective when differentiating between non-diegetic music and Gorbman's meta-diegetic music. This differs from Høyer's subjective POAs as it is clearly non-diegetic music, and not supposed to be understood as music only happening in a character's mind. Either the music is triggered by the character's memory, or the music seems to be objective until we understand that it has a subjective function. Stilwell (ibid.) tackles this ambiguity between diegetic and non-diegetic as a trajectory, or a place of destabilisation, and together with James Buhler in a film music roundtable for *The Velvet Light Trap*, they proposed the term *Fantastical Gap* to describe the abilities of this concept. This seems a fitting term, as the word *fantastical* captures the unrealistic, and has connotations to supernatural or fantasy, placing us in the world of the *unrealistic*. Stilwell proposes that the strategy of using this Fantastical Gap to draw the audience towards a character should include axes of multiple binaries, including empathy/anempathy, subjectivity/objectivity and the aural perspective, described by her as the opposition here/there. Stilwell notes that non-diegetic music tends towards subjectivity, and diegetic music tends towards “a kind of realistic objectivity” (id., p. 191), but she separates these terms from empathy and anempathy, as the latter concerns the audience's perception and response to a character, whereas the former are more technical terms reflecting the filmmaker's audio-visual choices. This approach seems to be an attempt to show a more holistic view on complex auditory phenomena, a view that dares to reach beyond the strict categorical approach often taken by theoreticians trying to grasp the concept of POA.

### **3.2 Neumeyer's categories**

As we have seen above, theoreticians discussing POA, including Høyer, do not account for the non-diegetic aspect of sound perspectives. David Neumeyer (2015) discusses sound perspectives with a terminology that cover some of the same phenomena as Høyer's

categories, but he does not use the term POA. Neumeyer emphasises two aspects that differs from Høyer: One is the non-diegetic dimension, and the other is the sound's relationship to what is explicitly and implicitly shown visually, namely if the sound can be said to be *onscreen* or *offscreen*. In this section Neumeyer's categories are presented, discussed and critiqued, and they will together with Høyer's categories forms the basis for this thesis' analysis and discussion.

### 3.2.1 Combining narrative levels with framing

In David Neumeyer's *Meaning and Interpretation of Music in Cinema* (2015), he proposes a set of five binaries as tools for analysis and interpretation. The first two are *clarity*, the prioritising of intelligibility and efficient presentation of the narrative, and *fidelity*, the prioritising of the accurate and realistic presentation of the narrative; they are related to the balance between the narrative levels. The second two are *foreground* and *background*, which refer to the mixing of the sound track elements influencing the degree of attention the sound draws in comparison with the visual. Then there is the category of *diegetic* versus *non-diegetic*, which touch upon the placement of the sound, either within the diegesis or on the outside of the narrator's level. The fourth binary tackles *synchronisation* and *counterpoint*, which portray, respectively, a realistic and non-realistic relationship between sound and image with regard to both time and logic. Lastly, we have *empathy* and *anempathy* where the coordination of sound and image are either sympathetic, or distanced or ironic (ibid.). See **Table 3** below for a schematic overview of Neumeyer's binaries.

**Table 3** Neumeyer’s “five binaries” with descriptions (Neumeyer 2015, p.63)

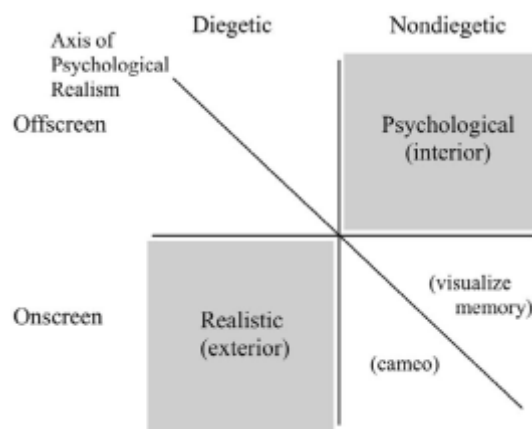
<b>Terms/oppositions</b>	<b>Description</b>
<i>(principle: balance of narrative levels)</i>	
1. Clarity	1. Priority to efficiency in presentation of narrative
2. Fidelity	2. Priority to accurate reproduction of the acoustic environment of the physical world depicted in the narrative
<i>(method: mix and balance of sound track elements)</i>	
1. Foreground	1. Sound that is forward in the mix, or a sound track that draws attention compared with the image
2. Background	2. Sound that is low in the mix, or a sound track that draws little attention
<i>(space: sound in relation to space/levels of narration)</i>	
1. Diegetic	1. Sound belonging to (anchored in) the physical world depicted in the narrative
2. Nondiegetic	2. Sound belonging to the level of narration or the narrator
<i>(time: temporal and logical coordination of sound and image)</i>	
1. Synchronization	1. Close, realistic coordination of sound and image
2. Counterpoint	2. Nonrealistic treatment of sound/image simultaneity
<i>(agency: emotional coordination of sound and image)</i>	
1. Empathy	1. Close, sympathetic coordination of sound and image
2. Anempathy	2. Ironic or distanced treatment of sound in relation to the image

Further on in his discussion, Neumeyer investigates the relationship between, on the one hand, diegetic versus non-diegetic sound, and, on the other hand, onscreen versus offscreen sound. By combining these four aspects, we get four combinations of audio-visual expressions which, as will be evident through this thesis, can be quite problematic as well. The first combination is *onscreen diegetic*, which seems to the audience as the least problematic, having synchronised sound and picture, and disclosing a straight forward narrative. It also has objective POA and POV, giving us only the *exterior* of a character’s persona, which appears as a more realistic portrayal. The second combination is *offscreen non-diegetic*, which is also relatively unproblematic due to the established cinematic conventions, leading the audience to accept sound and music appearing without a specific source. There are many examples of this phenomenon, ranging from the subjective perspective of a character, displaying a psychological insight, to the *voice over narration* which can take many forms in terms of narrative levels. Neumeyer draws what he calls *the axis of psychological realism* between these first two pairs, as shown in **table 4**, and notes that:



“most dramatic narrative film presupposes this axis as a normal state, which is one reason it seems so unproblematic to an audio-viewer. The normative state of this axis is also the primary reason that digressions from it can be so powerful: they can be used to channel desire, either ours, as a part of the narrative process, or a given character’s, as a representation of that character’s desire. Alternating between these options creates a particular kind of fluctuation along the axis of psychological realism.” (Neumeyer 2015:80).

**Table 4      The four quadrants divided along an axis of psychological realism  
(Neumeyer 2015:80)**

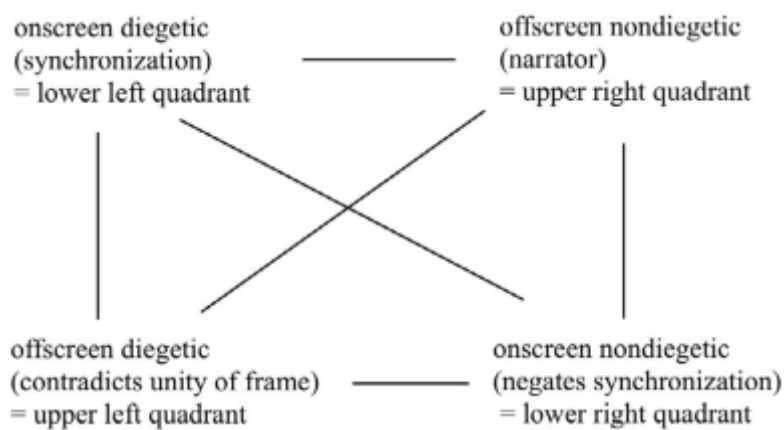


The third combination is *onscreen non-diegetic*, which is considerably more complex than the former, given the ambivalence afforded by “visualising diegetic and non-diegetic space simultaneously” (id., p. 81). Neumeyer distinguishes between onscreen non-diegetic understood as exterior and interior, where the former is typically an insert, a picture-in-picture or a brief comment on the narrative, whereas the latter is a direct visualisation of a character’s feelings or memories. He notes that this phenomenon is rarely used, a statement I will come back to, and if so it is usually part of a comedy seeing as it has a certain awkwardness to it (id., p. 276),

The last combination is *offscreen diegetic*, which can be unproblematic if understood as exterior, portraying neutral background sound or music which is a realistic and plausible

part of the diegesis. It is on the other hand more complex if the offscreen diegetic sound is understood as interior, where the sound is diegetic but at the same time a part of character's mind. Neumeyer exemplifies this by a case from the movie *Sons of the Pioneers* (1942) where two of the characters imagine their dead friend singing along with them, leaving the voice of their friend to initially be offscreen non-diegetic, but in their heads. However, it turns out that their friend is not dead after all, and is present, singing diegetic. This leads to a comedic event where the two of them slowly realises that the sound is not coming from their heads, and gets frightened because they cannot see their friend, as he is hiding (Neumeyer, 2015, p. 99). This is a rare example and, similar to onscreen non-diegetic, it is, according to Neumeyer, mostly used for comedic effect. Whether interior or exterior, the offscreen diegetic “undermines the distinction of narrative levels that is the particular province of the offscreen non-diegetic at the same time that it wrestles the realism of synchronisation by expanding diegetic space, threatening the camera's authority, so to speak, by contradicting the unity implied in framing” (id., p. 98). For a schematic overview of the relationship between these four pairings, see **table 5** below.

**Table 5 The relations of the four quadrants (Neumeyer 2015:82)**



### 3.3 A comparison of Høyer and Neumeyer's categorisations

Neumeyer (2015) and Høyer (2013) agree on some aspects and disagree on other. They complete each other on yet some, but they both fail to cover certain grey areas that are fairly common in modern film and television. In the following I will present what distinguishes

and unites them, before presenting a new categorisation combining their categories. The new categorisation will then be the foundation of the analysis.

### 3.3.1 Comparing Høyer and Neumeyer's categorisations

Høyer (2013) and Neumeyer (2015) base their theories mainly on the same sources, namely Gorbman (1987), Chion (1994), Altman (1992), Branigan (1992) and Raskin (1992). They do not, however, acknowledge each other in their articles, they therefore do not comment on each other's findings, and are not directly influenced by each other. There is of course a chance that Neumeyer might have read Høyer's contribution, given that it is published two years before Neumeyer's latest contribution in 2015, but there are no direct references to Høyer, and POA is not discussed explicitly in Neumeyer's work.

Firstly, and most importantly, the main difference between their two categorisations is the fact that Neumeyer take into consideration non-diegetic sound and music, while Høyer does not. This makes their models impossible to intertwine completely, but this is also what makes their two categorisations, seen in coherence, so interesting. The fact that Høyer excludes the non-diegetic aspects from his study might reveal a disagreement about the categorisation of non-diegetic and diegetic, and opens up for a discussion of the grey area between these, which Stilwell (2007) calls *The Fantastical Gap*. The second important difference between their categories is that Neumeyer emphasises the audio-visual aspect to a greater extent than Høyer, as his categories highlight the offscreen/onscreen dimension. This is what makes POA distinctive, the debate on whether it is useful to separate sound theory and film theory to a certain extent, in order to give justice to the independent science and art form that film sound is, a view represented in this thesis by Langkjær (2000). The third important difference is the fact that Neumeyer does not mention the debate surrounding POA, nor uses this term in his work. However, the basis of his theory does very much lean on the key-concepts linked to POA, namely subjectivity, objectivity, exterior, interior and diegetic. What he brings to the POA discussion, without taking an active part in it, is the realistic and unrealistic/psychological dimension, which has not been part of the debate. Seeing as Høyer does not mention the non-diegetic aspect, I will let Neumeyer's categories be the starting point of the comparison, and discuss Høyer's categories in relation to these.

### 3.3.2 Neumeyer's onscreen diegetic

Neumeyer's onscreen diegetic corresponds with Høyer's Observational POA in several ways. Neumeyer ascribes the following words to this concept; *synchronisation, realistic, objective, exterior and anti-psychological*. Høyer notes that with the Observational POA the audio matches the visual, thus ensures synchronisation, which leads to a more realistic perception. Høyer places the Observational POA under what he calls spatial POA, as opposed to subjective POA. Even though he uses spatial instead of objective as the overarching category, he mentions in parenthesis that spatial POA is more objective than subjective POA is (Høyer, 2013). Neither active nor Observational POA depict a character's psyche, which make them exterior, and hence also categorisable as anti-psychological. This is quite straightforward when dealing with Observational POA, but gets somewhat more complicated when trying to match Høyer's Active POA with Neumeyer's categories. The aim of Høyer's Active POA is to explain the phenomena where the audio does not match the visual. Even though Active POA is objective and depicts the exterior, and is therefore anti-psychological, it fails to fulfil the aspect of synchronisation, which leads to lack of perceived realism. We can therefore conclude that with his Active POA, Høyer covers an aspect that Neumeyer fails to acknowledge with his onscreen diegetic.

### 3.3.3 Neumeyer's offscreen diegetic

Neumeyer's offscreen diegetic is divided into two subcategories: Interior and exterior. The exterior aspect is unproblematic, since the sound must be anchored somewhere in the diegesis for it to fulfil the criteria of offscreen diegetic, like acousmatic voices, background and environmental sound, or music played from a known source. Høyer does not mention this, but it is reasonable to think that he considers this type of sound to have a spatial POA, that is either an Observational or an Active POA, because of the realistic aspect. This example highlights Høyer's exclusion of the visual aspect, because the determination of POA concerns the position of the microphone more than where the camera is looking. To clarify: The POA does not change in line with the POV. We can therefore conclude that Observational POA matches with both onscreen and offscreen exterior diegetic.

The offscreen interior diegetic can be tricky to make sense of, in that it must not be confused with offscreen non-diegetic. To avoid that, we need to define the boundaries between the two, which we have seen can be quite ambiguous. The commonly acknowledged difference between diegetic and non-diegetic is whether the source of the music can be located within the diegesis, and for now, in the case of offscreen interior diegetic, that will be sufficient. Høyer's Individual POA seems to go well together with Neumeyer's category, as the individual perspective of a character is still what Raskin (1992) calls actual sounds that exist in the diegesis. For them to be perceived as individual, however, the volume and/or quality of the sounds differ from the purely objective way of experiencing them, and Raskin therefore calls them differentially audible sounds. Examples of these kinds of sounds can be intense breathing from a person under stress, as heard from the character itself, or the loud sound of a wall clock ticking, as experienced by a character, both examples of sounds that can be heard objectively, but are perceived quite differently.

Neumeyer uses an extraordinary example to highlight the ambiguity of the offscreen interior diegetic, as previously referenced to, which serves as a bridge to the discussion of the onscreen non-diegetic. In the example from *Sons of the Pioneers* (1942) the characters think they are hearing the voice of their dead friend in their mind, remembering him from the time they used to sing together, which makes it a sound memory, something Høyer would classify as a Personal POA. But the sound turns out to in fact be diegetic, as the perceived dead friend is actually alive, making the audio both a sound memory and diegetic. The problem is however that if the sound were not anchored in the diegesis, it would be only in their minds, a phenomenon Neumeyer fails to address. Høyer, however, places all interior sounds as diegetic, and the Personal POA matches therefore with the offscreen interior diegetic in cases where, as in the example from *Sons of the Pioneers*, the mind of a character is depicted at the same time as it is anchored in the diegesis. It is unclear whether Neumeyer would consider unanchored interior sound as diegetic, as this is not exemplified in his work, and it leads us therefore to question what "anchored in the diegesis" actually entails. It is unclear whether Neumeyer would consider unanchored interior sound as diegetic, as this is not exemplified in his work, and it leads us therefore to question what "anchored in the diegesis" actually entails. My understanding is that Neumeyer interpret the inside of a character to be understood as a point in the diegesis to which sound can be anchored.

### 3.3.4 Neumeyer's onscreen non-diegetic

This leads us to Neumeyer's onscreen non-diegetic, which he has also divided into an interior and an exterior dimension. The interior aspect covers every direct visualisation of memories or thoughts. The fact that it is onscreen, and so commonly used, makes it quite straightforward. But a visualised memory with sound is also covered by Høyer's category Personal POA, which we know is supposed to be diegetic. It is distinguished from Individual POA in the way that it is not actual sound, and, according to Neumeyer, it should therefore be put in the non-diegetic category. Hence, the two theoreticians both cover the phenomenon in their theories, but they disagree on the labelling. The exterior aspect, however, is quite complicated in that the sound needs to be onscreen, directly linked to a visual image, not a part of someone's mind, and at the same time be non-diegetic. For this we need to leave Høyer's categories for now, as he clearly states that his theory does not cover non-diegetic sound and music. Neumeyer claims that examples of exterior onscreen non-diegetic is rarely used in mainstream cinema, and when it is used it is treated in a picture-in-picture manner, as an insert, or represented as a brief narrative intrusion or a comment, like a cameo in the corner of the screen (Neumeyer 2015, p. 80). Neumeyer does not go further into this, but this phenomenon has become more common in modern films and TV-series than we would think based on Neumeyer's description of it. I will come back to this in my discussion concerning Stilwell (2007) and her *Fantastical Gap* in the analysis.

### 3.3.5 Neumeyer's offscreen non-diegetic

Finally, we have the offscreen non-diegetic, which is by Neumeyer described as fairly straightforward as well, given the cinematic conventions resulting in the audience accepting sound without apparent sources. He places this category on the upper right in his diagram, leaving it on the far end of the axis of psychological realism. It is also given the subtext "interior" in parenthesis in his figure 2.10 (Neumeyer, 2015, p. 80), which indicates the strong focus on the subjective. When giving examples, however, in addition to non-diegetic music, which obviously can be very representative for a character's psyche, he mentions voice-over narration as a part of this sphere.

In order to highlight the small differences between the psychological and the commenting

aspect, and to make room for all the different types of narrators, I would suggest dividing also this category into two subcategories: Subjective and objective. In the subjective category I would first and foremost place non-diegetic music, which either represents the subjective perspective of a certain character, or gives insight into a character's feelings. Music with an empathetic attitude also fits this category. In the objective category I would place the voice-over narrators without an anchorage in the diegesis, the so-called heterodiegetic narrators. The homo-diegetic narrators would also be in this category until they eventually appear to be anchored in the diegesis, as this would entail a change in the audience's perception. In addition, I would add non-diegetic music with an objective attitude, which would be music that provides either a geographic placement, a temporal placement, or a general mood to the universe without being biased or influenced by a character's point of view.

### 3.4 Compilation of Categories

Before embarking on the analysis, it is worth summarising and schematising the different categories that have been dealt with so far. **Table 6** below show Neumeyer and Høyer's categories together, leaving out Høyer's Active POA, which I will come back to. Neumeyer divides two of his categories by *the axis of psychological realism*, to indicate what is considered exterior and interior. This concerns offscreen diegetic and onscreen non-diegetic, but is not mentioned in the two other categories. Above I suggested dividing offscreen non-diegetic into two as well, and this division is added to the table, as this aspect will be further discussed in the next chapter. Neumeyer's final category, onscreen diegetic, has only been presented as external so far, but as I will show in the following sections, it might be appropriate to divide this into an external and internal sub-category as well. In order to complete the table and as support for the analysis, I have added this division. By adding these categories I am able to cover some of the grey areas presented in the following analysis, and by adding on to existing categories, it becomes apparent that their categories are not far away from covering such phenomena.

**Table 6**      **Compilation of categories**  
**Diegetic**

**Non-diegetic**

<p>Offscreen diegetic Interior</p> <p>Individual POA Personal POA</p> <p>Offscreen diegetic exterior</p> <p>Observational POA</p>	<p>Offscreen non-diegetic exterior (added by me)</p> <p>Offscreen non-diegetic interior</p>
<p>On-screen diegetic interior (added by me)</p> <p>On-screen diegetic exterior</p> <p>Observational POA</p>	<p>Onscreen non-diegetic exterior</p> <p>Onscreen non-diegetic interior</p>



## 4. Analysis

---

Seeing as modern television shows have been more and more legitimised and equated with cinema, and since they use the same technological methods and rhetorical tools for building narratives (Newman and Levine 2012), there is no reason for distinguishing between the two media in this discussion. In the following I will therefore reference to scenes from both cinema and television. In part one I use the television series *13 Reasons Why* (Yorkey, 2017) as a main analytical object, because it presents good examples of auditory situations which can be problematic to describe using the established terminology. The following examples aim to show the gaps in the existing terminology in practice, to get a better understanding of the difficulties of applying sound perspective terminology. The second part uses the findings in part one, and takes it even further. Considering the terminological challenges faced in part one, it becomes apparent that the adjustment of the categories in **Table 6** are not sufficient, and the adding of Stilwell's (2007) Fantastical Gap helps situate Høyer's Active POA together with the established terminology, in order to create a complete framework.

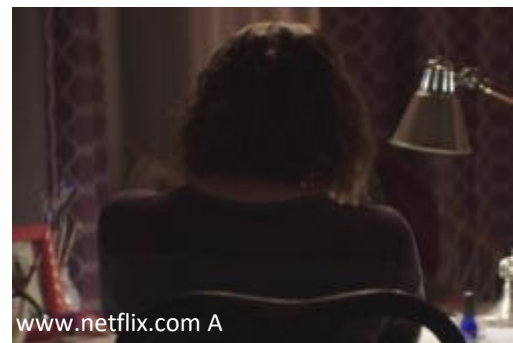
### 4.1 Part one: Several different POAs at the same time

The series *13 Reasons Why* (Yorkey, 2017) shows an example of a voice-over narrator who switches between being treated as diegetic and non-diegetic. This is used in very creative ways, which violates all rules of synchronisation and reality. The main character Hannah Baker has taken her own life, but before she did, she recorded a series of cassette tapes where she explains her thirteen reasons for taking her life, or the thirteen persons who influenced her choice. These cassettes were handed over to one of her classmates a few days after she died, and this classmate were to listen to all the tapes before handing them over to the next person. Each episode in the series represents one side of a cassette, or one of the thirteen reasons for Hannah to kill herself. We follow one of her classmates, Clay, as he listens and reacts to the tapes. The series thus gets a natural narrator, which is the recorded voice of Hannah, but because this is a television series and not an audiobook, the creators have been very creative in their use of this narrator. These creative techniques make it difficult to draw general conclusions about Hannah Baker's position as a narrator, and it is

therefore worth examining further. How are we supposed to understand the sound perspectives, what does it do with our experience of the narration, and what terminology and categories are suitable for analysing this?

#### 4.1.1 Defining the narrator's POA

begins in the same way as the first cassette tape in the diegesis, namely by Hannah introducing herself, and asking the listeners of the cassettes, and the audience, to settle in and prepare for a long story. If we base this analysis on a first time viewing without any knowledge of the series plot, Hannah's voice is in this moment off-screen non-diegetic exterior, and as it becomes clear that it is her life story told by herself that we are about to hear, we know that this narrator is homo-diegetic. The next time we hear from her is when her classmate and friend Clay, the series other main character, listens to the first tape (12:48). In this moment, we understand that not only was the introductory narration the same as the opening of the first tape, but given that we during the first ten minutes of the first episode learn that Hannah has taken her own life, and is no longer physically in the diegesis, the narrating voice of Hannah will exclusively be what she recorded on the tapes. Here we are faced with the dilemma of what to categorise as diegetic and non-diegetic, because the second time we hear Hannah, the voice is on-screen via a cassette player, and therefore very much a part of the diegesis, even though the owner of the voice technically is not anymore. Also shown is a brief visual of Hannah recording the tape (13:12-13:19), sitting in front of a desk talking into the microphone; a situation that has happened in the diegesis earlier, before the audience was invited to take part in the story. Having this in mind, it seems unspecific to categorise her as a non-diegetic narrator. Seeing this from a POA perspective, it is clearly Individual as we see Clay listening to the tapes, and our perception of Hannah's voice is the same as his.

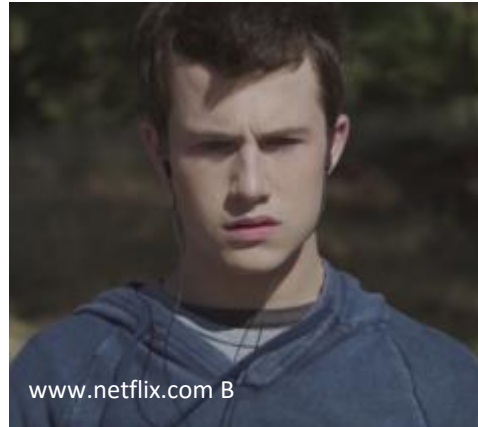


When comparing Høyer and Neumeier's categories, I conjoined Individual POA with Neumeier's off-screen diegetic interior, and it would probably be right to categorise Hannah's voice as off-screen in the cases where we do not see her speaking. However,

seeing as the sound of her voice has a clear source, namely Clay's Walkman, I would argue that Hannah's voice, in situations where the source is visible, is on-screen. Seeing this in relation to the interior aspect of subjective POA, I believe it would be more accurate to call this *on-screen diegetic interior*, as suggested by me in **Table 6** (p.30).

The narrator's voice also has the ability to bridge the diegetic and the non-diegetic universe by functioning regularly as an off-screen non-diegetic narrator. Whether Hannah's voice can be understood as non-diegetic depends on the narrative and aesthetic function it serves, but it is of course questionable if there should be an opportunity to define something clearly diegetic as non-diegetic for a period of time. Take for instance the following example from the opening scene of the movie *Guardians of the Galaxy* (Gunn, 2014) where, fittingly enough for our case, the main character puts on a Walkman and we share his Individual POA. As he dances around to his music, the title of the movie is shown in big block letters over the screen, and by default his diegetic individual music now serves as non-diegetic theme-music for the movie. This is often used for comic relief by surprising the audience who think they have figured out the POA, and it can go both ways, but it is maybe more effective when music you think is non-diegetic becomes diegetic. A great example of this is in the series *Sherlock* (Gatiss & Moffat 2010 -) where the villain Moriarty flies in on a private jet accompanied by a non-diegetic playing of *I Want to Break Free* (1984) by Queen (35:41-36:34). However, when he descends from the plane, he plugs out his headphones and we realise that we have shared Moriarty's Individual POA all the time. Seeing as Moriarty is a terrifying psychopathic villain, this is not someone you want to identify yourself with, so this creates a quite chilling effect, where the audience might feel tricked into empathising with him. In such situations the music can be said to function as both diegetic and non-diegetic at the same time. However, Jeff Smith (2009) argues that non-diegetic sound that turns out to be diegetic, has in fact always been diegetic; the film-makers have only withheld this information from the audience. I agree with him regarding situations where the sound quality and setting does not change, and where the audience has simply been fooled, but in some cases the sound quality changes, which makes this argument harder to sustain.

Returning to our case, we get an example of this later in episode 1 (18:49-20:17) when Clay is sitting on a bench listening to tape number one on his Walkman, looking out over the city.



Hannah explains that there is a map of the city attached to the cassettes, in order for the listeners to be able to visit the places of significance that she points out. Clay does not have this map with

him, and decides to ride his bike home to get it. He continues to listen to the tape as he picks up his bike and rides out of the frame. The camera stays on the view of the city, and day turns to night while the tape with Hannah's voice is still rolling. The tape hence functions as a bridge between the shots, and in the next frame Clay is riding his bike in the dark, still listening to the tape. It is significant that the tape has not made the same jump in time as the visual, and the audio is therefore making a jump from a realistic POA to an unrealistic POA to serve as an aesthetic element, and to ensure continuity.

Referring to **Table 6**, the sound moves between the on-screen interior diegetic and the non-diegetic off-screen objective. It is worth reminding that both of these categories are suggested, or at least refined by me, and that Neumeyer and Høyer's initial categories do not cover this phenomenon.

On-screen interior diegetic was added by me above in order to account for auditory perspective like this. Seeing as I argue that the source of the sound in this series is Clay's Walkman, it therefore is a part of the diegesis, and this shows that my category is necessary. The category fits with what Raskin (1992) calls *actual sounds* or *differentially audible sounds*, which also coincides with Høyer's individual sounds. What Clay is hearing can be heard by others in the diegesis, but not exactly from his perspective in this particular situation. In the example above, the sound transitions from this perspective to the off-screen non-diegetic, where I have added an objective dimension. I claim that it is objective, because the sound or content is not influenced by any character, as non-diegetic music often is, but the tapes have been recorded and are being listened to objectively. However, in this transition the perspective gets lost in what Stilwell (2007) calls *The Fantastical Gap*. We know that the only reason we can continue listening to the tapes, is that Clay has continued listening to them, and this makes the sound difficult to categorise as non-diegetic, at least if

we still use the argument that Clay's Walkman is the source. In this Fantastical Gap, the lines between diegetic and non-diegetic get blurred, and it is hard or even impossible to locate a precise POA. However, in Stilwell's contribution, the gap indicates a circumstantial location where everything is ambiguous, and the Fantastical Gap might be worth considering as a POA on its own, a point I will come back to.

#### 4.1.2 Simultaneous time zones

A little later in episode 1 Jorkey, 2017) (24:34-28:19), we see Clay riding his bike, listening to Hannah's voice telling him that he has arrived at her house. In the next frame, Clay rides up to the house, but outside there are lots of people having a party. Hannah explains on the tape that when she first moved to the city, she threw a party in which she met the subject of the first tape, her classmate Justin. It is dark when Clay rides around on his bike, and it was dark the night of the party, so it is not apparent whether the Clay riding the bike up to the house is the past Clay going to the party, or the Clay from the current evening imagining the night of the party. It can initially be interpreted as both, but it becomes clear when Clay enters the door and joins the party that the visual belongs to a flashback. The tape continues to roll, and as a result of this, the sound has the Individual POA of Clay in another time than the visual, and at the same time it functions as a voice-over narrator for us to understand the flashback. In the flashback of the party we hear a cell phone buzzing (28:05-28:19), and Clay reaches for his phone in his pocket. He looks at his phone, but as the camera switches for us to see who is calling, the narrative makes a jump again and we see Clay in present time looking at his phone, receiving a call from his mum. This indicates an ambiguity as to where the sound of the phone buzzing comes from, the flashback or the present time, and it can be interpreted as belonging to two places at once



The most intricate phenomenon in this series is when different time zones are depicted at the same time. This happens several times in the situations where Clay is riding his bike to the locations where Hannah is explaining what has happened to her. We often see Clay functioning as what Neumeyer calls a “cameo”, standing next to Hannah, looking at her, but in a different time. In these situations, it is impossible for him to be looking at her where she stands, given that it is impossible for him to know exactly where she stood, and it is therefore clear that this is his subjective image of what has happened, based on what he is hearing on the tapes. However, we do not see the action from Clay’s point of view, we see them together side by side, giving us two narratives depicted at the same. A similar situation can be found in the episode *Tape 6, Side B* (Jorkey, 2017), when Clay rides up to a house where Hannah went to a party (35:31-36:08). We hear the sounds from the party at the same time as we see Clay approaching the house, and we hear Clay’s bike and him stepping off it and the noise from the party simultaneously. Hannah’s voice-over narration is also present at the same time, but Clay is not listening to the tapes as he rides his bike, resulting in three different time zones: 1. The sounds of Clay in present time, 2. The sounds of the party in past time, even though they are the subjective sounds of Clay’s imagination, seeing as the actual sounds of the party is impossible for him to know, and 3. The sounds of Hannah’s voice, either when she recorded them, or more likely from when Clay listened to the tapes before coming to the house. This is therefore an example of Neumeyer’s on-screen non-diegetic exterior, but as I perceive it, this is not done in the same manner as Neumeyer portrays it. It is not a picture in picture manner; an insert, or represented as a brief narrative intrusion; a comment, or a cameo in the corner of the screen, but the complex phenomenon of showing different narratives at the same time without disturbing or confusing the overall story.

In the episode *Tape 2, Side B* (s1e4, Jorkey, 2017), the same three dimensions are depicted, but with some variations. Clay is walking beside a house listening to Hannah’s tape, and we are hearing the sound of him walking, which is onscreen, diegetic, objective and an Observational POA, while simultaneously listening to what he is listening to through his headphones, an Individual POA. This is used on and off as non-diegetic, or in fact homo-diegetic voice-over narration. Suddenly he sees, or at least pictures, Hannah sitting outside the window with a tape recorder, and as the camera hits her, the non-diegetic recording switches to diegetic (02:49-03:50). Clay’s Individual POA changes to diegetic, and we hear

Hannah's voice as if we were standing next to her. As she continues talking into the microphone, we hear Clay taking out a piece of paper from his pocket, giving us two time zones presented visually and audibly at the same time. Clay is still listening through his headphones, so the POA that we are given comes from Clay's picturing of Hannah sitting there with the microphone. He is imagining her being beside him as she speaks, and this is also the situation the audience is put in: A place between the subjectivity of Clay and the objectivity of the visual and audible position that we are given.

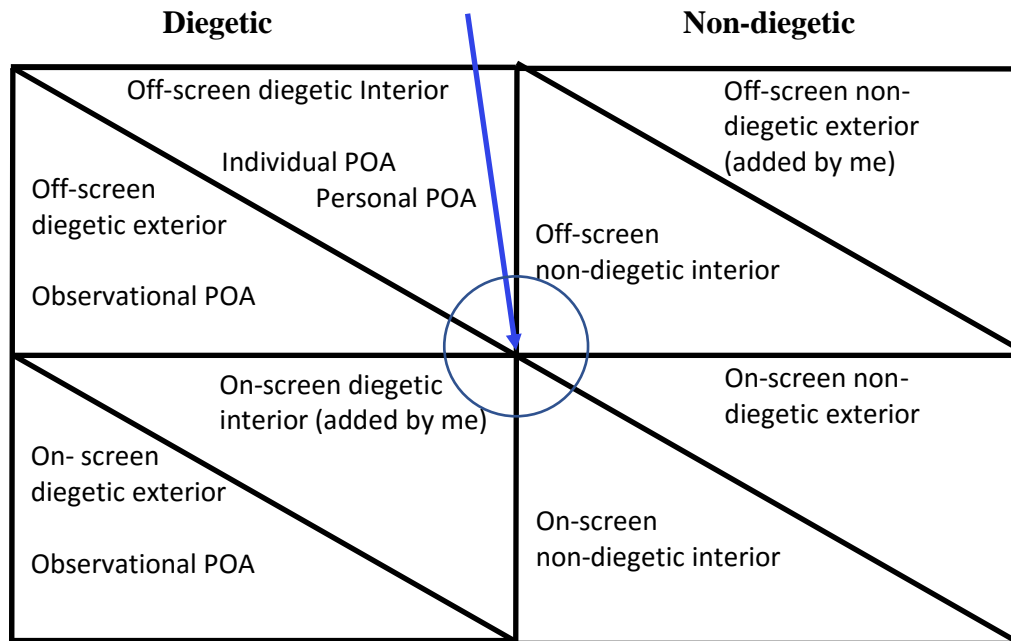
## 4.2 Part two: Active POA and the Fantastical Gap

So far, we have looked at situations where it is difficult to define the narrator's POA, and where several time zones are depicted at the same time, confusing our understanding of POA and what is diegetic and non-diegetic. The coming examples further explore such problematic situations, and investigate a few more aspects that Neumeyer (2015) and Høyer (2013) struggle to cover, namely the identified grey areas between the different POAs. Neumeyer covers more phenomena than Høyer, as Høyer chooses to concentrate solely on diegetic sound in his essay. Hence it is more appropriate to focus on what is missing from Neumeyer's categories. However, although Høyer concentrates only on a small part of what Neumeyer covers, Høyer's also has a category, the Active POA, which does not fit with the schematisation so far.

I suggested above to consider Stilwell's (2007) Fantastical Gap as a POA on its own, in order to address the many grey areas that seem to get lost in the diffuse zone that is this gap. Would it be easier to understand these complex situations if the Fantastical Gap was situated together with established terminology presented in **Table 6** (p.30)? And could this open for a better understanding of Høyer's Active POA? If we establish the Fantastical Gap as a POA, or rather a zone of audition, as suggested by Chion (1994), can there be different kinds of POAs within the Fantastical Gap as well? In **Table 7** below, the categories presented in **Table 6** are presented together with the zone of audition that is the Fantastical Gap. At this point in the discussion, we might ask ourselves if this is taking it too far, and complicating more than illuminating. Seeing as the overall scientific aim is to find a terminology that best enhances our understanding, and facilitates debates. I will in the

following point to some examples that show yet some potential grey areas in the current established terminology, but which can be located within Stilwell’s Fantastical Gap.

**Table 7 Situating The Fantastical Gap as a zone**



#### 4.2.1 The grey area between Active and Individual POA

In **Table 6** (p.30), Høyer’s (2013) Individual POA is placed together with Neumeier’s (2015) Diegetic Off-screen Interior which situates it in the upper left quadrant, on the right side of the “axis of psychological realism”. In the following example, I will investigate what happens when this point stretches towards the left side of the axis, and show that Høyer’s Active POA can find its place there.

*Modern Family* (Lloyd & Levitan, 2009 -) is an American sitcom television show which is made as a mockumentary about a big family divided into three households. Every episode contains interviews with the characters, where they sit in front of the camera telling the viewers what is going on in their lives, and “real life”-footage, where we see the characters go about their business followed by a camera crew, presumably making a documentary about them. The series like to play with the conventions, and deviate from the genre from time to time for comedic purposes, but often the genre traits is not being followed for story-telling purposes, and that is where it gets interesting from a POA point-of-view.



In the episode, *The Day We Almost Died*, (s6e11, Lloyd and Levitan, 2015,) we see the characters Phil and Claire together with their kids, driving in their car (00:05-01:19). Our POV is in the front of the car, giving us the impression that they have attached a small camera on the dash-board to capture everyone in the car. The POA is the same as the POV, and it allows us to hear the parents in the front seats more clearly because they are closer to us. This is a realistic, objective, Observational POA, or Onscreen diegetic exterior, which is the most straightforward auditory perspective. There is a lot of noise in the car and everybody is arguing and yelling at each other. Claire gets a phone call from her brother Mitchell, and since she cannot

hear him because of all the noise, she yells out that everybody needs to “shut up”. She decides to insert an ear-plug in one of her ears, to focus on her conversation and block out the other ones in the car,

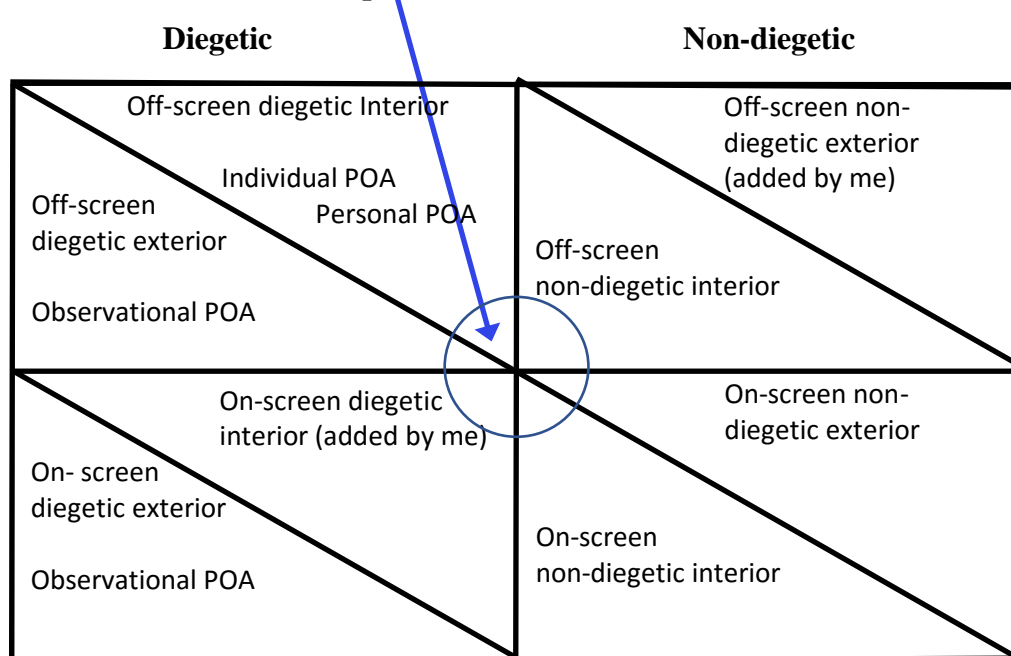


but the attentive eye will notice that the other characters in the car are still talking to each other, but we cannot hear them. This becomes evident also later in the episode, when we “go back in time”, and the POV switches to the other characters in the car, and we get to hear what they are saying. But in this scene, their voices are faded, and we can hear Claire speaking loudly, in addition to Mitchell who is talking on the other side of the line. This is not realistic from our POV, as we should not be able to hear Mitchell, and we should not be able to tune out the other conversations in the car. The auditory perspective becomes even more unrealistic when at the same time Phil starts a different conversation with someone on Bluetooth speakerphone. When the conversation is over between Mitchell and Claire, we see Phil ending his Bluetooth conversation, which can be heard in the far distance, but it is completely unintelligible during Claire and Mitchell’s conversation.

The example above shows a situation where it is quite difficult to locate a precise POA. It suggests an Individual POA from Claire, seeing as we share her auditory perspective, and it can function as Raskin’s (1992) differentially audible sound. Her other ear is still exposed to the other conversations, and it is impossible for her to tune it all out, not even if she has noise-cancelling ear plugs in both ears, and this makes it very unrealistic and not in line with the mockumentary genre. It would therefore not be precisely accurate to categorise this

simply as an Individual POA, seeing as the camera in front of the car is supposedly catching their conversation and blocking out Phil's conversation on Bluetooth. This suggests an Active POA giving the audience access to a situation that is inaccessible from the POV, and it leaves the auditory perspective in a grey area between Active and Individual POA. Both Neumeyer's and Høyer's categories are unable to cover this phenomenon, so it leaves this grey area in the Fantastical Gap found at the intersection between their categories, as shown in **Table 8** below. **Table 8** situates the grey area between active and Individual POA inside the zone of The Fantastical Gap, in the intersection between the interior and the exterior. The reason for it touching the border towards the external is that the Active POA is categorised by Høyer as objective, together with Observational POA.

**Table 8** The grey area between Active and Individual POA situated inside The Fantastical Gap



#### 4.2.2 The grey area between Observational and Individual POA

In Table 6 (p.30), the diegetic, off-screen objective sound is situated in the upper left quadrant together with Høyer's (2013) Observational POA. It is clear from this that we are talking about the exterior, because it is placed at the left side of the axis of psychological realism. In the following example, I will show what happens when this category stretches

towards the border of the interior and Høyer's Individual and Personal POA, and we find the grey area between Observational and Individual POA.

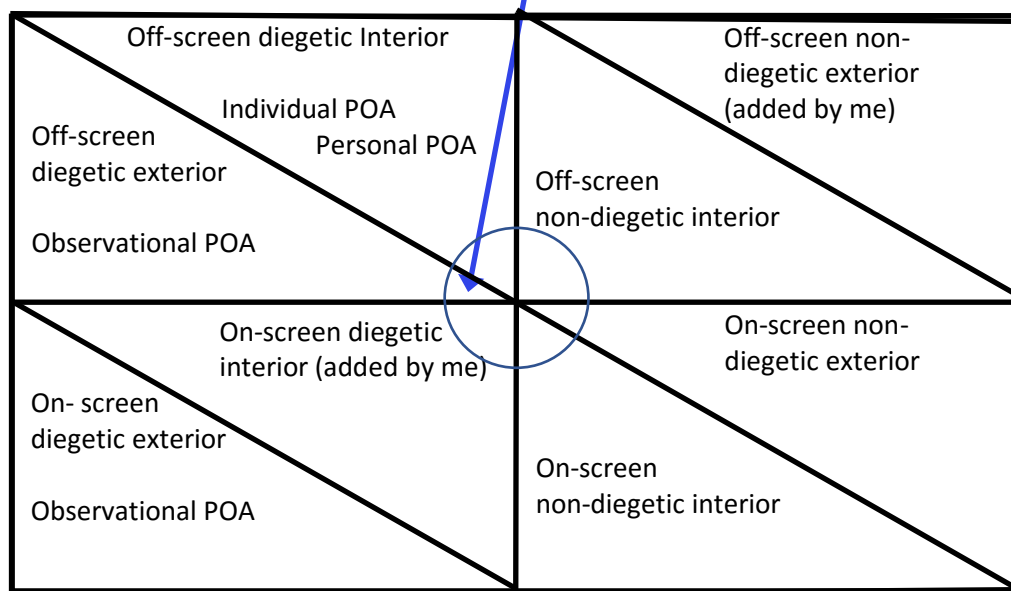
In the *Modern Family* episode *Clean for a Day* (s7e12, Lloyd and Levitan, 2016) (20:53-21:53), we are at the house of Phil and Claire's family, and they are all gathered in the hallway talking about a stuffed animal from the children's childhood that they have managed to save from a donation bank earlier that day. The stuffed animal is making a terribly annoying sound, and they agree that they have to make it stop. Phil is handed the toy and a rock, and leaves the room to "kill" the stuffed animal. We the spectators, and the supposed documentary film crew are left together with the rest of the family in the hallway, and cannot see what is happening in the other room, leaving us with an Observational POA. The stuffed animal continues to make the annoying noise, and as we, and the rest of the family wait for a moment for Phil to "kill" the toy in the other room, a suspension builds up. What is interesting from an auditory perspective in this example is that its sound stays at the same level of volume even though Phil leaves the room, and the thump that is finally heard is as loud as if it was in the same room. This creates an effect of discomfort, which is emphasised by the facial expressions of the remaining characters in the hallway, and we are therefore dealing with a subjective aspect as well as the objective sounds that are heard.



The example above shows a grey area between objective and subjective sound. The cringing response at the end shows the emotional ties that the family has towards the stuffed animal, and the fact that the sound seems much closer to the ones involved, indicates that we share their perspective, and their subjective POA in this moment. This can simply be said to be an Individual POA, although a shared Individual POA of all four family members in the hallway, and not a personal POA, because the sound is objectively present, and can be categorised as an *actual sound*, and a *differentially audible sound*, to use Raskin's (1992)

terminology. This is not, however, perceived in the moment as something subjective by us, because the sensation is not strong enough. Seen from the film makers' point of view, it does not seem to have subjectivity as its narrative objective, and because of that the exaggerated sound may seem surprising. The phenomenon fits better with what Buhler et al. (2010) and Høyer (2013) calls sweetening of the sound, a strategy used to make it more effectual and close-up, in order to inflict empathy with the characters. Although Høyer acknowledges this phenomenon in his essay, by addressing sweetening, it is still a grey area that does not fit within the categories in **Table 6**, The phenomenon does, however, fit within the Fantastical Gap as shown in **Table 9** below. The grey area between Observational and Individual POA is situated in the circle, divided by the axis of psychological realism (Neumeyer 2015). The reason for placing it more on the external side is the genre of the series, in which we are supposed to perceive what we see as objective, realistic and documentarist, even though they often deviate from the conventions of this genre for humour and narratological purposes.

**Table 9**      **The grey area between Observational and Individual POA situated inside the Fantastical Gap**



### 4.2.3 Grey area between diegetic and non-diegetic, and subjective and objective

Even though crossing the established borders is quite common, it can be problematic if the auditory perspective stays in the Fantastical Gap. This is the case in the “Camelot-scene” in the movie Jackie (Larrain, 2016), in which Jackie Kennedy, portrayed by Natalie Portman, puts on a record of the song Camelot



(Loewe, 1978), from the musical with the same name. The song begins playing while Jackie is standing in the same room as the LP player, but as she leaves the room the quality of the sound changes. We follow her visually, and auditory, showing signs of the camera crew walking behind her. Seeing as the camera crew follows her, it is an objective perspective, a purely Observational POA that stays so close to Jackie that we are almost hearing it from her Individual perspective, but we are two steps behind her. The situation is realistic in the sense that we understand that the music is diegetic during the whole scene, and it is in line with the genre and the established conventions in the rest of the movie. However, the significant happens when the music continues in the same manner even though the visual are cutting in time, like a montage.

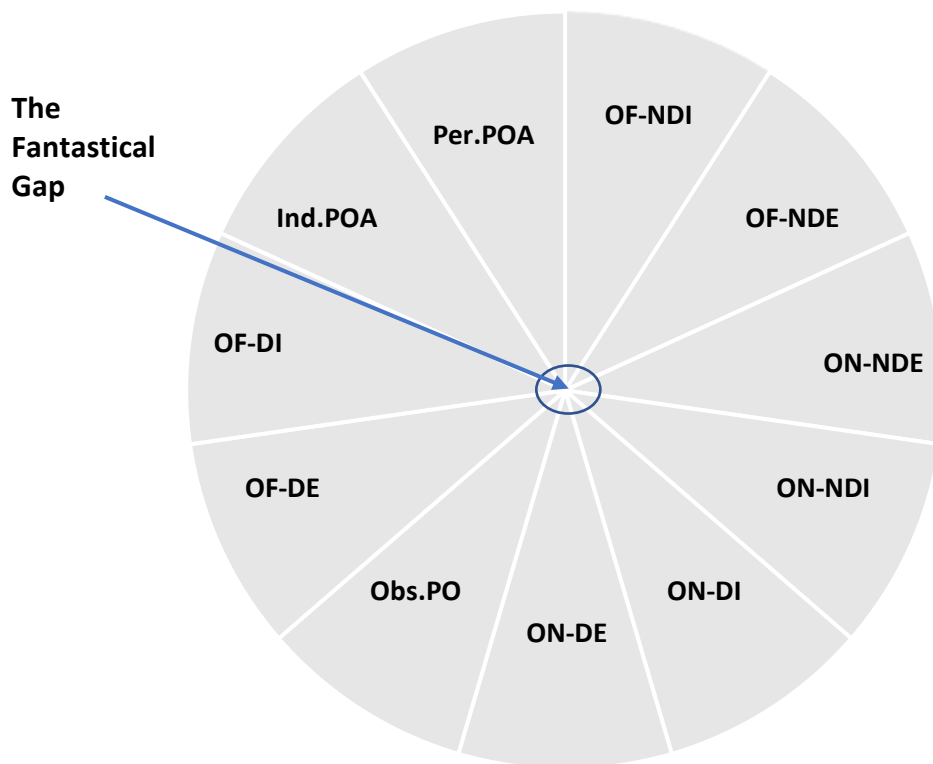
This is an excellent example of sound and music functioning as a unifier and a facilitator of continuity (Gorbman, 1987), but it deviates from the former conception of a camera crew giving us visual and audio from the same source, and therefore creates what Høyer (2013) calls audio visual dissonance. In what becomes a montage situation, the diegetic music crosses over and functions as non-diegetic music, but at the same time there is no question as to the music still being diegetic. This creates a grey area that is unrealistic, and will be best fitted within Stilwell’s (2007) Fantastical Gap. Jackie walks around the house changing dresses and jewellery, and the sound changes accordingly, so there is no question as to where the source is. However, in all the time she spends walking around, the record must have been played repeatedly to still be playing at the end of the scene. It can therefore be understood as what Bordwell and Thompson calls temporally displaced sound (1985). This

is not only her individual perspective, at least not always, but can instead be understood as a partially Observational and Personal POA because it might be perceived as her state of mind, and that this record is playing repeatedly in her head. However, the record might of course also be doing that in the diegetic reality, but in the montage presented to us it only plays one time, and it creates a blurred auditory perspective.

The example above shows music functioning as diegetic and non-diegetic at the same time, and showing the same character in different times during a montage scene. The auditory can depict Jackie at one point in time while the visual is showing her in another. The audience shares her auditory perspective even though it is the objective perspective as well, which lets us feel empathy for her without entering her mind or accessing a purely subjective POA. It is realistic because the sound changes when she moves around, but the sound does not make the same time jumps as the visual, which makes it unrealistic.

We have now moved over to the right-hand side of the axis of psychological realism, at the same time as the sound source stays on the left, and realistic side. The sound we hear is in fact diegetic on-screen exterior because we know the source. But the fact that the music continues to play independently, suggests a subjective interior aspect of the character. At the same time it functions as non-diegetic music, even though we know the source is diegetic. Given that she walks around, the sound is sometimes onscreen and sometimes off-screen, and the representation of the sound takes this into consideration, making it objective and realistic. Because of all the different ways the sound can be understood at the same time in this scene, this phenomenon can best be located in the inner point of the Fantastical Gap, the place where all borders meet. The design in **Table 7** does not give this point fully justice, so it is better exhibited as in **Table 10** below.

**Table 10** Grey area between diegetic and non-diegetic, and subjective and objective



OF-NDI	Off-screen Non-Diegetic Interior
OF-NDE	Off-screen Non-Diegetic Exterior
ON-NDE	On-screen Non-Diegetic Exterior
ON-NDI	On-screen Non-Diegetic Interior
ON-DI	On-screen Diegetic Interior
ON-DE	On-screen Diegetic Exterior
Obs.POA	Observational POA
OF-DE	Off-screen Diegetic Exterior
OF-DI	Off-screen Diegetic Interior
Ind.POA	Individual POA
Per.POA	Personal POA

## 4.3 Summary

In these analysis, we have seen several examples of grey areas not covered by Høyer and Neumeyer, which show that established terminology is not yet fully equipped to describe such complex auditory phenomena. First there was the dilemma of what to categorise as diegetic and non-diegetic, exemplified by the status of Hannah as narrator in *13 Reasons Why*, and the shifting status of the song *I Want to Break Free* in *Sherlock*. The crossing of boundaries showed that the categories added by me, onscreen diegetic interior and non-diegetic off-screen objective, are highly relevant. Secondly, we saw that the portrayal of several time zones at the same time also blurred our conception of not only what is diegetic and non-diegetic, but also objective and subjective. The best example of this is in episode *Tape 2, Side B*, where we participate in Clay's imagination of how Hannah's voice would sound like to him, if he was present when she recorded it, at the same time as he is listening to the recording.

Part two investigated the suggestion to add Stilwell's Fantastical Gap as a zone that captures the grey areas between Active and Individual POA and Observational and Individual POA. This facilitated a focus on the subtle boundaries between external and internal, seeing as Neumeyer's Axis of psychological realism divides these categories, and the Fantastical Gap gives an opportunity to place auditory perspectives such as the car scene from *Modern Family*, and the phenomenon of sweetening the sound. Lastly, the Fantastical Gap as a point itself was suggested, as a possibility to embrace the unstable ambiguity exemplified by the montage scene in *Jackie*. My suggestion is therefore both to place highly complex auditory phenomena like the *Jackie*-scene in the specific point that tie all the different categories together, as in **Table 10**, and to distinguish different phenomena within the zone of the fantastical gap, as shown in **Table 7**.



## 5. Concluding notes: scope and ambition

---

This thesis has made matters more complex and perhaps more complicated, in order to show that there can in fact be terminology which covers grey areas. The term POA can be argued to cover grey areas when the term is understood in more nuanced ways than has been proposed by scholarship until now, and you could perhaps expand with the Fantastical Gap as well, as shown through my analysis. I have not suggested specific terms, except for my refinement of Neumeier's categories in **Table 6** (p.30), but I described the more complex phenomena as grey areas between the established borders. This is because I believe in fact that a terminology that covers every specific phenomenon, will not necessarily be suitable as an analytical tool for researchers. Høyer says that "certainly, it can be fruitful to discuss sound perspectives with more detailed categories, but a categorization of POA should also be based on variations that are identifiable for scholars, students and practitioners" (Høyer 2013), which is a statement I highly agree with. Høyer states this as an explanation as to why he does not take it further than these four subcategories, as well as underlining that he knows the limitations of his categories, which has been the foundation of this thesis' investigation. This thesis shows that the limitations of Høyer's categories are in fact quite extensive, leaving out the non-diegetic vs diegetic aspect, and while acknowledging that there are grey areas, his categories does not help our understanding of them.

So, in what way has this thesis been helpful? A highly relevant question to ask is if the categories used in this thesis' analysis are useful, easily identifiable and easy for scholars to use, or if they are too detailed and complex. My answer to this question is that they certainly are too complex in the form which they are presented here, but my contribution is not a definite suggestion to a new categorisation of POA, but to start a debate on the relevance of using this term in its current shape. The even more important argument from my side is that if we are going to have a term like POA it should not be generalising, but it should be thorough enough to at least cover auditory phenomena which are commonly used today.

In the introduction, I argued that POA is important because of the focus on which "points" we are "listening from" within the diegesis. However, as this thesis has highlighted, these "points" might also be non-diegetic, or somewhere between those two dimensions. Chion (1994) suggested to speak of an auditory zone instead of a single point, because the position

that we are listening from is more encompassing than the given picture, which gives us a more rigid visual perception. On the other hand, the examples showed in this thesis capture quite specific auditory perspectives, giving credit to the “point”-analogy in some of the situations. If we take the examples where the Fantastical Gap is suggested as a narrow point where all the grey areas meet, it functions as a specific auditory “place”, if not a point, although it is ambiguous in its nature. The Fantastical Gap is a possibility to capture all the phenomena, or so far unnamed categories, that does not fit within either Høyer nor Neumeyer’s categories, but by suggesting this, a new important question arises: Is it enough to just put everything that does not fit within the borders inside the Fantastical Gap and leave it at that?

This thesis’ aim was, as presented in the introduction, to form a first step towards answering the larger question: Whether the conceptual framework for discussing sound in film and television is fulfilling its purpose, namely aiding our understanding of how the different auditory perspectives influence the audience’s experience of the visual narrative. As this thesis shows, film sound researchers use different terminology to describe auditory phenomena, which leads to several different terms for the same perspectives. This is understandable in a research field that is quite new, and the categories presented by Høyer and Neumeyer differ also because their focus points are differing. Shifting focus points are natural in a research field that constantly evolves, and sound theory has had to evolve quickly with the rapid development of technology during that last twenty years. But the other question presented in this thesis’ introduction was whether the refinements Høyer offers are helpful, or if it simply makes the term even more complex and subject for misinterpretations. Initially POA was a narrow term, only describing the auditory phenomenon of accessing a character’s inner thoughts, as Altman (1992) defined it, and this definition did not open for many misinterpretations. However, seen in light of Høyer’s and Neumeyer’s categories, Altman’s definition of POA can be said to generalise subjective auditory phenomena.

I believe that Høyer’s expansion was helpful when it comes to the specific term POA, and that it might have been necessary seeing as Chion divided POA into too broad categories in 1994. Maybe it was there it all started, and had it not been for this first expansion, POA would only refer to a specific subjective perspective. In that case, it would still be a single term. However, the challenge with POA in its current shape is that it is no longer just a

term, but a set of categories or a concept. Looking at it this way, Høyer's contribution has definitively made it more complicated and difficult to use. As this thesis show, they do not cover enough and at the same time they are too specific to be used in a generalising way. One can say that Høyer's categories fall between two stools. The same goes for Neumeyer's categories, although they are equipped to cover more than Høyer's. The fact is that neither sets of categories are cohesive enough to envelope the complex auditory phenomena in modern film and television series, and that only by including Stilwell's Fantastical Gap can we more or less account for the grey areas between the categories. The examples in the analysis show that these are not rare exceptions, and that such phenomena appear in both film and cinema, and in comedy as well as drama. I hope that this investigation has shown that more work needs to be done, in order to establish a cohesive theoretical framework, and that this has opened up for further thought regarding what terminology that adequately can be used for describing such auditory phenomena.



## 6. References

---

### 6.1 Literature

Altman, Rick (1992) *Sound Theory Sound Practice*. New York: Routledge.

Buhler, James et al. (2010) *Hearing the Movies – Music and Sound in Film History*. New York and Oxford: Oxford University Press.

Bordwell, David and Thompson, Kristin (1985) “Fundamental Aesthetics of Sound in the Cinema” in Elisabeth Weis and John Belton (eds.) *Film Sound: Theory and Practice*. New York: Columbia University Press.

Bordwell, David and Thompson, Kristin (2013) *Film art: an introduction*. New York: McGraw-Hill Companies Inc.

Branigan, Edward (1992) *Narrative Comprehension and Film*. London: Routledge

Chion, Michel (1994) *Audio-Vision – Sound on Screen*. (Edited and translated by Claudia Gorbman) New York: Columbia University Press.

Coyle, Rebecca (2010). “Point of Audition: Sound and Music in *Cloverfield*” in *Science Fiction Film and Television*, vol. 3(2), p. 217-238. Liverpool University Press

Genette, Gérard (1980) *Narrative Discourse. An Essay in Method*. (Translated by Jane E. Lewin). Oxford: Blackwell.

Gorbman, Claudia (1987) *Unheard Melodies: Narrative Film Music*. Bloomington: Indiana University Press.

Hibberd, Sarah and Nielsen, Nanette (2015) “Immersion and proximity: music, sound, and Subjectivity in *The Memory Dealer*” in *Journal of Sonic Studies* 9. Retrieved from:

<https://www.researchcatalogue.net/view/148488/148489/0/0>

Høyer, Svein (2013). “The relevance of Point of Audition in television sound: Rethinking a problematic term” in *Journal of Sonic Studies*. Retrieved from:  
<http://journal.sonicstudies.org/vol03/nr01/a04>

Neumeier, David (2009) “Diegetic/Nondiegetic: A Theoretical Model” in *Music and the Moving Image, Vol. 2, No. 1 (Spring 2009)*, p. 26-39. University of Illinois Press.

Neumeier, David (2015) *Meaning and Interpretation of Music in Cinema*. Bloomington: Indiana University Press.

Newman, Michael Z & Levine, Elana (2012) *Legitimizing Television – Media Convergence and Cultural Status*. New York: Routledge.

Raskin, Richard (1992) “Varieties of Film Sound: A New Typology» in *(Pré)Publications 132*, p. 32-48. University of Aarhus.

Smith, Jeff (2009) “Bridging the Gap: Reconsidering the Border between Diegetic and Nondiegetic Music” in *Music and the Moving Image, Vol. 2, No. 1 (Spring 2009)*, p. 1-25. Urbana: University of Illinois Press.

Stilwell, Robynn J. (2007) “The Fantastical Gap between Diegetic and Nondiegetic” in Daniel Goldmark, Lawrence Kramer, and Richard Leppert (eds.) *Beyond the Soundtrack*. Berkeley and Los Angeles, CA: University of California Press.

Tranås, Marie W. (2016) *Point of Audition – En litteraturgjennomgang/A literature review*. Unpublished. Department of Musicology, University of Oslo, Oslo.

Whittington, William (2011) “Sound Design for a *Found Future*: Alfonso Cuarón’s *Children of Men*” in *New Review of Film and Television Studies* Vol.9 (1) p. 3-14

## 6.2 Television-series and episodes

**13 reasons why, 2017 created by Brian Yorkey, Paramount Television.**

*Tape 1, Side A*, season 1, episode 1, directed by Tom McCarthy, 2017, Paramount Television.

*Tape 2, Side B*, season 1, episode 4, directed by Helen Shaver, 2017, Paramount Television.

*Tape 6, Side B*, season 1, episode 12, directed by Jessica Yu, 2017, Paramount Television.

**Sherlock, 2010 created by Marc Gattis and Steven Moffat, BBC Wales.**

*The Final Problem*, season 4, episode 3, directed by Benjamin Caron 2017, BBC Wales.

**Modern Family, 2009 created by Christopher Lloyd and Steven Levitan, 20<sup>th</sup> Century Fox.**

*The Day We Almost Died*, season 6, episode 11, directed by James R. Bagdonas, 2015, 20<sup>th</sup> Century Fox.

*Clean For A Day*, season 7, episode 12, directed by Beth McCarthy Miller, 2016, 20<sup>th</sup> Century Fox.

## 6.3 Film scenes, songs and images

### Film scenes

“Opening scene”, *Guardians of the Galaxy*, 2014 directed by James Gunn, Marvel Studios.

[https://www.youtube.com/watch?v=x\\_jRQBGKPaA](https://www.youtube.com/watch?v=x_jRQBGKPaA)

“Camelot scene”, *Jackie*, 2016 directed by Pablo Larraín, LD Entertainment.

<https://www.youtube.com/watch?v=i7lj0y4KRXg>

“Shower scene”, *Psycho*, 1960 directed by Alfred Hitchcock, Shamley Productions.

<https://www.youtube.com/watch?v=0WtDmbr9xyY>

## Songs

“I want to break free”, Queen, 1984 produced by Queen and Reinhold Mack.

“Camelot”, Frederick Loewe and Alan Jay Lerner, 1978 written for the musical *Camelot*.

## Images

Netflix.com A, Screen shot of *13 Reasons Why* (2017) season 1 episode 1

<https://www.netflix.com/watch/80117471?trackId=14277281>

Netflix.com B, Screen shot of *13 Reasons Why* (2017) season 1 episode 1

<https://www.netflix.com/watch/80117471?trackId=14277281>

Netflix.com C, Screen shot of *13 Reasons Why* (2017) season 1 episode 1

<https://www.netflix.com/watch/80117471?trackId=14277281>

Netflix.com D, Screen shot of *13 Reasons Why* (2017) season 1 episode 1

<https://www.netflix.com/watch/80117471?trackId=14277281>

Netflix.com E, Screen shot of *13 Reasons Why* (2017) season 1 episode 4

<https://www.netflix.com/watch/80117474?trackId=200257859>

Netflix.com F, Screen shot of *Modern Family* (2015) season 6 episode 11

<https://www.netflix.com/watch/80043946?trackId=14170289>

Netflix.com G, Screen shot of *Modern Family* (2015) season 7 episode 12

<https://www.netflix.com/watch/80085086?trackId=200257859>

Netflix.com H, Screen shot of *Modern Family* (2015) season 7 episode 12



<https://www.netflix.com/watch/80085086?trackId=200257859>

Youtube.com A, Screen shot of *Jackie* 2016

<https://www.youtube.com/watch?v=i7lj0y4KRXg> (00.36)