



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Acousmatic Sound in Film:

A Means of Perceptual Manipulation“

verfasst von / submitted by

Alina Baukholt

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Arts (MA)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

UA 066 844

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium
Anglophone Literatures and Cultures

Betreut von / Supervisor:

Univ.-Prof. Dr. Sylvia Mieszkowski, MA

Acknowledgements

This Master's thesis has been a challenging, nonetheless a joyous journey. I am fortunate to have had support of different kinds from several people during the writing process. I would like to express my deepest gratitude to every single person who supported me.

First and foremost, I would like to thank my supervisor Univ.-Prof. Dr. Sylvia Mieszkowski for her continuous support and advice, helpful feedback, and invaluable guidance during the process of writing this thesis. Thank you for encouraging me to pursue my ideas and empowering me to constantly improve my writing.

I am beyond grateful to my loving partner and family for their unfailing emotional support and patience throughout my studies and the writing process of this thesis. Thank you for encouraging me in every way and believing in me when I doubted myself. Without you, I would not have been able to achieve my goals.

Finally, I cannot forget to thank my friends for their unconditional support in this very intense process. Thank you for your motivating and inspiring words, for always listening, and offering me joyful breaks when I needed them. Special thanks to those who immersed themselves intensely in the analysis of the films as my interlocutors, which allowed me to gather different information and to carry out analyses that are more specific. My heartfelt thanks to all of you.

Table of Contents

List of Figures.....	1
Introduction	2
1 The Nature of Sound	7
1.1 Auditory Orientation and Processing	7
1.2 Categories of Sound Qualities	9
1.3 The Use of Sound in Film Production	12
1.4 From Physical Hearing to Interior Sound Perception	17
1.4.1 How Hearing Works.....	17
1.4.2 Sound Perception and Psychoacoustics	19
1.4.3 Listening	21
2 Sound in Film: The Acousmatic.....	24
2.1 The Concept of the Acousmatic	25
2.2 Auditory Phenomena.....	30
2.2.1 Acousmatic Zones.....	30
2.2.2 Music	33
2.3 The Voice	39
3 Strategies of Cognitive Processing	47
3.1 Mental Processing Models	49
3.2 The Listener's Activity.....	56
3.3 Imagination and the Simulation Process	63
3.4 A Cognitive Understanding of Emotions	67
Conclusion	73
Bibliography	80
Appendix A: Reviews	85
Appendix B: Abstract	92
Appendix C: Anti-Plagiarism Statement.....	94

List of Figures

Figure 1. Illustration of the structural relationship between the outer, middle, and inner ear (Menner 14).	18
Figure 2. View from the unknown caller's perspective [0:49:07].	42
Figure 3. Close-up of Stu [0:23:36].	43
Figure 4. Close-up of Stu [0:24:29].	43
Figure 5. Schemata of information-processing (Clarke, <i>Ways of Listening</i> 13).	50
Figure 6. Bob lying in bed, talking on the phone to his wife Lydia [00:54:33].	60
Figure 7. Bob in the hotel pool, talking on the phone to his wife Lydia [01:16:44].	60
Figure 8. Tony and J.A.R.V.I.S. talking to each other [54:52].	68
Figure 9. Harold outside of the bakery [0:13:14].	69
Figure 10. Harold from a bird's eye perspective, desperately shouting at the voice [0:13:20].	70

Introduction

Sound is ubiquitous – it surrounds all of us, no matter what time and which place we are in. The acoustic landscape in which we live, is generally composed of voices, music, the sound of machinery, vehicles, natural sounds, and electrical devices. Unfortunately, in today's society our aural environment is considerably reduced to a specific function: the accompaniment of our visual sense. We are used to consider the phenomenon of sound in the context of our other senses and have difficulties to engage with sound for its own sake. At a time when we are regularly flooded with visual material in every aspect of life, "sound is often simply a background for our visual attention" (Bordwell & Thompson 269), which causes that we become used to disregarding many of the sounds in our surrounding. As a musician, I am accustomed to situations in which I get involved with sound for reason of complete devotion to its singularity. I am solely confronted with a sound's properties and able to receive an acoustic impression without relying on my other senses. And although there are probably plenty of musicians or sound experts who are able to relate to my experiences or think similarly, this is not the general case for the perception of sound in ordinary life nowadays. Usually, as Bela Balazs points out, "[t]ogether with the sounds and voices of things we see their physiognomy" (Weis and Belton 117), meaning that we are inured to see the visible sources of the sounds we hear simultaneously. I have often witnessed in my social environment that the absence of a visible source raises problems in terms of a loss of control in people's minds, as the connection between the senses breaks off, which turns out to be problematic in dealing with sound. These personal experiences paired with the general superiority of the visual sense in society have caused me to seek answers to how the influential phenomenon of sound is discussed in the context of an audio-visual scope.

My subsequent interest was directed at finding suitable objects of analysis, which reflect this discourse and contextualise sound phenomena according to their crucial characteristics in relation to our sense of sight. I have chosen the film medium for my work, as it features prominently in contributing to the discourse and challenges people's urge to control their senses and their perception. The beginning of the sound era provided the film industry with the power of sound. Since then, the relationship between image and sound has been playing a major role in film production and its technical process. As a medium, film combines the senses of sight and sound for a coherent experience and at first glance, they

seem to form an interwoven entity with sound naturally accompanying the images on screen and with the images at the centre of attention. This unity basically arises from the right timing because when image and sound occur at once, they are conceived by humans not as two separate events, but as one (Bordwell/Thompson 270). With this in mind, the question becomes what impact the detachment of sound from image can have on the viewer, which is why this master thesis focuses on the interrelation between the use of acousmatic sound in film and its cognitive impact on humans within the scope of audio-vision.¹

In the larger context of audio-visual media, film is a suitable source and powerful form of art that displays the interconnectedness of images and sound considerably. It is well-known that sight and sound influence and reshape each other, as proposed by Walter Murch in the Foreword of *Audio-Vision* (2019), stating that “the sound makes us see the image differently, and then this new image makes us hear the sound differently” (19). However, vision usually dominates hearing, which is why the impact of sound is often underestimated and not sufficiently explored in terms of perception (Bordwell/Thompson 269). In his article “Aural Objects”, Christian Metz acknowledges the conventional analysis of film sound, as the object itself is conceptualised “in a way that [it makes sense] only for the image” (29). For this master thesis, I will diverge from this approach and immerse myself in the study of sound the way it is used in film and its qualities concerning the influence on perception. Nevertheless, paying attention to the relationship between the two senses is inevitable when investigating how sound modifies perception in an audio-visual medium. I consider the object of sound in terms of Michel Chion’s approach, who elaborates the combination of visual and aural components according to their mutual influence and transformation and the effects this has on perception (Murch 15, 19).

For this master thesis, I chose the following films as my objects of analysis: *Iron Man* (2008), directed by Jon Favreau, *Lost in Translation* (2003), directed by Sofia Coppola, *Phone Booth* (2002), directed by Joel Schumacher, and *Stranger Than Fiction* (2007), directed by Marc Forster. All films were produced in the same time period, but represent different genres: Comedy, Romance/Drama, Thriller, and Action. In respect of my research area, these films are suitable in that they are representative of the different types of sound I will investigate and

¹ A neologism coined by Michel Chion in 1990, with which he “wished to describe a unique mode of perception engaged by film, television, and other audiovisual media” and to state that “studying a film’s sound and image independently of each other would make no sense” (Chion xxi)

allow for comparative analyses and inferences of these. David Sonnenschein explains that sounds that are audible in a film originate from “an object, person or event that is either diegetic (on screen or off screen) or nondiegetic” (153). Both areas will be significant for my analysis, since the spatio-temporal dimension offers valuable insights into the nature of sound and its potentials as a manipulative device. They will be of particular importance in terms of acousmatic sound, the type of sound I will focus on in this thesis. The word ‘acousmatic’ is of Greek origin and was initially defined as “indicating a noise which is heard without the causes from which it originates being seen” (Schaeffer 64). The concept of the acousmatic was taken up again by Pierre Schaeffer in clarifying his theory of the “Sound Object” and explaining the idea of the acousmatic experience. It is important to connect this concept, which principally defines media such as the telephone and the radio, where the acousmatic concept is rather clear, to the domain of fictional film. I will devote myself to acousmatic situations in films since the separation of sound influences and changes the listening situation and the way we perceive it in the audio-visual context.

My central claim in this thesis is that, in film, the experience of acousmatic situations not only offers insights into how listening conditions are altered significantly, but also into how involved characters’ and the audience’s perception is manipulated in ways that can cause multifaceted mental consequences. I aim to answer the following research questions: Firstly, I will investigate in which way the dissociation of sound from image in film affects the characters’ and the film audience’s auditory perception. Secondly, I seek to answer to what extent acousmatic sound, with special emphasis on music and the disembodied voice, constitutes a means of manipulation. Finally, I will examine how the audience perceives acousmatic sound by means of various processing strategies on a cognitive level and scrutinise possible repercussions. Methodologically, I will rely on a qualitative data analysis (Mikos, Maeder 2013) of the films as well as on reviews of my interlocutors. I consulted four persons in terms of how they perceived acousmatic listening situations in specific film scenes, highlighting the individuality of mental processing.² My approach is twofold in that I focus on the effects of acousmatic sound on the characters in the films on the one hand, and on the audience on the other hand.

² Please see chapter 3 and the reviews in the Appendix for further details.

Offering an interdisciplinary approach, my analysis will include important elements drawn from Sound Studies (see Schaeffer 1994, 2017; Sonnenschein 2001; Chion 1999, 2019; Belton and Weis 1985; Kane 2016), Film Theory (see Bordwell and Thompson 2010; Chion 1999; Metz 1974, 1980) and Cognitive Theory (see Bordwell 1985; Currie 2008; Plantinga and Smith 1999). In chapter 1, background information on the essence of sound and how we deal with it will be provided. I will first explain its function in everyday life in chapter 1.1, and present how sound functions as a means of guidance in a human's acoustic environment. Additionally, I will clarify how acousmatic information is stored in our brains in terms of its capability to evoke certain responses relying on memories and emotions. In 1.3, I will turn to the use of sound and its essential role in film production and illustrate the different dimensions of sound with regard to their complex relation to other film elements. The transition from physical hearing to interior sound sensation will be highlighted in chapter 1.4. In subsection 1.4.1, I will present how hearing works physically in terms of perceiving and distinguishing sounds. Section 1.4.2 stresses the importance of inner sound perception in respect of psychoacoustics.

In chapter 2, I thoroughly discuss the notion of the acousmatic. Concerning Sound Studies, Schaeffer's concept of the acousmatic and the respective experience, through which the listener is directed away from the visible source that causes a sound, serves as a starting point in chapter 2.1 (64-65). To expand on the acousmatic situation in film, yet without neglecting the importance of it within the audio-visual context, Michel Chion is a helpful theorist to draw on. In chapter 2.2, I will explain the different types of auditory phenomena with special emphasis on the acousmatic zones, namely offscreen, onscreen and nondiegetic (Chion 73ff.)³, and music (Chion 79). Film Theory contributes to a critical analysis of sound in the medium of film, since it reveals several auditory techniques used in filmmaking. Due to its perceptual properties, sound is chosen according to distinctive types and manipulated to fulfil its purpose and achieve the effects desired (Bordwell/Thompson 269ff.). Of major importance is the instrument of the voice, including the concept of the 'acousmêtre', which Chion addresses in a separate work (*The Voice in Cinema*, 1999). I will analyse in chapter 2.3 the essential role of the voice in my objects of study and to what extent it contributes to my

³ This type of reference, without citing the name of Chion's work, will be continuously used in this thesis to refer to *Audio-Vision: Sound on Screen* (2019).

research questions as an acousmatic element. I will examine how the voice affects the audience in terms of perception and how it can influence mental processes.

In chapter 3, the issue of mental processing will be highlighted. Complementing the theoretical framework, Cognitive Theory plays a major role concerning the perceptual and cognitive aspects of film viewing in terms of the audience's activity (Bordwell 29). I will provide definitions of different strategies, focusing on mental processing models in chapter 3.1, and investigate whether perception changes depending on a certain model and which cognitive dimensions come to the foreground while processing. In chapter 3.2, I will devote myself to listening as an individual activity, since diverse perceptual capacities determine how sound is perceived by individuals. The variety of personal factors reveals the richness of thoughts during perceptual processes, which is suitable to the analysis of the qualities of acousmatic sound, as it helps to discuss the different degrees of manipulation according to individual perspectives. I will examine another strategy of cognitive processing in chapter 3.3, which is concerned with cognitive illusion. I will analyse how acousmatic situations in film enable the audience to make inferences about how the characters' minds function and base my interpretation on Currie's concept of 'Imagination' (144-147). The "mental mechanism" (ibid.), which functions via "simulation" (ibid.) to comprehend mental processes of someone else, is particularly important in terms of the intersection of fiction and reality. In addition to that, I will explore how emotions are represented in the films and how they are experienced by the audience (3.4). Here, I will employ Tan's and Frijda's discussion of 'witness emotions' and analyse in which ways emotions shape the acousmatic experience (52). Finally, I will compare my findings and consider the degrees of perception that are put forward in the films through the creation of acousmatic situations. I will discuss possible consequences these listening experiences can have for the film audience and scrutinise, in which way they can have an influence in the real world in positive and negative ways in the conclusion.

The analysis helps to substantiate my claim, which implies that when a situation in a film is created in which a sound is detached from its visible source, this has consequences for affected characters as well as the audience. It is important to mention that the degree of manipulative features of acousmatic sound and the repercussions in cognitive terms will be explored in detail through the qualitative data analysis. In addition, my analysis fills a gap in research on the processing of acousmatic sound in the cognitive sense because previous

research has mostly focused on the way in which fictional images affect mental processes. Thus, this master's thesis will contribute to academic research on the merging of Sound Studies and Cognitive Studies within the field of film analysis. It shall further raise awareness for the psychological consequences of experiencing fiction in the broader area of mental health.

1 The Nature of Sound

Sound not only produces a ubiquitous acoustic sphere but further considerably contributes to shaping how we experience our environment. As described in the introduction, undergoing everyday life is predominantly identified with the sense of seeing as the dominant sense. The superiority of the visual often entails that the other four senses are “filtered through a visualist framework” (Bull/Back 2), which causes serious problems regarding our capacities of understanding our social environment and using our imagination. This first chapter will introduce the equally essential role of sound and its impact on our understanding of the world, without superseding the visual sense as its counterpart. It will introduce the acoustic field of study in its diversity, ranging from auditory processing, separate sound categories and its different contexts to the use of sonic features in film production. Furthermore, physical hearing and interior sound perception will be presented, which provide a basis for understanding sound in its entirety. Acoustic fields according to physicist R.B. Lindsay's “Wheel of Acoustics”, which are of major importance to my work, are “Life Sciences” and the “Arts”. In the inner circles, these disciplines contain among others “Music”, “Speech”, “Communication”, “Psychology”, “Psychoacoustics”, “Physiology” and “Hearing” (Lindsay 2242). All of these are intertwined and play a fundamental role in my thesis, as they emphasise the crucial role of sound in human life as well as the richness and potential of acoustic qualities in regard to its impact on a society.

1.1 Auditory Orientation and Processing

Sound can evoke a multitude of different connotations, have various meanings to different people, and expresses itself in most diverse manners. Aside from its way of occurrence, what all sound has in common is its function as a means of guidance for humans in their acoustic

environment. As Scot Gresham-Lancaster and Peter Sinclair point out, “[s]ound encompasses our fundamental sense of where we are and what surrounds us at any given time” (71), emphasising the importance of aural surroundings for human orientation. This illustrates the essential role sound plays in our lives, which is underlined by how the brains works to comprehend new acoustic impressions: “[sound] triggers our memories of what we have heard before. Memories of sound work to cluster all new sounds into our understandings of the world, which are connected through the gestalt of those memories to meanings and feelings from other times and places” (71). This means, sound can not only be processed as auditory information, but it also stimulates the cognitive ability to remember similar sounds and associate them with certain sentiments. This will be of particular importance in chapter 3.

With respect to cognition, it is further interesting to consider the above in light of short-term and long-term auditory memory. The term ‘echoic memory’ is used for short-term memory, referring to “the few seconds of auditory information we seem to be able to store and reproduce immediately” (Bijsterveld 7). Although our ‘echoic memory’ does not last long – according to psychologists, certain tones fade away after only a few seconds – it still lasts longer than our ‘iconic memory’, referring to the visual input, which is usually gone after milliseconds (7-8). This is due to functionality: generally, we are able to look at a visual source over and over again for some time, while this does not work for sound because of its temporality. Thus, it is typical of the human brain to remember sounds longer than images (8). This is important to realise in terms of the relationship between sight and sound, which will be discussed in subsequent chapters. Regarding the long-term memory, it seems paradoxical, at first, that humans are capable of recognising auditory information such as familiar voices or other identifiable sounds, given that acoustic details such as tone fade away after a few seconds. To clarify the difference and equally emphasise the significance of the human long-term memory for this thesis, I will briefly elaborate on a study on the storing of acoustic information in the following.

In the context of Reactivation Studies, experiments by István Winkler and Nelson Cowan are designed to examine acoustic details in that the storage of acoustic information in human brains is shed light on and discussed. The psychologists have investigated which types of sounds can be kept long-term in our mind, why this is the case, and how these auditory

memories are accessible (Winkler/Cowan 3). On the basis of their experiments that used a model of auditory recognition, they suggest that sensory information lasts in memory due to the “Regularity-Record Plus Anchor” hypothesis (Winkler/Cowan 16). Winkler and Cowan explain that the acoustic memory is a record of recurrences, with the last auditory information received being the anchor for this record and consequently, the “reactivating reminder” in this process (16). Thus, the results obtained with the “auditory memory reactivation paradigm” cause the psychologists to argue “that the brain stores features of individual sounds embedded within representations of acoustic regularities that have been detected for the sound patterns and sequences in which the sounds appeared” (Winkler/Cowan 3), for instance familiar voices which one has not heard for a long time. This means that sounds which are recognised in a familiar auditory situation are more likely to be stored and that we can draw on these records by reactivating them through the anchor, which reminds us of the records remembered (Bijsterveld 8). What can be concluded from Winkler’s and Cowan’s Studies is that sound is capable of bringing back memories and trigger certain emotions connected to them. I will investigate in Chapters 2 and 3 what this means with regard to the engagement with fiction and highlight which implications come along with this type of auditory sensory memory.

1.2 Categories of Sound Qualities

There are several properties of sound with which we are familiar due to our everyday experience with sound. We have developed a personal sensitivity for how we perceive sounds, for instance as pleasant, unpleasant, soft, musical, or noisy. However, to be able to conceive and identify sounds and to understand how they can be manipulated, it is useful to classify them into categories that help to build a basis for the analysis of my objects of study. In the following, I draw on David Bordwell’s and Kristin Thompson’s classification as well as on the categories David Sonnenschein has established in *Sound Design* (2001). American film theorist and film historian David Bordwell is Professor of Film Studies, holds a master’s degree and a doctorate in film and has published numerous works on fiction film and cinema, among them the ninth edition of the film textbook *Film Art* (2010) together with his wife Kristin Thompson. Thompson, also an American theorist and Ph.D. in film studies, has not only co-written this seminal textbook, but published several volumes on the subject of cinema herself, as for

instance *Storytelling in Film and Television* (2003) or *Breaking the Glass Armor: Neoformalist Film Analysis* (1988), which is considered to be her breakthrough in the scholarly field of film studies. Musician and filmmaker David Sonnenschein was first noticed as a graduate student at USC School of Cinema-Television, where his thesis film won the Motion Picture Sound Editors Golden Reel Award for Best Sound. From that point onwards, with degrees in Cinema/TV, Music, and Neurobiology, he developed a unique career in the area of sound design. He not only became a producer and director, but also managed to write *Sound Design* (2001), a guide that explores the potential of sound in visual media and teaches the expanding possibilities of sound design by explaining step-by-step methods of creating sound.

The first category of sound that is introduced in both, Bordwell and Thompson's and Sonnenschein's work, is that of intensity. This type indicates the loudness, or volume of a sound. Since sounds result from vibrations in the air, the amplitude of this vibration produces our impression and sensitivity of softness and loudness (Bordwell/Thompson 273). The range of the intensity of sounds, meaning how high or low the vibrations are, can be measured in decibels, a "scale of sound energy with each ten points representing ten times the loudness" (Sonnenschein 66). For orientation, what is considered the lowest threshold of hearing is zero decibels, whereas the highest threshold of hearing is around 130 decibels, which can cause actual damage when the sound is continuous. Within this range, hearing rush-hour traffic, for instance, is usually measured at 70 decibels and a jet take-off at around 120 decibels (ibid.). The intensity of a sound is particularly relevant to human's perception of distance, as sounds appear louder, the closer we perceive them (Bordwell/Thompson 273). This implies that the intensity of a sound significantly influences human awareness of it, depending on its volume, and determines the perception of the acoustic environment.

Pitch is the second category, which is all about the frequency of sound vibrations, meaning how high or low a sound is perceived (ibid.). Frequency is measured in Hertz, or cycles per second, which serves to track how fast a sound wave goes from the highest to the lowest pitch and back up. The measurement of how fast this cycle is repeated is then called frequency, the parameter of pitch (Sonnenschein 67). In normal hearing, frequencies range from about 20-20,000 Hertz. The lowest pitches, called infrasonic, are barely audible but rather felt in the body, and the highest pitches, termed ultrasonic, are also not audible as they are above the human hearing range (ibid.). For orientation, what is ranged in at 27 Hertz is the

lowest note on the piano, whereas the highest note is ranged in at 4,186 Hertz. Nevertheless, most sounds we hear are not pure, but complex tones that are composed of different frequencies (Bordwell/Thompson 273). A significant feature of pitch is its ability to help us differentiating various sounds or objects, for example in film, since we can distinguish noise, speech, and music because of its frequencies (ibid.). I will examine the characteristics of pitch in the respective objects of study and investigate it in terms of acousmatic sound in the films.

Another category, that is especially important for the description of a sound's texture, is that of timbre. Timbre is also referred to as tone, or the colour of a sound, since "the harmonic components of sound give it a certain color" (Bordwell/Thompson 273). When a tone is played, what is audible is not only the fundamental tone, but also other subtle frequencies: the harmonic overtones (higher frequencies) and subtones (lower frequencies), referred to as harmonics. Thus, in terms of technique, they can be best exemplified based on instruments since any instrument has a fundamental frequency. These different harmonics "have regular waveform patterns, defined as timbre" and define the overall colour of a sound (Sonnenschein 68). Timbre can also be linked to chapter 1.1, in which I illustrated the impact sound has on memory and the processing of our acoustic environment, since in everyday life, we recognise familiar sounds mainly because of several aspects of timbre (Bordwell/Thompson 273). This category is interesting not only in terms of recognising different voices, but also in respect to the disembodied voice and which effects timbre has on our experience of hearing a voice without seeing its visible source.

Although intensity, pitch and timbre are the most essential categories of sound, there are further important acoustical parameters that contribute to a sound's overall texture, which is why I will combine them in this subsection. Sonnenschein elaborates on these parameters by a further classification of the qualities of speed, shape, and organisation in *Sound Design*. Speed simply describes the velocity at which sound travels, meaning how slow or fast "acoustic impulses are repeated" (Sonnenschein 68). How we perceive the speed of a sound firstly depends on physical conditions of the environment, and secondly on the actual beats per second of which a sound consists of. When a sound stops or is at least not audible longer than a second, as for instance a resting cardiac pulse, we are not attentively listening anymore, whereas we have problems to follow individual sounds if these are emitted at or over 20 beats per second (ibid.). A sound's shape, also called "envelope" (ibid.), is about the

duration of a sound. It is characterised by attack, body, and decay, which describe how quickly the sound grows and falls off, which is measured in ranges “from more impulsive to more reverberant” shapes (ibid.). A quick and intense gunshot for instance would be considered impulsive, whereas the wind howling through an alley or tunnel would be called reverberant. How humans perceive a sound’s shape is determined by the conditions of the environmental space, as well as by the sound source’s waveform (Sonnenschein 68,69). In addition, the feature of organisation is mainly connected to personal sensibility and social background, as it refers to how organised or chaotic sound appears to our ears (Sonnenschein 70). On the one hand, this can be applied to foreign languages or dissonances in music that would seem chaotic to our ears, and on the other hand to languages we understand or musical harmonies, which would sound organised (ibid.). All categories mentioned are significant to comprehend the fundamental properties of sound and serve as a basis for the analysis of my objects of study. All of the above-mentioned qualities are intertwined and, together, contribute to the overall texture of a sound, which shapes the way we experience our acoustic surroundings.

1.3 The Use of Sound in Film Production

Now that I have introduced crucial characteristics of sound, it is not only important to consider these in everyday life, but to investigate them within the world of film production. Similar to our environment in real life, we perceive a film primarily through our visual sense, which becomes apparent merely through the expressions regularly used: we “watch” a film, or we are the “viewers” of a film. As Bela Balazs states, this is mainly because our ear “is not less sensitive, it is only less educated than our eye” (122), implying that we have not grown accustomed to using our senses equally effective. We have learned that our visual sense can be sufficient to explore everyday life without conscious support of our sense of hearing. As Bordwell and Thompson rightly assert, what makes sound a more difficult and challenging discipline than vision, is its fast transience, since we cannot stop a film to freeze a certain moment of a sound in order to analyse it, as it is the case for images (269). We can freeze pictures and look at them for as long as desired, but to achieve a similar effect with sounds, we would need to play back a recording repeatedly. It is therefore true that “the sounds and the patterns they form are elusive” (ibid.), which conveys the power sound has to “achieve very strong effects and yet [to] remain unnoticeable” (ibid.). Thus, my aim is to raise

awareness for the importance of sound in film in equal measure, to then comprehend and investigate how sounds can be employed and manipulated and how this influences our perception of the film in its entirety.

In order to understand the development and changes that sound has undergone in the film industry, I will illuminate important milestones in the history of sound, which have made film sound what it is today. The revolutionary inventions of the first electroacoustic devices such as microphone, loudspeaker and telephone in the mid-19th century have not only stimulated a huge technological development but also provided the opportunity to handle and control sound for the first time (Goldsmith 6). They further signified the “birth of the electronic communications revolution” (Ihde 4), which enabled people “to listen farther than any previous human generation” (ibid.). The invention of audio communication technology has further advanced progression in the realm of the amplitude of electronic signals, which was then first developed in 1906 by Lee De Forest in the shape of the triode vacuum tube. This first audio amplifier and its further development revolutionised electroacoustic technology, since “[l]ife could be captured, recorded, and analysed as never before” (7).⁴ Concerning the history of the coming of sound in the film industry, Warner Bros.’s release of *The Jazz Singer* in 1927 can be considered a milestone, as it marked the end of the silent film era and heralded the start of sound films (Sergi 12). Through this film, regarded as the first talkie, meaning the first film with synchronised dialogue, “sound became a tool by which ordinary films could be made more appealing to an audience than they had been” (Tankel 24-25). This is because the introduction of sound makes the viewer’s experience more complete since image and sound are unified and the viewer’s identification between these as such becomes solidified. After the adoption of sound in the industry, the 1930’s and 1940’s were an era in which early equipment for sound mixing and equalisation, as well as reverb devices and film-editing techniques emerged, which have helped sound films for a solid presence in the industry (Alleyne 25). Whilst the following years were marked by sonic improvements in terms of tape recording and the mass marketing of stereo records, “the last third of the twentieth century [was] largely defined by a handful of groundbreaking technological innovations” (Alleyne 30), such as the Dolby Corporation’s noise-reduction technologies, which changed the way film

⁴ For a more detailed contemplation of the technological innovations in the 19th and 20th century and its effects, see “The Seven Basic Technological Effects” in *Sound: An Acoulogical Treatise* (2016).

sound was created, recorded, and reproduced. Characterised as another milestone in the late 1980's was the digital instrument revolution, which innovated the music industry through synthesizers, drum machines and samplers, which was simplified by a solid level of standardisation (Alleyne 37). Since then, technological developments have continuously evolved, "characterised by a remarkable ability to identify new markets and possibilities", such as "digital broadcasting", the "computer world [...] and personal stereos" (all Sergi 181). The constant evolution in sound technology has offered and is still offering and shaping the film industry most importantly in two respects: filmmakers can draw on a sophisticated repertoire of powerful storytelling tools in terms of creative opportunities, and the audience is able to relive filmmaking and experience the results differently as time progresses.

In addition to technological developments, I will focus on the filmmakers' creative use of sound and its effects, as well as its relation to other film elements in this passage. Sound in film as we know it today, or, as Barbara Flueckiger states, in "the classical Hollywood era" (151), is mostly produced in a studio environment, which has caused a change in the production of sound effects, as they are, unlike in the past, produced afterwards. There are mainly three types of film sound: speech, music, and noise, all of which play a crucial role in the creation of a film and the production of the soundtrack. As Gianluca Sergi rightly asserts, it is important that the soundtrack is considered as a whole, consisting of combined sound elements, rather than being assumed to be only the film's music track (6). For Sergi, it is the complex combination of effects, music, dialogue, and silence that composes the structure and core of the soundtrack (6-7). Revealing this intricacy, it is useful to examine the manner in which sounds in film are created and manipulated. Important to realise is the fact that the majority of sounds are not recorded during filming, but added freely, so that images and sounds can be synchronised deliberately. For this purpose, sound designers often have recourse to a repertoire of sounds which are stored in a type of sound library in order to re-use them if desired (Bordwell/Thompson 274-275).

There is not only a sound library, but several other tools in the sound studio, which are helpful for creators to implement sound in film. For instance, equipment and mechanisms such as "directional and shielded microphones, sound engineering and editing" and "camera blimps to absorb motor noise" are frequently used tools (Bordwell/Thompson 275). With the help of these, editors can select sounds and manipulate them in a way that a certain sound

fulfils a particular function in terms of the soundtrack. This is done intentionally to guide the viewer's perception not only of sound, but also of the respective image and the action happening in order to manipulate the cognition (ibid.). Achieving a shift in the viewer's perception and attention is not only viable through the equipment's possibilities in a sound studio, but also through the change of acoustic qualities, such as discussed in chapter 1.2. For example, the property of loudness can easily be modified in order to determine the perceived distance of a sound. Pitch is another quality, which is frequently manipulated and adjusted to emphasise one sound or set it apart from another (Bordwell/Thompson 275-276). I will analyse in the respective chapters how the selection, manipulation and mixing of sound shapes the characters' as well as the viewers' experience and in how far this process can be misleading.

Since it will be important in subsequent chapters how sound and sight influence each other, I will draw on Bordwell and Thompson's description of the dimensions of film sound in this last section of the first chapter. There are four main dimensions sounds can occupy when it relates to other film elements: rhythm, fidelity, time, and space (Bordwell/Thompson 280). The first dimension, which is rhythm, might be the most famous among laypersons since it is best known for occurring in music. There are, of course, other sounds than music that can have a rhythm, as for instance a clock tick, a heartbeat, or the sound of a galloping horse. However, it is essential that the rhythm involves "a beat, or pulse; a tempo, or pace; and a pattern of accents" (Bordwell/Thompson 281), as it "characterizes sound through time" (Sonnenschein 65). There are two main different uses of rhythms in film and how they can be combined. The first use consists of a matching of audible and visual rhythms in order to achieve a coordination among sound and image. The second, contrasting way to employ rhythm is to generate disparities between the images and the rhythm of sound. Here, the editing of images does not correspond to the rhythm of the sound. The editor even can create a "counterrhythm" (Bordwell/Thompson 282-283), which can help to enhance the desired effects in terms of atmosphere.

Fidelity plays a role in regard to a sound's source, conveying the sound's level of faithfulness in relation to its source. This means, a sound can be considered faithful when it corresponds to what the viewer sees. This would be the case when the viewer sees a car passing by, hearing its engine. If the car was replaced by a truck, but the viewer still hears the

car's engine noises, then a disparity between the image and the sound would occur – “a lack of fidelity” (Bordwell/Thompson 283), since it is not realistic for a truck to be as quiet. In this case, the actual source in production does not matter, because the point is whether the viewer takes the sound to be coming from the diegetic world of the film. Faithful and nonfaithful sounds or shifts in fidelity can have various consequences on different levels, as for instance in a comedy or in a drama, which lead to distinct effects on characters and viewers. The next dimension concerns space, since “[s]ound has a spatial dimension because it comes from a source” (Bordwell/Thompson 284). Here, it is distinguished between diegetic and nondiegetic sound, with diegetic sound defined as coming from the story world, and nondiegetic sound as belonging to the outside of the story world (ibid.). I will come back to these two types of sound in the next chapter, still, it is significant to mention both types at this point, since “a film's narration deliberately blurs boundaries between different spatial categories” (Bordwell/Thompson 285). I will investigate in how far the blurring of boundaries is present in the respective films and how this works in respect of acousmatic sound.

The last dimension is that of time, which is crucial in terms of synchronisation of the soundtrack and the image. Time can be represented in different ways, for instance as “synchronous” (Bordwell/Thompson 294) or “asynchronous sound” (ibid.). The best known example for synchronous sound is that of lip synchronisation in a dialogue, since we normally hear the voice of someone at the same time as we see a person's lips moving. This is in contrast to the case in which sound is labelled asynchronous when a sound is “out-of-sync” (ibid.). Sometimes this happens due to an error in production, but it is also often used intentionally for reasons of distraction or humour in order to play with the viewer's expectation (ibid.). I will further examine the illusive effects of the representation of time through sound in the analysis part of my paper. Overall, the overlapping and merging of the different dimensions of film sound introduced, is created for the desired effects to be achieved flexibly. In addition, sound's complex relations to other film elements demonstrate the multi-faceted nature of sound and its pivotal role in the world of film production.

1.4 From Physical Hearing to Interior Sound Perception

In this section I will highlight the complex way that hearing and listening work, as it is an essential part of the physical and psychological effects sounds have on humans. Often, we neglect the difference between hearing and listening, because we are unaware of the significance and diverse implications both have on our lives. Therefore, I will explain the difference between the two, but also illustrate the connectedness of physical hearing and psychoacoustics. Michel Chion states that through the absorption of sound through the ear there is “a link between different worlds (real and imaginary) and different registers (physical and mental)” (*Sound: An Acoulogical Treatise* 18). I agree with him that the ear’s special function is to be an external and internal organ at the same time. The connection between these areas is crucial to the understanding of how sound affects the perception of auditory experiences. Accordingly, I will first illustrate the physical process of how sound reaches our ears and is transferred to our bodies, and then clarify the difference between hearing and listening by focusing on the interior perception of sound. Subsequently, I will elaborate on the different listening modes through which sound is perceived to show how mental capacities influence human perception of sound.

1.4.1 How Hearing Works

The ears are our organ for hearing and responsible for perceiving and distinguishing sounds. The hearing process is essential as it connects us to the soundscapes in our environments. As early as the 1970s, the composer R. Murray Schafer used the term “soundscape” to refer to “any acoustic field of study” (7), for example, a piece of music, a radio or television programme, or any other kind of acoustic environment, which “consists of events *heard* not objects *seen*” (8). In my thesis, I will regard and isolate the films, or film scenes, as individual soundscapes to establish a connection between the media and the way these films are heard and interacted with. It is important to consider the different dimensions of soundscapes and how humans perceive them, since this process is complex and multi-faceted in its entirety. As Per Magnus Lindborg argues, “[p]erception depends on expectations [...] [w]hile we react physiologically to some aspects of soundscapes in ways determined by our biological setup, other aspects are dealt with in cognitive, individual ways [...], moderated by personality traits”

(48). It is true that perceptual phenomena vary depending on specific personal aspects, especially in terms of cognition. Evidence for this will be provided in the detailed analysis of the objects of study in subsequent chapters.

The human hearing system is capable of absorbing sound waves from the environment, before sending the sounds on an internal journey to our brains where they are processed. The ear is divided into three basic parts: outer, middle, and inner ear.

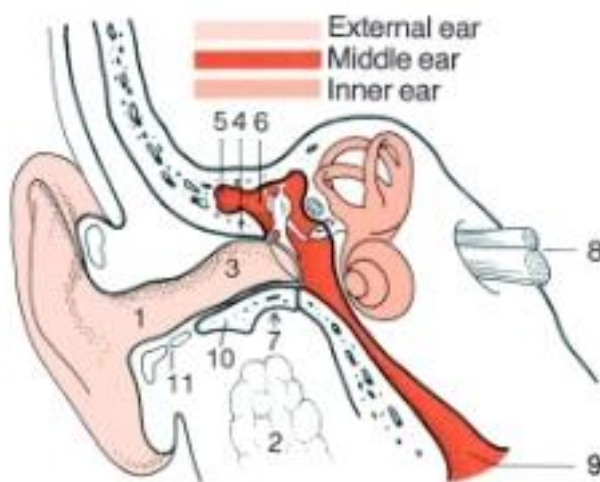


Figure 1. Illustration of the structural relationship between the outer, middle, and inner ear (Menner 14).

The outer ear, also called pinna, acts as an “apparatus both of *protection* [...] and of *resonance*” (Chion, *Sound: An Acoulogical Treatise* 18). This means for the latter, it acts as a funnel to gather sound waves and then direct them towards the eardrum, or “tympanic membrane” (ibid.), which is located in the middle ear. The sound waves cause the eardrum to vibrate, which actuates the ossicles, a series of three tiny bones that form a chain and are active in the air-filled middle ear (Goldsmith 56). The motion of the ossicles is supposed to transform “airborne pressure vibrations into mechanical vibrations” (Chion, *Sound: An Acoulogical Treatise* 19), which are then passed to the gateway in the inner ear: the cochlea. This main organ of the inner ear is filled with a fluid, which is activated and moved by the vibrations of the oval window. Through this movement, the sound waves now push against hair cells, which in turn are connected to around 30,000 neurons (ibid.). These neurons are set in motion, meaning that the hair cells turn the vibrations into electrical nerve impulses, which contain different pieces of information, for instance how many nerve fibres are signalling simultaneously or the length of intervals between bursts of fibre signals. These are decisive

factors in the brain's processing of pitch and loudness of a sound (Goldsmith 58). The electrical impulses then travel along the auditory nerve, which connects the cochlea with the auditory centres of the brain, where speech, language and hearing activities are located. The "primary auditory area" (Goldsmith 59) is responsible for recognising nerve impulses as sounds, as it extracts specific features from the input data to use them for the improvement of a "mental model of the thing being listened to" (ibid.). This process is called "hierarchical encoding" (ibid.) since the brain is considering which element of a sound serves the meaning. The decodification of sound waves, which have to come a long way until they are experienced as sound, is a deeply interesting and complex mechanism, which enables us to recognise certain sounds or familiar sound patterns (ibid.). 'Hierarchical encoding' plays an important role in relation to acousmatic sound as it allows the brain to process a sound and determine whether it is accompanied by visual physiognomy, as will be further analysed in subsequent chapters. As demonstrated in this section, physical acoustics is indispensable for human perception to connect with surrounding sound fields.

1.4.2 Sound Perception and Psychoacoustics

Since individuals experience sounds differently, an important part of the research on the human hearing process is the analysis of the cognitive perception of sound in the realm of psychoacoustics. The field of psychoacoustics is defined as "[t]he study of the relationship between sounds and their psychological effects" (Colman n.p.), which represents the interface between physical acoustics, meaning how sound waves are sensed by the ear, and the corresponding interior sound perception. In psychoacoustics, the aim is to measure the human perception of sound to associate the observations with psychoacoustical phenomena (Gold/Morgan/Ellis 2009). The fundamental question dealt with is why we perceive sounds the way we do, including how the ear responds to different intensities or frequencies (ibid.). I will elaborate on these issues relying on scientific studies that measure varied sound characteristics and their differences, to discover why sound events are perceived distinctively by humans.

One of the factors that evokes individual reactions to sound events is in the frequency of sounds. As Tomlinson Holman states, "human hearing does not respond equally well to all

frequencies in the audible range" (33), because humans are most sensitive to the high frequency ranges and less sensitive to lower frequencies, as investigated in diverse psychoacoustic studies. A decisive aspect that plays a role in this context is that of loudness and how it varies with frequency (Holman 34). Both sound characteristics, of which "the perception of frequency is called pitch" and "the perception of intensity is called loudness" (Gold/Morgan/Ellis 209), are interconnected elements, which have a significant impact on the human auditory system. Pioneering experiments conducted in the 1930s by Harvey Fletcher and Wilden A. Munson contributed notably to the insight into human perception of loudness, as they produced so-called 'curves' of equal loudness in one such series of experiments. The researchers have established definitions for the intensity level of sound, its physical composition, the manner of listening to sound, and the psychological and physiological conditions of the listener. These allowed them to determine the 'loudness', which is the "magnitude of an auditory sensation" (Fletcher/Wilden 83), of a sound. As Leon Gunther explains, the equal loudness curves "reflect an average of individuals' responses to sound" (284), resulting from the measurements of how a "typical listener[s]" (Fletcher/Wilden 84) ears are sensitive to equal loudness within the frequency range of sound. For example, in one of the experiments people are exposed to two different sounds - the "reference sound" (Fletcher/Wilden 83) and another added sound - both of different frequency, but of the same intensity level, meaning of equal loudness. In a nutshell, the outcome of the experiment was that the majority of participants reported that they had perceived the loudness of the two sounds differently (Gunther 284). This shows that the change, mostly the increase in loudness, depends on its frequency in relation to the reference sound. Fletcher and Wilden state that "the value of the intensity level of the equally loud reference sound depends upon the manner of listening to the unknown sound and also the standard of reference" (83), revealing that the individual mode of listening is responsible for how the ear responds to loudness. The factor of the "*integration time* of hearing" (Holman 35) must be considered in this context, as loudness changes differently for more rapidly changing sounds than for continuous sounds: "[l]onger sounds of a given intensity sound louder" (Gold/Morgan/Ellis 216). Another equal loudness curve-based incident that is related to a shift in frequency is provided by Holman, who illustrates how hearing works at levels higher or lower than threshold. He exemplifies that when the frequency of a sine-wave tone at 1 kHz is changed to 100 Hz, it is "dropped below the threshold of hearing and no sound is perceived at all" (Holman 34). Thus, he states that

“the frequency of a tone affects loudness, as does its sound pressure level” (ibid.), which contributes to the scientific explanation for the individual reactions to sound events. The correlation between frequency and loudness is significant in that it is often played with and taken advantage of to create film sound.

A powerful film technique often applied in the context of a sound’s loudness and frequency is that of masking. The idea of “frequency masking” (Holman 36) conveys that “[l]ouder sounds cover up softer ones, especially those that are nearby in frequency” (ibid.). This means that acoustic impressions can be deliberately induced to cause a certain effect in the listening experience. For instance, this effect is regularly used in post-production to adjust background noise from different shots to an equal volume level, or to contrast background sound, such as music from the sound of voices in a conversation (Holman 36-37). Experimental results have not only shown that tones close in frequency have a greater effect as they can easily mask one another, but also that “a tone more easily masks a tone of higher frequency than one of lower frequency” (Gold/Morgan/Ellis 214). In some cases, this means that the weaker tone may not even be heard anymore, especially if it is a higher frequency. When a sound is distracted by the presence of another sound not precisely at the same time, this case is called “temporal masking” (Holman 37). Consequently, the masking effect intentionally occurs as a manipulative device in film sound to confuse the listener’s perception of the simultaneous sounds (Gold/Morgan/Ellis 214). This psychoacoustic phenomenon constitutes one specific aspect of the complex hearing system and demonstrates how sensitive our perception of sound is, as well as how easily it can be out of balance.

1.4.3 Listening

As a multi-dimensional process, listening not only enables the perception of sound, but it is further responsible for how humans perceive sound subjectively. This subsection will build on the previous one by portraying the difference between hearing and listening to explore listening as a cognitive phenomenon. As I have illustrated in chapters 1.4.1 and 1.4.2, the process of hearing encompasses the functionality of physical mechanisms by which the ear absorbs environmental sounds, which thus constitutes a pure physiological process. Listening, however, is far more complex, as it is not only about receiving sounds, but also about

interpreting and converting these into meaning. While we automatically hear the sounds around us, since we cannot close our ears the way we can close our eyes, we listen to them by choice, depending on whether we want to make sense of the sounds we hear.

As one of the major communication competencies, listening as a social function plays a huge role in everyday life. In academia, it has been defined in a number of ways, yet the condensed definition proposed by the International Listening Association (ILA) seems to be a suitable starting point for further investigations. Listening is “the process of receiving, constructing meaning from, and responding to spoken and/or nonverbal messages” (qtd. in Worthington/Fitch-Hauser 7).⁵ As Debra Worthington and Margaret Fitch-Hauser state, there are three major complex categories, in which most definitions of listening can be classified: “affective processes”, “behavioral processes” and “cognitive processes” (8), the latter of which “address attention, understanding, reception, and interpretation” (ibid.). I will focus on this category in my work, since listening as a cognitive activity plays an essential role in analysing sounds’ capabilities to influence and manipulate human perception. Key elements included in cognitively oriented approaches are “attention and memory” (Worthington/Fitch-Hauser 10) among other related elements such as “inductive reasoning”, “cognitive complexity”, and “receiver apprehension” (ibid.). When analysing why these aspects are important characteristics of cognitive processes, it is crucial to consider “[a]ny act of listening [as] interpretive” (Schnell 3). Equally significant is the individuality of these interpretive processes, since every human comes from a different background and is confronted with various experiences, which often induce situations in which two individuals listen to the same sound and understand it differently (ibid.). This is especially the case in “interpersonal communication” (ibid.), for instance in conversations among people or in general when listening to someone else speaking. According to Jim Schnell, listening, however, can be categorised as “intrapersonal communication” (ibid.), which happens within ourselves as we process the sounds heard and reflect on their meaning. Again, this process is mainly dependent on past experiences (ibid.) and listener expectations (Schnell 4), as well as “physical and psychological differences between people”; for example, temperament and personality traits (Worthington/Fitch-Hauser 69). An individual’s context or situation plays a major role in the listening process, as it clearly affects the way humans listen and react to sound. I will

⁵ See Worthington/Fitch-Hauser, especially table 1.2 in the introduction, for a detailed overview of scholarly definitions of listening.

explore how humans process and store information individually in greater detail in chapter 3 when talking about strategies of cognitive processing.

The listening process is a significant component within the scope of humans' active and passive perception of sound – a number of elements make up this process as there are different ways of listening. Being a multi-dimensional skill, listening competency is reflected in the three modes for which Michel Chion coined the terms: "causal listening", "codal listening" and "reduced listening" (22). "Causal listening" is the most prevalent mode and consists of listening to a sound in order to detect its source or cause (ibid.). This type of listening becomes more complicated when the sound's source is not visible, and thus, the sound itself becomes the main source of information. Although the invisible source can be "identified by knowledge or logical prognostication" (ibid.), causal listening is also prone to manipulation, as it can easily be influenced and become deceptive for the listener. As Chion argues, causal listening "can take place at various levels" (23), meaning that there are cases in which we can identify the exact cause – for example, a familiar or unique person's voice or the sound produced by a specific object – and instances where this is not as easy. There is another type of causal listening, in which we are unable to recognise an individual but rather a group of certain sounds, such as car engines, sirens or birds singing (Chion 24). In another kind of causal listening, in "still more ambiguous cases" (ibid.), it is likely that all we recognise is "only the general nature of the sound's cause" (ibid.) – for example, an animal or human voice, which we cannot allocate. These types suggest that causal listening is definitely "associated with context, and [...] manipulated by the audiovisual contract" (ibid.). "Codal listening", also referred to as "semantic listening", implies that the auditor "aims to decode the signal to get the message" (Chion 25). This means, that a code or language is interpreted to understand its content. The most common example would be spoken language, but also Morse or other such codes are included here (ibid.). It follows that causal and codal listening can be employed at the same time while listening to a sound, for instance when someone is speaking to us via telephone and gives us information through speech (ibid.). The third category, which is called "reduced listening", deals with "the traits of the sound" itself, "independent of its cause and of its meaning" (ibid.). This implies that this type of listening examines the sound as "the object to be observed" and not as a deliverer for any other kind of message (ibid.). I would argue that reduced listening is the rarest mode because, as I have indicated, we are not used to listening to a sound for its own sake. This is also shown in certain reactions people have when paying

attention to sound in themselves, since a majority is “laughing off” (Chion 26) the project in different ways. Anyway, when for instance a musical piece is analysed and a regular beat or the pitch of a note is identified, this is called reduced listening, since pitch is “an inherent characteristic of sound” (Chion 27) and “independent of the sound’s cause or the comprehension of its meaning” (ibid.). I agree with Chion, who illustrates the advantages of this kind of listening, namely that it “open[s] up our ears and sharpen[s] our power of listening” (28). This is often played with in films, especially in acousmatic situations (ibid.). In further chapters I will investigate how far ‘reduced listening’ and the above-mentioned types modify listening and the way it behaves in relation to perception.

2 Sound in Film: The Acousmatic

Sound is not only a complex phenomenon that is still often subordinated to the visual sense, but it is also a wide-ranging and influencing technique used in film production. In this chapter, I will build on the introduced dimensions of sound and analyse them especially in the light of the concept of the ‘acousmatic’. I address the ‘acousmatic’ by drawing on Pierre Schaeffer’s theory and his definition of the concept in *Treatise on Musical Objects* (2017), and on Brian Kane’s interpretation of acousmatic sound, who is taking a different approach in *Sound Unseen* (2014). Additionally, the use of auditory devices in film according to film theorist Michel Chion will be examined. Here, the focus will lie on his understanding of the interdependent interplay of sound and images, as well as on how sound influences human perception in different ways (Chion 34). The following subchapters discuss these auditory phenomena and focus on the acousmatic zones (Chion 74ff.), ambient sound (Chion 75), music (Chion 79), and added value (Chion 5ff.), all of which contribute to the functioning of the audiovisual system in film. I will investigate the function of the voice in this chapter’s last subsection, drawing on Chion’s theory of the voice in cinema, illustrating that the voice takes on an exceptional role in the analysis of the sound domain. Furthermore, I will refer to the selected films and give examples of scenes in which listening conditions are changed by means of acousmatic situations.

2.1 The Concept of the Acousmatic

Visualised sound is the most common type of sound employed in film: it is coupled to the image, which usually means the viewer can see the sound's source on screen simultaneously, as opposed to acousmatic sound. The term 'acousmatic' was first used by French composer Pierre Schaeffer in *Treatise on Musical Objects* (2017), originally published in 1966, which is regarded as his most important work on music and its composition. In his book, the adjective 'acousmatic' is defined as describing "a noise that is heard without the causes from which it comes being seen" (Schaeffer 64). To begin with, these 'noises' can be any sound whose source is invisible to the characters in the story world, the audience, or both. Nondiegetic sound such as background music, as in the very first scene of *Iron Man* [00:01:11-00:01:32], can be characterised as acousmatic. As audience, we hear ACDC's rock song "back in black", which we know is digitally integrated but whose source we cannot see. In contemporary films it is common that the soundtrack opens with a nondiegetic piece of music, as in this example, or more generally with any kind of sound effect before the images appear on screen. At the beginning of *Lost in Translation*, we hear a voice making a loudspeaker announcement at the airport [00:01:09-00:01:16] while the screen is black and we neither see the source of the voice nor any other image. As soon as the announcement is over, one of the main characters is introduced on screen, which is accompanied by the song "Girls" by the electronic music group Death In Vegas [00:01:16-00:02:33], which is known as the original soundtrack of the film. The examples described represent the most common type of acousmatic sound through nondiegetic music, often used to introduce a new setting, character, or surrounding in a film.

Aside from that, acousmatic sound can also occur as taking place within the story world. Diegetic sound with its source in the story space, is complex and can be detected in all of my objects of study. To allege some examples: in *Iron Man*, an acousmatic voice which is clearly heard by the character in the story, starts talking and reports on several facts concerning the weather and climatic conditions for the upcoming day in Malibu [00:09:30-00:09:41]. In the following scene, the exact same voice startles the woman in Tony Stark's house, as she touches a regulator of the house's computer system [00:10:33]. Soon afterwards, Pepper Potts introduces this voice as J.A.R.V.I.S., an artificially intelligent system to assist Tony. Another classic example of acousmatic sound in the diegetic world is that of phone calls. In *Lost in Translation*, Bob occasionally calls his wife from the hotel room [e.g.,

00:53:16-00:55:07]. Different than in *Phone Booth*, where the audience can see the person at the other end of the phone at the same time, we never see Bob's wife but only hear her as an acousmatic voice on the phone. Sometimes, the distinction between diegetic and nondiegetic sound is blurred in films, as for instance in the beginning of *Stranger Than Fiction*. We are introduced to the storyline by a seemingly all-knowing narrator. The female acousmatic voice accounts numerous attributes and facts about Harold, the main character who is introduced on screen simultaneously. First, the narrator reports on mundane situations in Harold's everyday life, such as getting ready in the morning and nearly missing the bus every day. Harold performs the actions the same moment the narrator tells us about them [00:00:55-00:04:14] and it becomes clear that the narrator must be omniscient, since she knows Harold inside and out. Up to the point Harold interrupts brushing his teeth, the narration does not seem to belong to the story world, as the narrator is neither visually exposed nor heard by Harold. The apparently nondiegetic voice becomes diegetic the moment Harold hears it as well [00:04:14]. Based on Bordwell's and Thompson's approach, the narrator's voice can be viewed as "internal diegetic sound" (290), as it "comes from inside the mind of a character; it is subjective" (ibid.). This becomes obvious later when Harold asks people about the voice, but who are obviously not able to hear it, as for example the woman at the bus stop [00:06:08-00:06:33] or his colleague Dave at work to whom he reveals that he is being followed by a voice [00:07:35-00:07:50]. The extraordinary approach to acousmatic sound in *Stranger Than Fiction* will be of enormous importance in the further analysis. In *Phone Booth*, it is also the acousmatic voice which is at the centre of examination. This film is exceptional in its own way since the unknown caller's voice can be considered the pivot around which the action revolves. In this special case in which only Stu is able to hear the voice, it will be interesting to assign the voice to an acousmatic zone and to ascribe it to the outside or inside of the story world. Nevertheless, it is a different scenario than in *Stranger Than Fiction*, where the narrator's voice is internal and cannot be heard by other characters because it seems to be in Harold's mind.

Acousmatic sound further needs to be considered with regard to space. Brian Kane expands the basic definition put forward by Schaeffer in *Sound Unseen* (2014), stating that "the very acousmaticity of sound – it's quality of being acousmatic – depends on the spacing of source, cause, and effect" (Kane 149), assuming that there can only exist acousmatic sound when these elements are spaced. Schaeffer further states that "[the acousmatic] [...]"

emphasizes the perceptual reality of sound” (64), which takes up the idea of ‘reduced listening’ again, focusing on listening to sound as such. In the traditional acoustician’s approach, “pure listening” (Schaeffer 66), especially when it comes as a surprise, can unsettle the human sense of hearing, since a visible sound source is missing and “sound no longer appears as a medium or placeholder” for something else (Kane 28). This confusion means “we discover that much of what we thought we could hear was in reality merely seen, and explained, by the context” (Schaeffer 66), meaning that another sense, usually sight, is missing to identify the sound source. These approaches serve as a basis for the further analysis and will be unfolded in the following subsections.

Listening in an acousmatic situation challenges the audience especially in respect to the source of the sound, which is often unidentifiable. The ‘sound object’ is neither the tape of a recorded sound, nor the instrument played or any other tool producing a sound, but rather a sound whose source the listener is unable to identify (ibid.). Schaeffer expresses that it is “independent of any causal reference covered by the terms sound body, sound source, or instrument” (67) and an “acoustic signal” (ibid.), which can have different meanings to various people of different cultures or civilisations. The ‘sound object’ is “the object of our listening alone” (ibid.) and “entirely contained within our perceptual consciousness” (ibid.), meaning that it is relative to different manners of hearing and modes of perceptions as well as to the ways sound can be manipulated. Drawing on this, I would like to emphasise one point Schaeffer made about the listening intention, namely that through listening in an acousmatic situation, “values are surpassed, forgotten, redefined, for a unique [...] perception” (115) of the ‘sound object’. Although this is a suitable basic insight in terms of an acousmatic experience’s effect on the listener, it is important to enhance Schaeffer’s definition. The following example may serve as a criticism of this existing view: the main character Stu in *Phone Booth* is bound to the phone in the phone box as the unknown caller is extorting him. In this case it is true that the voice Stu is hearing is the object of his listening and that this acousmatic situation turns his life upside down since he is undergoing a personal crisis. It can be argued that his values change and are redefined through the experience of perceiving the ‘sound object’, as becomes obvious in his emotional outburst when he confesses his infidelity to Kelly, his wife [01:06:52-01:08:15]. However, the question is why Stu is ‘purely listening’, as Schaeffer calls it. Since Stu is being blackmailed, the decisive factor to be considered is that of voluntariness, which is not existent. Rather it is his subjective motive that forces him to stick

to the caller, which calls into question Schaeffer's approach of 'reduced listening' the way he likes it to be conceived.

With respect to the concept of the acousmatic, Brian Kane's theory on sound can be seen as a contrast to the Schaefferian approach. His central claim is that "acousmatic listening is a shared, intersubjective practice of attending to musical and nonmusical sounds" (Kane 7). Kane is explicit on his topic and argues that Schaeffer's ideas on sound do not adequately explain "the cause, the source, or even the production of sound" (10). Therefore, he works with a model of sound that consists of the components of "source, cause, and effect" (Kane 7), constituting the "tripartite ontology of sound" (Kane 224). For this pattern to work, Kane proposes that "[e]very sonic effect is the result of the interaction of a source and a cause" (7, 134), which does not mean that a listener is sure about the sound source by hearing the effect alone. Usually, the environment gives information on the relationship between a sound and its source. However, there are cases in which the source cannot be detected or the sound itself is unclear or dubious (Kane 7). Kane challenges Schaeffer's approach of the acousmatic reduction, which he argues was used by Schaeffer to establish the sound object as part of the phenomenological consideration of sound (138). Rather, he focuses on "the relationship of sonic source, cause, and affect articulated in the acousmatic reduction before any ontological separation has been asserted" (Kane 138). As such, his approach differs in that he investigates sound separate from distinct modes of listening and the ontological sound object (ibid.). The detachment from the phenomenological aspect of sound as occurring in Schaeffer's theory, leads to Kane's idea that there is no need for acousmatic "sonic effects" to be "ontologically separated [...] from their sources" (ibid.). Whereas in Schaefferian theory, a separation of eye from ear would equally mean that the auditory effect is separated from its source and the acousmatic sound "takes on an aspect of intensified profundity" (143), Kane illustrates another attempt within his theoretical framework. Among others, he theorises acousmatic sound in his analysis of the short story "The Burrow", written by Franz Kafka in 1923. There he demonstrates that the main character, who is a mole, is constantly searching for the unknown sources of sounds it is hearing, and frantically detecting possible meanings by constructing different hypothesis about the origin and essence of the sounds (Kane 143-144). Kane outlines the anxiety of acousmatic listening, by claiming that the "numerical identity [of a sound] is always insecure" (147), which evokes uncertainty in the listener. This is what constitutes acousmatic sound, since "the auditory effect, when unaccompanied by contributions from

other senses, underdetermines ascriptions of source and cause” (Kane 147), meaning that the unknown origin cannot be detected without the support of another sense. Not only does Kane use his analysis of the story to critique Schaeffer’s method, but also to promote his own perspective. He criticises that Schaeffer draws a clear line between the sonic effect as the sound object and a sound’s source and cause, as suggested in ‘pure listening’, which grants the sound object independence and suppresses the source entirely. Instead, Kane is convinced that the source and the cause of a sound cannot be completely separated from the auditory effect, as is exemplified in the mole’s constant search for a source (ibid.). Kane’s approach can also be detected in several situations in the selected films, especially in *Phone Booth* and *Stranger Than Fiction*. Both main characters, Stu in the former and Harold in the latter, are continually searching for the source of the sounds they are hearing. Therefore, I mainly agree with Kane’s theory and will explore in more detail how it corresponds to my claim of perceptual manipulation in acousmatic listening in the following chapter.

According to Kane’s definition of the term, the ‘spacing’ of a sound’s components is important in regard to uncertainty and anxiety in acousmatic listening situations. Rather than a strict separation of these components, it offers space for imagination, which I claim is a neglected area in the field of sound studies, especially in terms of cognitive repercussions (see chapter 3). Although it seems as though Schaeffer’s and Kane’s theories both argue that acousmatic sound involves unsettling elements, Kane contrasts strongly with Schaeffer in that he assumes that Schaeffer “misrecognizes what precisely is unsettling about acousmatic sound” (148). Kane emphasises “underdetermination and uncertainty in acousmatic listening” (ibid.) and is missing these aspects in Schaefferian theory by arguing that Schaeffer would not be able to logically explain the mole’s emotional state in the short story. The crucial proposal in Kane’s approach is that “acousmatic sound is constituted by a structural gap” (149), which conveys that the “acousmaticity” of sound, which is the “quality of being acousmatic – depends on the spacing of source, cause, and effect” (ibid.). Kane is convinced that acousmatic sound can only come into being by spacing, which is in turn dependent on the listener’s cognitive state and also on techniques that are applied in the creation of film sound (225-226). I will deal with these approaches in depth in the following, exploring this type of sound’s characteristics in the selected objects of study.

2.2 Auditory Phenomena

Different auditory phenomena assist in involving the audience in the multidimensional world created in film. Building on my previous findings, which have revealed the intertwined interplay of images and sound, I will investigate the different ways in which the selected films emphasise the use of acousmatic sound to challenge the visual superiority on screen. I will explore how sound editors apply and play around with specific auditory means to evoke a cognitive impact on humans within the scope of audio-vision. Afterwards, I will expand on how the human perception can be manipulated through changed properties of sound in chapter 3. Following film theorist Michel Chion's approach of several auditory functions in *Audio-Vision: Sound on Screen* (2019), I will analyse the most important ones in the four films selected. My focus is on acousmatic zones and music in this chapter and on the voice as a major object of investigation in chapter 2.3.

2.2.1 Acousmatic Zones

This subchapter provides a basis for the division of acousmatic sound into subcategories. An acousmatic situation in film can develop in different ways, wherefore theorist Michel Chion originated the three-zone circle, including the acousmatic zones of "onscreen", "offscreen", and "nondiegetic" sound (73-74). For the sake of completeness, diegetic sound is defined as "sound that has a source in the story world" (Bordwell/Thompson 284), such as music coming from instruments in the story space, voices of characters, or sounds emitted by objects in the story world. Diegetic sound can be either onscreen or offscreen, depending on where its source is located. As stated by Chion, "offscreen sound" is sound that is "acousmatic relative to what is shown in the shot: sound whose source is invisible, whether temporarily or not" (73). A suitable example of this is the scene from *Lost in Translation*, in which the main character Bob is sitting at the hotel bar and the audience hears two men talking about him, without seeing who these men are [00:04:35-00:04:50], up to the point where the camera pans to the other side of the table and the men become visible. This first part of their conversation can be considered 'offscreen' since the source is not apparent. Nonetheless, their conversation is part of the story world and thus can be classified as diegetic sound. A different example of offscreen sound is the scene in *Iron Man*, in which J.A.R.V.I.S. is

introduced [00:10:33-00:10:38]. As he is Tony's computer system and has an electronic voice, it is difficult to determine how the source of his voice's sound looks like at this point. However, I would argue that it is offscreen, since we would expect a physical body onscreen when we hear a voice speaking. The woman's reaction proves this, as she is startled when J.A.R.V.I.S. starts speaking [00:10:34]. A sound coming from offscreen should not be conflated with acousmatic sound, since the former means that the sound is outside the visible field but included in the film frame. This implies that 'offscreen sound' is always acousmatic, but this is not necessarily the case the other way round (Chion 72). For completeness, "onscreen sound" is what Chion characterises as "sound [...] whose source appears in the image and belongs to the reality represented therein" (73), which is for instance representative of the sound of voices during a conversation. Although 'onscreen sound' is important in terms of the distinction of the individual zones, it is not my main interest in this study. As "nondiegetic sound", the theorist labels the kind of sound that is "not only absent from the image but also external to the story world" (Chion 74). Particularly characteristic about nondiegetic sound is its temporal unrelatedness to the story world since it has no existence in it (Bordwell/Thompson 297). A typical example of this type would be the narrator's comments throughout a film. We find this for example at the beginning of *Phone Booth*, where an omniscient narrator introduces the audience to the action, starting off by accounting numerous facts about New York City life, its telephone exchange lines, cell phone users etc. [00:02:31-00:03:36]. Afterwards, the narrator's voice introduces the main character Stu [00:03:36-00:03:43] and does not appear again in the film. This nondiegetic narrator remains acousmatic and is clearly external to the story world. A different example of nondiegetic sound can be detected at the beginning of *Iron Man*, where an invisible narrator introduces Tony Stark and summarises his company's history up until today [00:04:51-00:05:52]. What is different compared to the previous example is the fact that we have two audiences here, both being able to hear that voice: the viewers of the film and the audience shown on screen, listening to the presentation accompanied by pictures. Thus, although it seems we are dealing with a narrator being in the nondiegetic zone, it soon becomes obvious that the audience in the story world can hear this voice as well. This scene reveals how it can be played with acoustic means in film and that a sound can exceed or even change acousmatic zones. This phenomenon can also be detected in *Stranger Than Fiction*, concerning Karen Eiffel's voice. At the beginning we do not know that the narrating voice belongs to author Karen Eiffel, one

of the main characters in the story, but are rather presented with all the facts about Harold by a seemingly nondiegetic narrator [00:00:55-00:04:14]. It is only later that her voice is deacousmatized in an interesting way, which I will conduct a closer analysis of in chapter 2.3. This example confirms that there are no fixed boundaries between acousmatic zones, but that these are often blurred to create a certain effect: the manipulation of perception. Chion's basic division between the different zones initiates a better understanding of the fundamental types of acousmatic sound in film and will be analysed in terms of its perceptual manipulation in chapter 3.

After developing his first version of the three-zone circle, Chion revised it as it became more complicated in terms of the variety of sounds that needed to be considered. New categories were added to enlarge the scope of acousmatic film sound. Chion emphasises the importance of regarding the refined circle "as consisting of interlocking sectors" (77), which are enhanced by "ambient", "internal", and "on-the-air" sounds (75-76). 'Ambient sound' is defined as sound that "envelops a scene and inhabits its space without overtly raising the question of identifying or visually confirming its source" (Chion 75), such as everyday background noises of traffic or the twittering of birds in the morning. These sounds are usually also called "territory sounds" (ibid.) since they often designate particular areas through their omnipresence, such as the background sounds of the busy streets of New York in the opening scenes of *Phone Booth* [00:01:29-00:02:30]. The category of "internal sound" (Chion 76) complies with "physical and/or mental interior of a character" (ibid.) and is subdivided into "objective-internal" and "subjective-internal" sounds (ibid.). Whereas physical sounds such as breathing or yawning fall into the former group; inner thoughts, memories, and mental monologues are characterised as "subjective-internal" (ibid.). The voice Harold hears narrating his life in *Stranger Than Fiction* could be argued to belong to the latter category since it seems that the voice is not coming from the real story world, as none of the other characters can hear it. However, I argue that Harold's case is special as it becomes clear that the voice he hears not only comes from within him, but is audible to the audience likewise. Sounds that are "transmitted electronically" (Chion 76), for example via phone calls as in *Phone Booth* or *Lost in Translation*, or radio, are referred to as "on-the-air sound" (ibid.). What is specific about this type is its ability to "travel through space even while remaining in the real time of a scene" (ibid.). In many cases of "on-the-air sound" it becomes more difficult to clearly distinguish between the three main zones of onscreen, offscreen, and nondiegetic sound (Chion 77), as

will become apparent in the following chapters' analysis of acousmatic music and the voice in the selected films.

2.2.2 Music

As one major auditory device in filmmaking, music largely contributes to the functioning of the audiovisual system and greatly affects the human psyche while watching a film. It is well established that music accompanying the visual events on screen provides the emotional aspect of the sound design and emphasises the mood conveyed through the images. In this section, I will investigate acousmatic music's properties and illustrate how they are used as a tool to influence perceptual processes. Composer Andrew Lewis has established four characteristics of acousmatic music, one of which I would briefly like to elaborate on as an important basis for the further discussion. He points out that in the process of creating acousmatic music "there is nothing to see, *and we know it*" (Lewis n.p.), which means that the invisibility of a music's source is created intentionally. That is an important difference to music which we happen to hear in an acousmatic way, for example a street musician around the corner who we cannot see *yet*. A typical example of this kind would be the scene from *Stranger Than Fiction*, in which Ana is listening to Harold singing and playing the guitar in her living room while she is washing the dishes in the kitchen [01:06:14-01:06:24]. In terms of Chion's interpretation of acousmatic sound, I would argue that the music Harold produces is a sound offscreen from Ana's point of view since it is outside her visible field, at first, but included in the film frame. Beyond that, she is aware that Harold is the one singing and playing the guitar, although she cannot see him the moment he starts playing, which is rather coincidental. I observe similar scenarios in scenes of *Lost in Translation*, for example when Charlotte and Bob come out of the lift [01:02:39-01:02:46] and hear a woman sing, not knowing who it is until they recognise it is Kelly a few moments later. From their point of view, the chant is outside their visual field which is why I consider it as offscreen. Furthermore, they assume they are going to see the music's source soon as they at least know the direction it is coming from. Accordingly, in both examples the musical sources become visible relatively soon for the characters, which implies that the period of time in which they are not visible were not intentionally created as acousmatic sound sources. It is important to state that I will put more emphasis on sound that is acousmatic on purpose.

The way sound is experienced through musical listening is an important aspect of the identification of sound for its own sake. In his *Treatise on Musical Objects* (2017), Schaeffer provides a description of ‘pure listening’ (66). This type of listening is central to his conception and theory of acousmatic music, since a sound’s source is not only not visible, but further neglected for the purpose of solely focusing on the music’s characteristics. For Schaeffer, what he calls “concrete music” is “made of raw sounds: thunderstorms, steam-engines, waterfalls [...]” (Hodgkinson 2), which “are not produced by traditional acoustic musical instruments” (ibid.). Rather, it is a form of music captured on tape, such as the environmental sounds mentioned, which is then “manipulated to form sound-structures” (ibid.). Schaeffer has used these natural and non-musical real world-sounds as musical material in his experiments of sound composition and has thereby raised their musical potential, which is reflected in the concept of the acousmatic. According to him, this potential of the sounds recorded can only be revealed when listening is reduced to the sound’s own properties, independent of its possible meaning. However, I claim that in general it is nearly impossible not to search for the sources of acousmatic music, but that it is human nature to detect these, as it is represented in the short story “The Burrow” (see chapter 2.1). Because of this I agree with Simon Atkinson, who is concerned with the development of the acousmatic history as well as the role of acousmatic sounds and music as signifiers, which is how they can “stand for something” (114). In his article, Atkinson points out the logic of searching for a meaning and the origin of sound and claims that sounds “function as a sign in semiotic terms” (114, 116). I consent to this semiotic approach in terms of the science of signs as applied to music, since it proposes that a particular kind of music can represent a certain action, experience, or object that it refers to. For example, I regard the scene from *Iron Man*, in which Pepper is hacking into the mainframe in Tony’s office, as such a semiotic understanding of music [01:30:50-01:31:57]. The background music influences the viewer’s interpretation of the scene without us being aware of this. The instrumental music functions as a complementary element to the images on screen to emphasise Pepper’s operation. The instrumentality is highlighting the visual even more since there is no song text or familiar melody the viewer could concentrate on or be distracted by. The music is not remarkably loud, but nevertheless intense, why it creates suspense and a certain anxiety which we can detect in Pepper’s behaviour. This scene reveals how such feelings can be infectious in terms of the audience, as I argue that the music is in a key position when creating a filmic experience that is to affect the viewers in a specific way.

In the next chapter, I will examine how such acousmatic music will have lasting influence on the viewers' perception of such scenes. Considering musical meaning, Atkinson claims that "musical material could move from the iconic [...] to the symbolic if its use became conventional, or a cultural 'code'" (119). This assumption seems reasonable to me since there are affirming examples. For instance, Christmas songs played in several public buildings or in the radio at Christmas time become symbolic of the season. Another example would be a national anthem performed at a major sporting event, which can be viewed as a symbol of patriotism and pride. Both paradigms have a cultural meaning and are based on a shared understanding of a certain event or season. Thus, music conveys the associated mood or feelings of the occasions it represents. What is crucial here is that the same applies equally to acousmatic music, provided that the music is generally acknowledged. A striking example of this is a scene from *Stranger Than Fiction*, in which Harold has a tantrum because he hears the acousmatic voice in his head but does not know what is going on or where it is coming from. Out of rage he destroys items in his apartment and is obviously desperate and sad, which is emphasised by the emotional symphonic background music in this scene [00:22:10-00:23:01]. The effect this diegetic music can have on the viewer is enormous, since it activates a multi-layered process in our system of cognition. The underlying musical excerpt is from the reissued piece "On the Nature of Daylight", produced by British contemporary classical composer Max Richter. In an interview on his album "The Blue Notebook" (2004), where the theme is from, Richter shares what he liked to express with the entire work and what the impetus was. The sentimental album was written against the backdrop of the beginning of the Iraq War, which the composer names "the beginning of a sort of politics of unreality where facts didn't really matter anymore" (Deutsche Grammophon - DG 00:46-00:53), as a protest against violence. I detect an atmosphere that is created in a similar vein in the scene from *Stranger Than Fiction* in terms of its spirit. Especially regarding the tune of "On the Nature of Daylight", it is striking how the idea of this piece of music is reflected in Harold's behaviour. Richter explains that he liked to "create something which had a sense of luminosity and brightness but made from the darkest possible materials", a kind of "transmuting" process from "metal into gold" (03:42-04:00). Not only do I admire the idea, but I can also see it in Harold's situation in the sense of a prognostic sign. The associated mood and the feelings conveyed through this musical piece in cultural terms becomes clear by using a kind of melancholic sound in minor keys, in this case it is written in the key of B minor, to emphasise Harold's emotional state. As

conventionally used in film, there exists a shared understanding of considering slow and sentimental music played by a string section in minor keys as usually reflecting negative feelings of a character. As I have indicated, I argue that there is more to the choice of the musical piece than manipulating viewers in terms of their empathy for characters. In Harold's case, the music hints at a transformation that will unfold throughout the film. He is in a state of despair in this particular scene, but the 'luminosity' Richter is talking about, is resonating as well. Harold is calming down towards the end of the scene, which corresponds to increasing the pitch in the excerpt we can hear of "On the Nature of Daylight". With ending on a slightly higher tone and the main character trying to come to terms with what is happening, the music lends emotional resonance to the scene but helps to create a hopeful forecast at once. On the one hand, this scene is exemplary in terms of its collective appeal to emotional empathy, on the other hand, this is also what is problematic about it as there is neither another character involved, nor is the audience's diverse approach to this kind of music considered. Acousmatic music as used in this scene is manipulative for a specific purpose, but I further argue that the audience's individual reactions are a neglected area in the field of film sound analysis, since a certain kind of music means something else for different viewers. I will elaborate on the listener's role, its individuality, and different mental processes at work more precisely in chapter 3.2.

The main limitation of a semiotic approach to music, whose central idea is the music representing a particular experience, action, or object, is in terms of its scope. As discussed above, the examples mentioned are true for collective listening situations, but they neglect the "subjective interpretative experience" (Atkinson 117), even on a cultural level. That is why it is important to have a closer look at the individual perception of music in an acousmatic listening situation in this passage. A prime example of this is another scene from *Stranger Than Fiction*, in which Harold is playing the guitar in the living room, while Ana is washing the dishes in the kitchen [01:06:14-01:06:24]. I have argued prior that the music he is producing is offscreen sound that Ana is hearing in an acousmatic way. From the perspective of perception, I claim that the music in this scene precisely reflects the problem I have highlighted: perceptual processes are affected and manipulated by acousmatic sound in an individual way. This means, the musical meaning is not only reliant on the subjective experience of listening, but also dependent on the characters or viewers' kind of relationship to this music. At this point, I would like to introduce theorists Eric F. Clarke, who is concerned with the implications of

music in terms of the listener's perception and experience. Clarke adopts a "view of perception and cognition [...] that places the issue of musical meaning at the center of the whole endeavor" (*What's Going On* 334). The 'whole endeavor' can be equated with what Ruth Finnegan calls the "complexity of musical experience – mind, body, movement, environment, sensation" (357). It is the effects of music on the listener's experience at all of these levels which I investigate when exploring individual processing of acousmatic sound. Clarke applies an ecological perceptual theory of music, foregrounding the act of perception in the context of musical meaning. His approach contributes to my analysis in that it conflates listening and meaning. He argues that listening to a sound means understanding the perceptual meaning of it, and that understanding causes a reaction (Clarke, *Ways of Listening* 7). He further explains that "[p]erception is essentially exploratory" (Clarke, *Ways of Listening* 19), building a bridge to actions, which are "the result of, and response to, perception" (ibid.). With this in mind, the example introduced above can shed new light on the process of individual manipulation through acousmatic sound, in this case music. The exemplary scene from *Stranger Than Fiction* illustrates Ana's musical experience, since she is the one who is listening and receiving the emotions arising within her. I think that Ana's sense of listening is composed mainly of two different factors here. Firstly, Ana seems to know the original song "Whole Wide World" by Wreckless Eric very well, as she is moving her lips in synchrony when she is watching Harold sing [01:07:04-01:07:11]. That is why I would conclude that she likes the song and, or that it has a specific meaning for her. I would also argue that the song text is appealing to her, as the lines "I'd go the whole wide world, I'd go the whole wide world just to find her" reflect their or more precisely Harold's situation and feelings towards her. Secondly, it is the personal relationship to Harold that influences the way she is listening in this scene. She likes Harold dearly and has a special connection with him, which is why she is particularly correspondent to the feelings that take her by surprise as it seems. I contend she would not react in the same way when a stranger or someone who means less to her would sing an acoustic version of the song. What is remarkable as well in this scene is the complexity of the listening experience, whose importance Finnegan is alluding to in her article. Looking closely at Ana's body movement, it is striking how her motions coincide not only with what she is hearing, but with how she is processing the acousmatic music at the same time. For example, does she pause instantly when she is hearing the first tone [01:06:14-01:06:21]. She closes the water tap and remains in the same position for a moment, as if she is assuring

herself that she is hearing Harold sing indeed. Her sensations are a further part of the musical experience: emotions come up and she seems to be deeply touched by his singing since she stops washing the dishes and walks halfway through the doorway to stop again and watch him making music [01:06:23-01:06:33]. I argue that this movement is already part of what Clarke considered the result of and response to perception: she shows a reaction, namely listening carefully and appreciating and enjoying the intimate moment. They both seem to be lost in that musical world for a short time, climaxed by the genuine kiss Ana gives Harold at the end of the scene due to her overwhelming emotions [01:07:35-01:08:03]. I elaborated on this example in detail to emphasise the individuality of the process of perception. Ana's listening situation highlights just how important it is to explore the subjective understanding of music corresponding to a character's or viewer's environment. I argue that her mode of listening is influenced to the effect that she is not expecting Harold to sing when she is in the kitchen, which initiates her reaction triggered by her emotions. This example indicates how individual processing of acousmatic music can look like, since we can observe Ana's behaviour throughout the entire scene. This illustration helps to comprehend and relate to her perception, still it is important to emphasise that this is only one subjective interpretation of the musical piece. The first question to keep in mind is about the diverse notions of the viewers, and the second question I would pose is whether the scene would have happened like this without the acousmatic music, which leads back to the manipulation of the acousmatic music's perception. I hope I made clear that the concept of acousmatic music is as important in a coincidentally acousmatic situation as it is in an intentionally acousmatic condition, regardless of whether the sound source is deacousmatized or not.

There are multiple kinds of adaptations in social and cultural mechanisms, leading to a change in conception over time. Clarke is interested in examining perception from the viewpoint of adaptation, suggesting that all individuals are "immersed in a continual process of perceptual learning" (*Ways of Listening* 22), which helps them becoming more familiar and confident in perceiving their environment. Luke Windsor, also coming from the background of ecological psychology of music, is important to add to the discussion in terms of acousmatic music. In his article, he emphasises the difference between the perception of real events and that of an "acousmatic piece" (Windsor 20), which is significant in regard to the listeners' ability to perceive sound sources (Windsor 19). Windsor argues that this capacity marks "the gap between 'everyday' listening and 'musical' listening" (19), since with acousmatic music,

listeners try to detect the sound's cause. I agree with Windsor, since I have previously shown that this is the case in my objects of analysis. In all exemplary scenes analysed so far, the character involved is either actively searching for the sound's source or seems to be automatically attracted to the cause. This is what he calls "an important facet of musical interpretation" (Windsor 9), emphasising that the search for the sound source is personal (Windsor 31). This assumption seems fully plausible to me, as I have illustrated in my analysis of different scenes including acousmatic music. He further illustrates that the acousmatic listening situation creates a special communication between a music's composer and the listener. The composer can play with and manipulate musical sources, which entails that the listener might not hear the actual cause but something else (ibid.). "The notion of real and virtual causation" (ibid.) implies that "neither composers nor listeners are fully in control of what will be perceived" (ibid.). This is of particular importance for the analysis of the viewer's perception since the characters' interpretation of an acousmatic situation is obviously part of the film script and thus controllable. In contrast, the audience's reaction is diverse and not entirely predictable. According to Windsor, this insecurity is represented in an acousmatic experience, in which the listening situation is often dubious (ibid.). Subsequently, I will undertake a detailed analysis of the listening perspective in terms of perception in chapter 3. Considered within the social context of listening, the way we understand acousmatic music can be actively manipulated and arranged. The emotional response to such scenes and consequences I will examine in a cognitively oriented approach in the next chapter.

2.3 The Voice

In the domain of film sound, the voice takes on an exceptional role. It reveals new aspects in terms of how humans perceive and process auditory information and proves to be the most remarkable type of sound affecting the human cognitive processes. Among other cinematic elements of 'audio-vision', which I have introduced in the previous subchapters, Michel Chion profoundly discusses the voice, which has become the focus of his approach of film sound. In one of the first major studies of the human voice in film, his work *The Voice in Cinema* (1999), first published in 1982, he solely devotes himself to the function of the embodied and disembodied voice. For my approach, the latter plays a crucial role. As Chion claims in the prologue of his work, the voice is a sonic component often confused with speech, why the

voice itself is often neglected (*The Voice in Cinema* 1). In the context of cinematic sound and the audience's perception, he has coined the terms 'vococentrism' and 'verbocentrism'. By 'vococentric' he refers to the voice as the privileged sonic element over all others, meaning that there is a hierarchy in the sonic space. He centralises the voice by reducing other sounds, arguing that "[h]uman listening is naturally vococentrist" (*The Voice in Cinema* 6). First of all, we – the audience – focus on the character's voice as it centers our attention, although Chion also asserts that "identifying a voice is an elusive and difficult proposition" (*The Voice in Cinema* 33). Because an acousmatic voice is difficult to assign to a certain person as long as it does not have distinct characteristics, it is often used for reasons of deception and manipulation in film (ibid.). I will use the term 'acousmatic', which Chion has already introduced in *Audio-Vision* (2019), interchangeably with 'disembodied', which basically describes a voice that is heard without a physical body seen. Chion further argues that cinematic sound is also a 'verbocentric' phenomenon, which focuses on "the voice as the medium of verbal expression" (6). Thus, speech, language and the meaning of words are at the heart of 'verbocentrism' (ibid.). It is the interplay of both 'vococentrism' and 'verbocentrism' that constitutes my object of investigation and is important for discussing disembodied voices in the selected films. Against the backdrop of acousmatic sound, the voices I am particularly interested in in my research objects are the following: in *Phone Booth*, it is clearly the unknown caller's voice that is at the center of attention. The voice runs like a red thread through all the developmental stages of the film's plot and has an immense power to disrupt the protagonist's whole life. In *Iron Man*, especially the artificial intelligence system invented by Tony Stark – J.A.R.V.I.S. – is interesting for my interpretation, since it is the only unnatural voice that contributes to my analysis. Yet a greatly different concept of the voice is presented in *Stranger Than Fiction*, in which the disembodied sound, whose identity is revealed step by step, is introduced as a mysterious female voice. In the beginning the voice seems to be the narrator of the story, until later it becomes visible at first for the audience and later to the main character Harold. In my fourth object of analysis, *Lost in Translation*, the acousmatic voice is mainly used in different situations of communication as a means to depict the protagonists' lives and their emotional states during their journey. I would say that disembodied voices in this film characterise distancing as well as connection, as both Bob and Charlotte are going through a process of inner estrangement from their private lives, represented by their partners, which brings them closer together.

There is a special kind of acousmatic voice which must be subject to the discussion, as it holds special powers that are significant for the analysis of perceptual manipulation. By creating the voice as an object of study in *The Voice in Cinema* (1999), Chion rearranges the term 'acousmatic' as he develops the phenomenon of the 'acousmêtre' (*The Voice in Cinema* 17). His basic definition of this term reads as follows: "[w]hen the acousmatic presence is a voice, and especially when this voice has not yet been visualized [...] we attach the name acousmêtre" (Chion, *The Voice in Cinema* 21). Thus, it can be assumed that an invisible character speaking is created through its voice, rather than through its presence on the screen. This is the case in *Phone Booth*, in which the unknown caller intimidating Stu in the phone booth meets all the criteria for being an acousmêtre. Throughout the entire film, his identity remains mysterious, up until the last scene. This kind of 'acousmêtre' is called the "complete acousmêtre" (ibid.), which differs from the one that is already visualised. Chion further defines the 'acousmêtre' as "[n]either inside nor outside" (*The Voice in Cinema* 23), not inside because the embodiment of the person speaking is not included. Nor is it outside, as the voice is not clearly heard offscreen, which I have explained in chapter 2.2.1 (Chion 125). His notion of this phenomenon shall not be confused with a "radio-acousmêtre" (Chion, *The Voice in Cinema* 21), which is "acousmatic by nature" (ibid.). The crucial difference is that it is universally known there is no way of seeing people speaking on the radio, whereas in film, it can be played around with showing the physical embodiment of a voice (ibid.). This becomes especially obvious in the presentation of Karen Eiffel, who is introduced and visualised in an unconventional way first to the audience and then to the main character Harold, which I will return to later. According to Chion, there are four distinctive powers attributed to the 'acousmêtre', namely "ubiquity, panopticism, omniscience, and omnipotence" (*The Voice in Cinema* 24). These ensure that "[t]he acousmêtre is everywhere" (ibid.) and it can see everything that is onscreen, which is why "the acousmêtre's character seems to inhabit the image" (Chion 129). Accordingly, 'omniscience' derives from the power of the panoptic view: because the acousmêtre is all-seeing, it is able to know everything, even a character's thoughts. The omnipotent property, which is usually ascribed to God, provides the acousmêtre with unlimited powers to have complete control over a situation. As I have indicated, I would classify the unknown caller in *Phone Booth* as an acousmêtre according to Chion's concept. He fulfils all the criteria stated above: he is ubiquitous, and he sees everything not only Stu is doing inside the booth, but what is happening around him as well. There are a

few camera settings that illustrate this, as for example in this frame, where the viewer seems to look through a telescopic sight from the sniper's point of view:



Figure 2. View from the unknown caller's perspective [0:49:07].

Looking through his gun from the caller's point of view demonstrates the power he inhabits, as it becomes clear that he is able to control the proceedings at all times. Beyond the panoptic view he is also omniscient, which means he knows everything about Stu. He knows about Stu using immoral means at work, but even more important, he knows every detail of his private life, such as Stu's affair with Pam. When the caller speaks to Pam and reveals to her that he is married, Stu and we, the audience, begin to realise the gravity of the situation [0:15:15-0:17:14]. In the course of the call, the acousmètre is revealing more and more information he gathered about Stu via a microphone he installed in the phone booth and a telescopic sight. The caller's omnipotent power coupled with possessing a rifle put him in a superior position and make him occupy all the essential qualities of the acousmètre as defined by Chion. It is often these powers coming from a "non-localized body" (Chion, *The Voice in Cinema* 24), that make this phenomenon unsettling or uncanny. Chion calls it "obsessional *panoptic fantasy*, which is the fantasy of total mastery of space by vision" (ibid.). This is entirely true for the main character in *Phone Booth*, since Stu, who is presented as a self-confident, arrogant, and successful businessman at the beginning, is evolving into a scared and intimidated person throughout the film. At a certain point, he is terrified of the acousmatic voice, who becomes more determined to punish Stu for his offences, because he realises, he is under total control and not able to escape without something hurtful happening [0:22:30-0:23:08]. By saying "I'm aiming at you right now" [0:23:34-0:23:36], the caller is challenging Stu to test his nervous limits. The utterance is followed by an alternation of different close-ups of Stu's terrified face

and the different surrounding skyscrapers which we see blurred from his perspective [0:23:36-0:24:28].



Figure 3. Close-up of Stu [0:23:36].



Figure 4. Close-up of Stu [0:24:29].

Stu is desperately searching for the physical source of the voice but has to resign as he sees how many thousands of windows there are, and that there is no chance to find out where the person is hiding. Eventually, the caller drives Stu into confessing in front of everybody that he cheated on his wife and that he lied. In this scene, everyone, all the other characters and the viewers, are witnessing Stu's emotional outburst, which marks the highlight not only of the entire film, but especially of the effect that an acousmètre can have on another person [01:05:38-01:08:48]. By means of Stu's character, we can observe a personal development with an outcome that is thought-provoking. Aside from the fact that the intense process urges the viewer to reflect and can cause different emotional responses, which I will further analyse in the next chapter, it is important to highlight the time span of listening in an acousmatic way. I claim that the duration of time an acousmatic voice is heard is important and decisive in terms of its functions, which I would consider a neglected area in the field of film studies so far. In *Phone Booth*, the caller's voice remains acousmatic almost until the end and has a huge impact on the characters. An example of a voice that functions according to the same principles but is acousmatic for much less time, often has a completely different effect on characters or audience. The character of Karen Eiffel in *Stranger Than Fiction* for example has a distinct effect, since it is visualised much earlier than the unknown caller in *Phone Booth*. Yet, it possesses all the powers of the acousmètre according to Chion at the beginning, as becomes obvious the moment the voice starts reporting on Harold's life [0:04:05-0:06:10]. The female voice is coming from a "non-localized body" (Chion, *The Voice in Cinema* 24), while being ubiquitous for the audience and Harold, who can clearly hear the voice. Since the acousmatic speaker is describing numerous details of Harold's everyday life and his actions as

well, it seems to be omniscient and to have a panoptical view as well. These are the qualities why I do not categorise other acousmatic voices from my research objects, such as J.A.R.V.I.S. in *Iron Man* or the family members in *Lost in Translation*, as acousmêtres, since they do not fulfil all of the powers as defined by Chion.

An acousmêtre can be dissolved, meaning that a voice does not have to be continuously acousmatic, but that there is a process in which the function and effect of the disembodied voice changes through visualisation. The fundamental quality, that immediately deprives the acousmêtre of its powers, is “*de-acousmatization*” (Chion, *The Voice in Cinema* 23). When the source of the voice is revealed and the character speaking embodied, the person is “*deacousmatized*” (Chion 128). From this moment on, the voice loses its “magical powers” (Chion, *The Voice in Cinema* 23), which is often a dramatic process of unveiling an unknown character and transforming it to a visible human being by assigning a voice. A prime example of this is the last scene of *Phone Booth*, in which Stu is in the ambulance and a man is passing by who says “nice shoes” at first [1:13:16]. It is interesting how the acousmêtre is partly ‘deacousmatized’ simultaneously to the audience and the main character here. But according to Chion, the revelation does not necessarily happen all at once; it is often a process in which voices are only partly revealed. When there is “an end point of deacousmatization” (Chion, *The Voice in Cinema* 28), as for instance when we have not yet seen the speaker’s face, Chion calls it ‘semi- acousmêtre’. When we have seen parts of the speaking character, like different body parts, but not the mouth, he defines it as ‘partial de- acousmatization’ (ibid.). I would argue that in the example stated above, there is a flowing transition from a partial disclosure to a complete ‘de-acousmatization’, since the camera is moving from the person’s hands grabbing Stu’s shoes to his face, which is briefly shown [1:13:15-1:13:18]. Chion claims that as long as the character’s face and mouth have not been completely unveiled, the “de-acousmatization is incomplete” (*The Voice in Cinema* 28). Intriguingly, we find a special case of ‘de-acousmatization’ in this example. For, although the acousmêtre is embodied as it is connected to a face and mouth, it is worth noting that both the character and the viewer have a limited field of vision due to the blur [1:13:15-1:14:04]. The setting of the camera is never really sharp in this scene with the purpose that the person remains mysterious despite being ‘deacousmatized’. Consequently, I would argue that for Stu it is not a complete revelation of the voice’s embodiment because he is still in a trance and not able to react to what he is experiencing. Furthermore, for the audience it is similar because we cannot see clearly,

although there are some close-up shots of the person's face. It is only in the next shot when the person is walking away from the ambulance that the camera image becomes sharp again and we are able to see the man from above [1:14:07-1:14:23]. Paradoxically surrounded by policemen who do not recognise his identity, the character stops at the booth and the camera zooms onto his grinning face [1:14:24-1:14:33]. This time we are provided with a sharp image of the acousmêtre, which I would consider, is hereby completely deacousmatized. The voice is finally embodied.

A different kind of 'de-acousmatization' have I found in *Stranger Than Fiction* in the shape of Karen Eiffel's voice. What is distinctive about her voice, which is acousmatic in the beginning, is the way it is deacousmatized. At first, the viewer as well as Harold can only hear the voice in an acousmatic way for the first thirteen minutes of the film. But then, the acousmêtre is partially deacousmatized to the viewers by suddenly showing a woman, who later turns out to be Karen Eiffel, from the back. The camera zooms in on the character, still from a back view, to then move toward the subject and pan to a bird's eye shot, from directly above her head [0:13:23-0:13:33]. From this camera position, the viewer looks in the same direction as the character, which shapes the viewers perspective of the scene. Here, Karen Eiffel who is standing on a rooftop surrounded by the skyline of Chicago, is looking down on the busy streets of the city. I argue that all the powers of the acousmêtre are perfectly implemented in one shot here: the woman seems to watch over the city from above, apparently seeing anything that is happening below, and obviously knowing every detail of Harold's life. The camera's movement and the bird's eye view from directly overhead emphasise that she is in a superior position, which further demonstrate how things in this, and the previous scene relate to each other. Right before Karen Eiffel is shown from the back, the other scene ends with Harold being furious about the voice he is hearing and shouting upwards to the heavens [0:13:17-0:13:22]. The camera zooms out and moves further away from Harold, which provides narrative information concerning the next shot. In this case, the camera's motion to move to the next shot reveals a new subject that was offscreen before. I claim that this technique is significant in terms of the viewers hypothesis that the person shown is the acousmatic narrator. The 'partial de-acousmatization' thus effectively creates tension in terms of the character's identity. Only in the next shot, in which there is a close-up of the woman's face, the acousmêtre becomes fully embodied and is completely deacousmatized [0:13:33]. From this point on for the audience, Karen Eiffel is not an unknown

acousmatic voice anymore, but a vivid person who is part of the story world. This does not apply to Harold since he discovers only much later who the acousmatic voice belongs to. This highlights the distinctiveness of the de-acousmatization process in *Stranger Than Fiction*, as the acousmatic voice is revealed at various times to the viewers and the main characters. Harold finally detects the voice when he is at Professor Hilbert's office: Karen Eiffel is being interviewed on television and introducing her new book called "Death and Taxes" [1:10:13-1:10:16]. Harold cannot believe his ears as he recognises her voice as the one in his head, narrating his story. He walks to the television and listens for a few seconds, when he says, "Oh my God, that's her" [1:10:45]. The deferred 'de-acousmatization' leads to different point of views among the characters in the story world and the audience outside of the diegetic space. This indicates that an acousmatic voice plays a different role concerning the acousmatic zone it can be classified in, and that interpretations differ according to a time-shifted deprivation of the acousmètre's powers.

The technique of voice-over is another way of deploying acousmatic sound in film, often referred to as an off-stage narration among laymen. Mostly for informative reasons, voice-overs are commonly known as commentaries that only the audience, but not the characters in a film can hear. Film theorist Mary Ann Doane, whose field of research mainly includes sound and the organisation of space and time in cinema, is one of the well-known scholars who addresses the role of voice-over in cinema. First of all, she defines the conception of voice-over as being "necessarily presented [...] outside of [the diegetic] space" (168). A characteristic example of this would be the introducing voice-over in *Phone Booth*, employed to provide the audience with a quick exposition of facts about the New York City telephone network [0:02:30-0:03:43]. Doane characterises this voice as radical different "with respect to the diegesis which endows this voice with a certain authority" (ibid.) over the characters in the film. Since this voice speaks directly to the audience and neglects the characters, it "establish[es] a complicity between itself and the spectator" (ibid.), who are then able to understand and interpret the image together. This is particularly true for Karen Eiffel's voice in *Stranger Than Fiction*, who is reporting on Harold's life, not talking to him but exclusively being turned towards the audience. Because the voice-over is non-determinable, it owns the power of "possession of knowledge and [...] the privileged, unquestioned activity of interpretation" (ibid.), which helps to provide important information according to the image on screen. This is the case for both voice-overs stated above, since each gives detailed

information in form of facts, that can only be knowledge-based. There are different kinds and a great number of voice-overs, as in the interior monologue or the voice-over commentary in a documentary. But what is characteristic of the voice-over most conventionally used in film, is that it is mostly detached from a particular figure, meaning that it is a disembodied voice (Doane 169). In the discussion of the voice-over phenomenon, film scholar Sarah Kozloff also tackles this topic in her work *Invisible Storytellers: Voice-Over Narration in American Fiction Film* (1988). Interested in communication through a narrative, she breaks down the definition of the term to make clear its use in fiction film. Kozloff states that the voice “determines the medium: we must hear someone speaking” (2); the word ‘over’ refers to the audience’s inability to see the person who is speaking. She emphasises that ‘over’ “implies more than mere screen-absence” (Kozloff 3), since otherwise it might be confused with the term ‘voice-off’, which indicates that a speaker is merely “temporarily off-camera” (ibid.). In contrast, the voice-over belongs to the spatial-temporal dimension of the discourse (ibid.). As stated above, Kozloff’s focus is on the narrative, arguing that ‘narration’ “relates to the content of the speech” (ibid.), which is telling an audience a “series of events” (ibid.). According to the criteria offered by Doane and Kozloff, I consider both of my examples from *Phone Booth* and *Stranger Than Fiction* as actual voice-overs. Furthermore, I agree with Kozloff’s classification of different kinds of voice-over narrators. She differentiates between a ‘third-person’ or ‘authorial’ narrator and a ‘character’ or ‘first-person’ narrator (Kozloff 6), based on Gerard Genette’s *Narrative Discourse* (1980). For my further analysis, it is important to be able to rely on such a precise classification of these types of voice-over narrators. Although the voice-over’s purpose in film is often to provide the audience with important information regarding the content, and to “implicat[e] us in his world and his value system” (Kozloff 50), as she explains, my main interest diverges from this content-based approach. I would like to take a step further and rather examine the voice-over’s attitude and especially its impact on the characters and effect on the audience in terms of perception.

3 Strategies of Cognitive Processing

This chapter sheds light on the influence of acousmatic sound on human perception according to the cognitive aspects of film viewing. Central concepts of the physical qualities of sound and the analysis of acousmatic sound in film as laid out in the previous chapters serve as a

basis for investigating further how acousmatic sound is processed. The influence of acousmatic sound on viewers has rarely been explored, particularly not in cognitive terms, which is why I aim to investigate the audience's perception and interpretation of acousmatic sound in film scenes and its manipulation conveyed through fiction in this chapter. For this purpose, I will discuss different strategies of cognitive processing by highlighting the manipulative aspects of acousmatic sound that arise from the analysis. To be clear on the further use of the concepts of perception and cognition, I will clarify how I understand both terms. Psychologist Ulric Neisser focuses on the balance between approaches to perception and to cognition, while clearly marking the transition himself. He posits that "perception is where cognition and reality meet" (Neisser 9), since "perceiving is the basic cognitive activity out of which all others must emerge" (ibid.). I rely on Neisser's differentiation as a basis for this chapter, as I will refer back to these terms regularly in the following.

Based on my analysis, I will then illustrate the cognitive repercussions that are revealed during this process. To support my main claim in this paper, I will incorporate mainly my own insights, but consider different views of my interlocutors. I conducted a mini study with two small focus groups, one of them consisting of two women, and the other consisting of two men. This division provides for a more complete and diverse understanding of the audience's interpretation of acousmatic film sound. In terms of demographic characteristics, all respondents share the following characteristics: they are aged between 25 and 30, German native speakers with good English language skills, and they hold a bachelor's degree as regards their educational background. I chose these four people as suitable participants because on the one hand they share these characteristics, on the other hand I personally know them and about their passion for film. I asked them to watch the four films selected with special focus on specific scenes, but I did not provide them with information concerning my research questions, to allow for an unbiased debate on the issue. Their perspectives made important contributions to my understanding of cognitive processes and helped me to clarify my findings in terms of perceptual manipulation.⁶

Concerning mental activities, I will base my analysis on the following theorists, who are engaged in the investigation of how cognitive processes work in various ways and examine how acousmatic sound influences the audience's perception differently. First, I will draw on

⁶ Please find the entire reviews in the Appendix A.

different mental processing models and highlight how the information-processing model by Eric F. Clarke and the ecological approach based on James Gibson's perceptual theory both contribute to the audience's mental processing of acousmatic sound (see chapter 3.1). In section 3.2, I will examine in more detail the viewer's activity as an active process concerning individual listening styles and conditions. Accordingly, I will rely on Debra L. Worthington and Margaret E. Fitch-Hauser, both interested in listening research. They explore the role and process of listening in communication as a social function, as a critical professional competency, and as a cognitive process, the latter of which I am interested in. Film viewing as a dynamic psychological process is also what is significant in chapter 3.3. Here I will deal with Gregory Currie, who is coming from a background of Philosophy, and his work on imagination and the simulation process and analyse in which ways fiction appeals to the imagination as a mental mechanism. Finally, I discuss the audience's emotional processes and empathy while watching a film from a cognitive perspective in chapter 3.4. Here, I will mainly base my argument on film theorists Ed Tan's and Nico Frijda's interpretation of sentiments in film viewing.

3.1 Mental Processing Models

The relationship between the physical properties of sounds and their psychological effects largely depends on the mutuality of listeners and their complex environments. But how do humans interact with everyday acoustic occurrences and interpret auditory phenomena in cognitive terms? To determine the audience's way of processing, I rely on a combination of two different accounts of sound perception. Both are concerned with how people interpret the acoustic events happening around them, but in different manners. The basic and dominant approach in auditory perception that is significant for my intent to investigate how listeners process sound is that of "information-processing" (Clarke, *Ways of Listening* 11). The difference between this view and the ecological approach to perception, which I will explain subsequently, is in the structure of the environment. Psychologist William James considered the infant's environment a "blooming buzzing confusion" (16) early on, illustrating that the entirety of the world is complex and puzzling and thus set to be inscribed upon. His central statement became a pivotal model of experience and is adopted by Clarke, who assumes that "structure is not in the environment [but] imposed on an unordered or highly complex world

by perceivers” (*Ways of Listening* 12). What follows is a “conception of [sound] perception as a set of stages or levels” (ibid.), which includes a schematic portrayal of information-processing from basic, more physical properties, via mental through to more complex and social characteristics. This means, the basic structure of the information-processing model consists of a “route from simple to complex properties” (Clarke, *Ways of Listening* 14), as can be viewed in figure 5. The simple features of a sound serve as a starting point of perception and lead to more multi-layered mental processes in the form of “bottom-up processing” (Clarke, *Ways of Listening* 15). I will give a potential example of a viewer’s perception according to this schema in the following paragraphs.

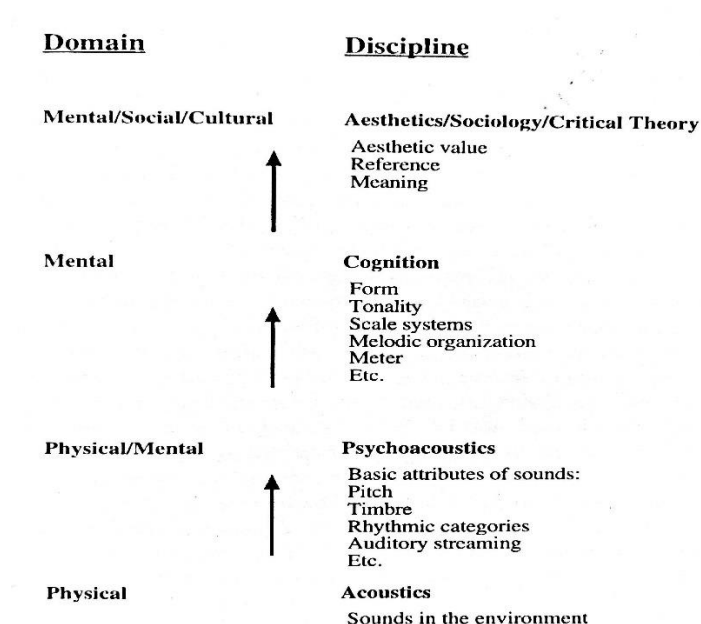


Figure 5. Schemata of information-processing (Clarke, *Ways of Listening* 13).

A less constructive process is found in the ecological approach based on James Gibson’s perceptual theory. It presupposes that the environment is already structured in itself and that “perception [serves] as the pick-up of that already structured perceptual information” (Clarke, *Ways of Listening* 17). This notion of picking up information relies on “[t]he directness of our perception of the world”, which “is the consequence of adaption, perceptual learning, and the interdependence of perception and action” (Clarke, *Ways of Listening* 47). Thus, ecological theory emphasises the relationship between perceiver and environment as the central element of perception (Clarke, *Impact of Recording* 48). I will continue with the analysis including aspects of both approaches, since they equally provide valuable contributions to the question of how the viewer is manipulated through acousmatic sound. I agree with Ulric

Neisser, who questions whether cognitive processes solely work in terms of the structure of the environment, or that they are exclusively based on radical forms of information processing, which leaves the perceiver relying on himself alone (53). It is a balanced combination of both forms that I will rely on in the following. As Neisser rightly states: “Although perceiving does not change the world, it does change the perceiver. (So does action, of course.)” (ibid.). In a broader sense, this coincides with my main assertion that the implementation of acousmatic sound in fiction serves as a means of manipulation, which can lead to multifaceted mental consequences concerning the viewer’s state of mind.

Acousmatic sound can impact the audience in a variety of different ways, depending on how individual information processes look like. What I detected while watching and analysing the films chosen as objects of analysis myself and through my interlocutors’ responses, is that distinct listeners make sense of sounds according to a number of different factors. These factors, in turn, contribute to the degree of possible perceptual manipulation. The most important ones can be deduced firstly from Clarke and his approach to the perception of a sound’s meaning, and, secondly from my research and observations of a viewer’s listening experience. The processing is thus dependent on the following aspects. Past experiences of viewers (Clarke, *Ways of Listening* 11) play a major role, considering that previous contexts and situations, including memory processes, are highly individual. In terms of past auditory events, we need to distinguish between long-term and short-term memory. Bob Snyder rightly points out that they can differ in the degree of their potential influence, since recent memories are more present and easily accessible than long-term memories, which are often not more than a trace of an auditory event in the past (215). Furthermore, he claims that especially long-term memory is not a fixed matter, but lively since “memories constitute who we are, and who we are is always changing” (Snyder 71). This well-founded assumption is crucial to bear in mind in regard to how past auditory events can manipulate the perception of current acousmatic sound. Furthermore, the listeners’ expectation (Clarke, *Ways of Listening* 14, Schnell 4) is a significant factor since it also influences the way in which the audience perceives sound in acousmatic situations. Based on diverse personality traits (Worthington/Fitch-Hauser 69), I would add that an individual’s mindset, including aspects such as imagination and intention – meaning the level of engagement with a film while watching – are decisive elements in the matter of cognitive processing and its manipulation. I will further elaborate on the imagination process in chapter 3.3.

Interpreting the listeners' acousmatic sound experience requires an examination of different personal factors that contribute to the meaning that the audience ascribes to this kind of experience. I will illustrate two of the main criteria stated above using relevant scenes from the selected films and thus give examples of the possible handlings of listening to acousmatic sounds. I will begin with the matter of past experiences and exemplify how they can affect the perception of acousmatic sound.

In *Phone Booth*, I have already discussed the issue of the acousmatic voice respectively the concept of the 'acousmètre'. I have analysed the main character's behaviour, for which research has shown that the acousmatic voice by the unknown caller changes everything for Stu (see chapter 2.3). But what about the audience's perception and their emotions in regard to past experiences and memory? It became clear from my interlocutors' remarks that a particular sound, in this case a voice, does not evoke the same emotions in every person, let alone the feelings of the characters in the film. Whereas the omniscient narrator speaking at the beginning [00:02:31-00:03:36] is generally regarded as neutral, since he is not causing any noticeable emotions, the unknown caller's voice is perceived in different ways. First of all, this stresses that when a nondiegetic narrator gives impersonal and fact-based information and is clearly external to the story world, though the voice is acousmatic, the viewer's emotions are not influenced significantly. It is otherwise with the voice which I identified as 'acousmètre', since personal feelings get involved, be it on the part of the listeners, or the character in the film. The correlation between our previous experiences and the way we perceive present happenings is remarkable because "[p]ast experiences are associated with particular feelings and when similar present experiences are encountered, humans tend to rely on past feelings to interpret the phenomena perceived" (Lewis 42). My findings correspond to Lewis's assumption, especially in terms of noise sensitivity. There are a few situations in which the acousmatic voice gets angry, louder, and more forceful, as in the scene in which the caller threatens to shoot the police officer [00:38:32-00:39:03]. It became clear from a conversation with my focus group 1 that person B could hardly bear these scenes because the voice becomes more threatening and louder. Not seeing the person speaking in combination with the raising noise level made her feel extremely uncomfortable. She revealed to me that when she was little, her dad used to raise his voice against her mother often, when she was secretly listening to their arguments. This discovery provided an insight into how memory processes work and how they can be triggered, as well as how present experiences are encountered.

Her reaction further proves that our long-term memory in fact plays a major role when processing acousmatic sound, since her childhood experiences still seem to have a huge impact on perceptual processes. Whereas person B reacts rather sensitively to sound and its volume and perceives a loud voice as intimidating and terrifying as the main character does, others respond in a totally different way. When I asked person A about her feelings about the caller raising his voice, she responded that she found the voice highly interesting as it sounded intelligent and calm, nevertheless mysterious, and that she wanted to know more about that hidden character who was behind this voice. Thus, she was attentively listening and examining the voice, but she was never feeling afraid or intimidated by its sound. In the first place, I derived her reaction from the fact that she does not have any negative connotations in terms of noise-sensitivity related to voice in the past. Other than that, I know that she listens to numerous podcasts and thus is used to various acousmatic voices and ways of conversations, and that she is an attentive listener who is interested in voices. These two perspectives alone are illuminating with regard to the cognitive processes at work in an acousmatic listening situation. In this case, the described situations represent contrasting examples of person-related factors that are decisive for the handling of such situations: whereas the first person is mentally challenged and marked by past events and memories to some extent, the second person is in a more impartial and neutral condition, not struggling with past experiences which could influence the way of dealing with the acousmatic voice. The examples prove the complexity of the experience, since the different factors contributing to perception are diverse and highly individual. I would further argue that both processing models introduced at the beginning of this chapter are reflected in these examples. Among others, the ecological approach to perception is characterised by an environment that is already structured and whose pick-up of information relies on adaption and perceptual learning (Clarke, *Ways of Listening* 17, 47). I would categorise person B's interpretation of the voice into this scheme, since she makes sense of the acousmatic voice through perceptual learning, which is experience-dependent and thus determines how the present sound is mentally processed. Based on my explanation of the information-processing schemata (see figure 5), I could classify person A's way of perception as belonging to this model of processing, since it may seem that she starts to perceive the more basic characteristics of a sound first. However, this does not mean that she is not impacted by what she experienced earlier. The problem I see with the pure information-processing approach is that it is too abstract and too stringent for

a specific and generally valid allocation of a viewer's perception. It does not seem plausible that every person, as unbiased and natural they might be, picks up the environmental information the way this model suggests and necessarily starts to perceive, as Clarke suggests, "lower-level features" first (*Ways of Listening* 16). Everyone is somehow formed, prejudiced, or influenced by the surrounding environment, social factors, or even past experiences, which means that processing does not always work bottom-up.

Another crucial aspect that contributes to the mental processes in an acousmatic listening situation and that determines the interpretation of a sound scene is the listeners' expectation. This factor is closely related to and based on past experiences, still worth elaborating on separately. As proposed by Michael Gatt, the listener is not able to build any expectation "[w]ithout a compatible schema to decode the [sound]" (208), which refers back to the processing models I have introduced previously. I agree that if the listener were unable to draw on a scheme that establishes the connection between perception and mental processing, they would be "left with nothing but the present" (*ibid.*). This would lead to the assumption that the audience could not even form some kind of expectation, which I have already invalidated by presenting the processing models and its influential factors. Neisser confirms that the listener "continuously develops more or less specific readiness (anticipations) for what will come next, based on information he has already picked up" (27). He claims that these anticipations are not very specific, stating that "[p]erceptual hypotheses are rarely definite" (Neisser 28). This corresponds to my assumption that acousmatic situations in film can serve as a means of perceptual manipulation, since a surprise effect is created knowing that expectations are principally not concrete. I would argue that different expectations, though not highly specific, are also part of the reason why the consequences which result from acousmatic experiences are so diverse for different auscultators.⁷ The following example from my personal listening experience concerning the expectations in light of the introduced accounts of perception, supports this claim. In the beginning of *Stranger Than Fiction*, Harold is introduced by the acousmatic narrator while exercising his daily

⁷ In the context of narratology, Melba Cuddy-Keane explains "the aural placement of the listener" (71) and coins the term 'auscultation' and its derivatives in relation to the visually dominated terminology of focalisation (for a further definition see Mieke Bal 2009: 145ff.). She states: "Auscultation, the action of listening, would parallel focalization; auscultize [...] would be the verb signifying the presentation of sound as listened to; and auscultator [...] would signify the person doing the listening". (71)

routine. This procedure is accompanied by instrumental music for over two and a half minutes as a complementary element to the images [00:01:27-00:04:03]. It is interesting that the diegetic music itself, which in this case is intentionally created as acousmatic for the audience, shaped my expectation on how this story would unfold.

As I have indicated earlier, the music's specific properties are used by composers as a tool to influence human perceptual processes, even if the audience's reaction is diverse and not entirely predictable in the view of the filmmakers. My expectation, when I first listened to the music accompanying the images on screen, was to experience a positive, optimistic story introduced by a powerful and cheerful main character. This was mostly because the song, which I later identified as the instrumental version of *The Way We Get By* by Spoon (2002), has a pulsing and steady beat, introduced by the piano sound at the beginning. When the percussion joins, the rhythm becomes still more joyful and lively. In technical terms, I used the metronome to define the tempo of the song, which in 4/4-time can be measured as around 147 beats per minute (in short: bpm). In musical terminology, this tempo would correspond to 'allegro', which indicates a generally fast and lively beat. It is applied to pieces of cheerful nature and thus emphasises the mood I experienced. However, I realised that my anticipation induced by the music was in significant contrast to the images and the narrator's description of a character, who seems to be a creature of habit, and a boring and conscientious person who is maybe satisfied, but not truly happy with himself.

I felt validated when I studied the lyrics of the song, which are left out in the scene, but which are also contrasting with Harold's visual characterisation. In my interpretation, *The Way We Get By* is supposed to be a rebellious song about a young generation growing up, doing whatever they enjoy whenever the opportunity arises. First of all, the lyrics seem to portray the complete contrast to Harold's tedious lifestyle, why I felt that my expectation has been proven to be correct. Yet, the chorus suggests that this carefree lifestyle might not be as easy and joyful as it seems, since "that's the way we get by"⁸ is sung several times in a row. It is more likely that the described way of living is a means to escape reality, implying that their actual situation is not as light-hearted and easy-going, but rather uncomfortable or even worrying. This is illuminating in the sense that my expectation would have been different in regard to how the story unfolds if the song had been played with lyrics. This is mainly because

⁸ Please see the appendix for the complete lyrics of the song *The Way We Get By* by Spoon.

the underlying negative connotation would not only have been present, but also would have matched the visible appearance and the narrator's description of the main character. However, the lyrics have intentionally been omitted, probably in order to create this dichotomy of Harold's situation visualised on the one hand, and the accompanying music with a different, more positive vibe.

As this analysis indicates, our expectation determines the way in which acousmatic sound is processed, especially in combination with images. This is clearly only one individual expectation, but it shows how music can manipulate what we perceive through the images. Jim Schnell argues that "[w]e sometimes look for the meaning we expect or want" (4), which I would agree on since I also tried to see the positive atmosphere created by the music, in Harold's everyday life, that in fact looks different on screen. Based on my musical knowledge and previous experience, which set up my expectation, I processed and interpreted the scenes the way I described above. The examples concerning past experience and expectations thus effectively illustrate how these factors contribute to the mental processes at work and determine not only how acousmatic sound is processed, but also how its handling can be changed by our beliefs and expectations. Moreover, interpretation is influenced by the composer's creation of the acousmatic listening experience and can thus become a prime manifestation of the manipulation of perception through film sound.

3.2 The Listener's Activity

As I have demonstrated in previous chapters, human perception of film sound is diverse due to different contextual factors such as cultural diversity and the individuality of past experiences, but also based on cognitive structures known as schemas. The latter must be examined in the light of the listener's subjective level of activity. Viewers are more frequently represented as passive perceivers remaining in a downplayed role by many theorists (Bordwell 29), than as an audience that is taking an active role in the process of perception. Bordwell's perceptual-cognitive account of the spectator's activity relies on a Constructivist theory of psychological activity, in which "perceiving and thinking are active, goal-oriented processes" (31). This assumption seems to be well-grounded in constructivist psychology, where inference making plays a major role in the realm of mental processing. I agree with Bordwell's approach, as I have indicated in the previous subchapter how the

perception process is working through certain schemata. Bordwell is emphasising that “seeing is [...] not a passive absorption of stimuli” (32) and while he is applying his theory exclusively on visual perception of filmic narration, I transfer this to auditory perception, for which I consider this theory as equally valid.

Perception as an active process primarily affects listening. Thus, as Worthington and Fitch-Houser rightly assume, “how you handle narrative information [in film] will have a profound effect on how you listen” (54). This assumption is strengthened by Melba Cuddy-Keane, who argues that “listening [not only is] the activity that constructs a meaningful representation of the narrative, but [that] hearing is the more inclusive, more integrative sense” (*Modernist Soundscapes* 386). I claim that listening is a highly complex, even skilled activity, investigated in the following as opposed to any passive notion of an audience’s perception. In chapter 1.4.3 I have demonstrated the difference between hearing and listening and portrayed the latter as a cognitive activity. I have further touched upon the individuality of the listening act, thus the “diversity of perceptual capacities of different listeners” (Clarke, *Ways of Listening* 47). In the following, I will investigate how these psychological varieties affect the listening process in detail.

Nondiegetic sound – especially music – as one sort of acousmatic sound accompanying the images on screen, can manipulate how the audience perceives and interprets a film scene. What is special about this type of acousmatic music is its exclusive audibility to the film viewers, which emphasises our difference to the characters in the story world, who are solely able to listen to sound in the filmic reality. This might sound obvious, but is often not as easily separable, since boundaries between different spatial categories, such as story space and nondiegetic space, are intentionally blurred by a film’s narration (Bordwell/Thompson 285). The following examples from *Iron Man* will help to illuminate whether and how music and images can form a unit and explain how this affects perception. In the opening scene of the film [00:01:11-00:01:32], the nondiegetic background music, which is the song “back in black” by ACDC, accompanies images of armoured vehicles driving one after another through a barren desert region in Afghanistan. To highlight the music’s effects, I asked my interlocutors of group 2 to watch this scene without the film sound for the first time and describe their thoughts. Afterwards, they watched it with sound and shared, which effect the sound had on their perception. When only seeing the images, both of them reported that they saw a

sequence of dangerous, almost threatening images that made them feel very uncomfortable. The scene had no positive connotations at all and caused a feeling of anxiety and negative expectations, especially because of the machine guns on top of the armour, which indicate brutal war scenes. Both of them tended to give different, slightly more positive interpretations of the scene when the images were accompanied by music, which both perceived as very loud. Person D said that it felt weird because the music was not what she expected and that it did not match the images. She further stated that the music made the scene look much more spectacular because of the song choice. Person C affirms the gap that opens up between what we hear and what we see, since he felt that this particular song by ACDC in a sense ridicules the scene and calls into question its credibility based on the images. Nevertheless, he also remarks that the music implies the military's firm commitment to start a new attack. This observation indicates that the music accompanying this scene shapes the viewer's expectations about what events may follow next. Their agreement on the gap between images and sound affirms what Bordwell and Thompson call the diegetic music's "temporal unrelatedness" (297) to the story world. What I detect from my interlocutors' comments is exactly this – the interplay of music and images seems to be incoherent in terms of logic. I would argue that the listener's perception is manipulated by the music because the song changes how we process the entire situation: we see armoured vehicles and listen to a rock song, simultaneously. The music is intentionally employed to interrupt the audience's connotations that usually would come up through these images and adds a rather witty touch to the situation portrayed. I detected a similar scene occurring a few minutes later in the film that supports my claim. In this scene, in which Tony Stark is racing by on a highway in his sports car followed by a Rolls Royce [00:12:43-00:12:57], nondiegetic music accompanies this scenario. Similar to the first scene, a rock song, "Merchant of Death" by Ramin Djawadi, serves as background music to the moving images. Although this time, the incongruence between the music and the visual is not as immense as in the first scene, since it seems to fit better together, it still affects our impression of the scene. I would argue that this song makes the scene look far more spectacular and impressive than it would appear without any kind of nondiegetic music. Thus, our perception might be manipulated in the sense that we discern the scene and therefore also the characters – in this case especially the main character – as more awesome, perhaps even intimidating, than they actually are.

By way of contrast, nondiegetic music is at the same time able to form a unity with the images of a scene to which it is applied. In the scene in which Pepper is on her secret mission to hack into Tony's computer in his former office to save essential data [01:30:44-01:31:57], the excerpt from the song "Save my Love (Remastered Version)" by Irakli Kolbaja serves as background music to complement the images on screen. The fact that the music is instrumental and rather quiet, creates a more instinctive and genuine relatedness of the images and the acousmatic sound. My interlocutors specify that they feel with Pepper because she is a sympathetic character and that the risky situation's tension felt by her is transferred to them. These insightful accounts highlight the integration of the sound into the scene, as they both did not explicitly mention the sound and its effect in the first place. It was not until I asked them specifically about how they perceived the music in this context, that they reported how the sound influences their perception of the scene. Both gave similar accounts of how the sound functions in a significant way: the music makes the scene appear more dramatic and far more suspenseful than it would be without music. One of them even stated that the suspense felt almost unbearable. I argue that these impressions and the resulting emotions are primarily elicited by the nondiegetic music applied to this scene. Ben Winters suggests that when filmmakers attempt to "create a feeling of verisimilitude" (229), the "music's presence seems entirely natural, rather than a troubling element that needs to be assigned to a separate level of narrative" (229), and I maintain that this is exactly what comes to light in this exemplary scene. It proves how story world and real world can be blurred in film due to the perceptual manipulation acousmatic sound can elicit.

Sound that is acousmatic to the listener but originating from the diegetic world of the film, can have different effects on the listener depending on individual listening styles. Some of the factors that have an influence on cognitive processes on which I have elaborated in the previous section and chapter. Still, I would like to include the following varying types of listening categorised by Worthington and Fitch-Hauser. These help to structure the analysis of the following examples from my objects of study: "personality traits", "listening styles" and "empathy" (Worthington/Fitch-Hauser69). I will address the category of "empathy" as a significant area on its own in chapter 3.4 in greater detail. Film scholar and semiotician Christian Metz uses the notion of diegesis to describe the 'reality' of fiction (*Film Language* 10). For him, this 'reality' "comes only from within us, from the projections and identifications that are mixed in with our perception of the film" (ibid.). He indicates what I hope to explain

in this chapter, namely that the sound we perceive through fiction – in this case film – is dependent on various factors that concern personality and individual perception.

As an example of individual listening, I will discuss the scenes from *Lost in Translation*, in which Bob is making phone calls to his wife Lydia. As the audience, we are never able to see her, but to hear her acousmatic voice offscreen, which implies that the sound's source is invisible.



Figure 6. Bob lying in bed, talking on the phone to his wife Lydia [00:54:33].

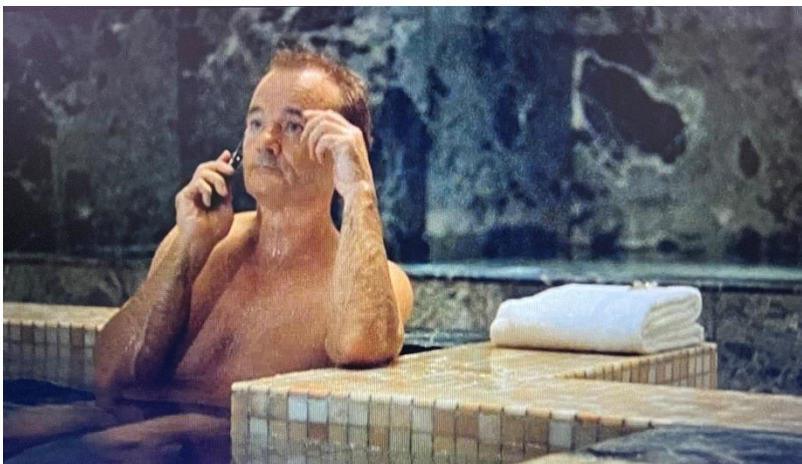


Figure 7. Bob in the hotel pool, talking on the phone to his wife Lydia [01:16:44].

Her voice is introduced to us when Bob receives a fax from her, which she reads out loud, asking him to choose new shelves [00:05:30-00:06:01]. The following phone calls mainly revolve around trivial things such as the colour of the carpet. But underneath these monotonous conversations various listeners perceive different sentiments, since I argued that listening is a highly active and individual process. The different listening types come into play here, as they can help to analyse different interpretations of the acousmatic voice in the

presented context in the film. Firstly, ‘personality traits’ differentiate us from each other, which can be illustrated by the example of introverts and extraverts (Worthington/Fitch-Hauser 69). Whereas extroverts are characterized as generally slower listeners as they are quicker to express their feelings to the outside, introverts are portrayed as quicker listeners but as personalities who rather suppress their own emotions (Worthington/Fitch-Hauser 71). And although traits are “believed to be motivational in nature” (Worthington/Fitch-Hauser 70), they are also shaped by someone’s social environment, including experiences and the people one is surrounded by. What does that mean in terms of perception? I would argue that an extrovert person listening to the phone calls could possibly miss what is underneath the surface of these conversations: a shattered marriage in which Bob and Lydia seem to be separate entities rather than a unity. The disembodied voice, which is never deacousmatized to the audience, reaffirms my assumption of Lydia being a distant character who apparently lives in a different reality. Additionally, this presumption is supported by Bob’s blank look and annoyed facial expression. An extrovert as characterised by Worthington and Fitch-Hauser, who does not seem to be an accurate listener in the first place, might be more engaged with his or her own feelings and reaction to the presentation of their relationship. However, an introvert listener, who tends to listen more carefully, might devote himself to the characters’ emotions and try to figure out the issue of loneliness and isolation in their emotionally cold interaction. There is a profound classification of further personality types⁹, which contribute to the individual perception of acousmatic sound to the same extent.

Subsequently, I will consider different ‘listening styles’ regarding the change of listening conditions in an acousmatic situation. In chapter 1.4.3 I have introduced different ways of listening, among which “causal listening” (Chion 22) is the most prevalent mode, implying that a sound is listened to in order to detect its source or cause. But the detection of a sound’s cause is often not even possible when acousmatic sound is not deacousmatized in a film, like Lydia’s voice in these examples. Worthington and Fitch-Hauser are concerned with listening styles as well and have established different listening orientations. A “Relational listener” (Worthington/Fitch-Hauser 78) is someone who “will be more concerned with, ha[s] a greater awareness of, and be responsive to the feelings and emotions” (ibid. 78) of others. In contrast, “Analytical listening” (Worthington/Fitch-Hauser 79) means “to listen

⁹ For a summary of personality types and attributes associated, see Worthington/Fitch-Hauser (2018: 76).

systemically” while “focus[ing] on the speaker’s entire message” to “then form a judgement” (ibid.). These different types¹⁰ not only validate the act of listening as a highly initiative activity, but also reaffirm that its mode is central to our perception. A person that is characterised by analytical thinking “prefer[s] to be logical, objective and fair” (ibid.), which would suggest that this person, listening to the phone calls respectively Lydia’s acousmatic voice, may be rather unbiased concerning the characters and their relationship. He or she would probably consider their connection impartially to form an opinion based on rational factors and seek for possible solutions. A ‘Relational listener’, by contrast, is a less practical and analytical thinker because he or she is far more emotionally involved when perceiving the characters’ relationship. This type of listener, for example, would be concerned with the inflection of Lydia’s voice and interpret it in terms of her feelings, whereas the analytical listener is more concerned with what is said rather than how it is conveyed. The relational style of listening “is central to establishing and sustaining interpersonal relationships” (Worthington/Fitch-Hauser 78), which suggests that these listeners tend to build a stronger connection with Lydia, although she is invisible, or precisely for this reason. Due to the acousmatic situation, listeners of this type fully engage with the characters and their emotions, why they tend to “draw upon [their] view of the situation and [their] relationships with others” (ibid.). This means, the voice being acousmatic has a great impact on how different listeners perceive what is happening, since it leaves no space for a listener’s analysis of the voice’s visual source. Each individual perceives acousmatic sound differently, especially a voice that centers our attention (Chion, *The Voice in Cinema* 33), which implies that the degree of manipulation differs according to individual perspectives of listening. Psychological differences with special regard to personality traits and listening styles, have been revealed to make a difference in the perception of acousmatic sound, which confirms that listening is a highly active and individual process. We are certainly able to influence how we listen, nevertheless, we are pre-shaped by the above mentioned components, which define our mode of perception and the consequences this has for our reality.

¹⁰ There are further types of listening that can be distinguished according to Worthington and Fitch-Hauser, which I did not consider to be of great concern for the analysis. For the complete overview of listening types, see Worthington/Fitch-Hauser (2018: 76ff.).

3.3 Imagination and the Simulation Process

Another important strategy of cognitive processing is that of imagination. According to philosopher Gregory Currie, who is coming from the background of cognitive film theory, imagination is supposed to be a 'mental mechanism', which works by means of "simulation" (145) in order to understand how the minds of other people function. The "Simulation Hypothesis" (ibid.) occurs in the following way:

I imagine myself to be in the other person's position, receiving the sensory information the other receives. Having thus projected myself imaginatively into that situation, I then imagine how I would respond to it: what beliefs and desires I would have, what decisions I would make and how I would feel having those perceptions, beliefs and desires, and making those decisions. (Currie 144)

For Currie, imagining in this context means to take on the other person's mental processes momentarily to experience these as one's own, with the difference that those processes work "off-line" (ibid.), which means, they are disconnected from real beliefs and desires. As listeners engaging in the process of mental simulation, we can use our own minds as "a reliable model of the other person's mind" (Currie 145) to observe how we would respond to a situation in imagination. I chose this approach because it contributes significantly to the auscultator's understanding of fiction and thus, the characters' minds in a film. Currie claims that in fiction, when we simulate the mind of another, "what we acquire instead of beliefs is *imaginings* which simulate belief" (148). Hence, "[we] imagine being in a situation [we] could be in but actually [are] not in" (ibid.), which is the reason why Currie regards our real "belief-desire system" (ibid.) as completely separable from that of a film character for example. I will analyse whether this assumption is generally valid and applicable to sound in film, since I claim that the characters' feelings influence our emotions not just "off-line".

According to Currie, there is a distinction between two kinds of imaginings concerning the understanding of fiction. "[I]magining what is fictional in a story" (152) is what he calls "*primary* imaginings" (ibid.). The other kind, which he coined "*[s]econdary* imaginings" (ibid.) appears "when we imagine various things *so as to* imagine what is true in the story" (ibid.). I would like to clarify this with the example of Stu's situation in *Phone Booth*. Simply imagining Stu standing in that phone box being threatened by the anonymous caller represents our 'primary imagining'. Clearly, Stu is not only standing there but experiencing an auditory

situation in which he has certain thoughts and anxieties, which requires ‘secondary imagining’. Although the accuracy of this experience is up to the filmmakers, it is the “imagination, to feel as the character feels” (Currie 153), that constitutes ‘secondary imagining’. In *Phone Booth*, I would argue that a large part of the audience empathises with Stu and sees things from his perspective. This is mainly because he shows his vulnerable side and confesses his mistakes as he is threatened and thus turns into a much more likeable and sympathetic character as the film progresses. Furthermore, Stu and the listeners have in common that the unknown caller is an acousmètre almost until the end of the film, which emphasises the commonality of the character and the audience. By putting ourselves in Stu’s position in imagination, we are able “to have imaginary versions of the thoughts, feelings and attitudes [we] would have were [we] in that situation” (ibid.). By way of this “empathetic reenactment” (ibid.), we are able to identify the characters’ emotions to then “imagine that the character felt *that way*” (Currie 154). I would say that a successful ‘secondary imagining’ also depends on the diverse factors and schemata I have discussed in previous subchapters, which differentiate us from each other in terms of perception. Nevertheless, the ‘Simulation Hypothesis’ assumes that based on evolutionary theory, humans are mentally similar enough in order to simulate the mind of another effectively (Currie 146). It is important to remark that the simulation process is not thought of as “a conscious or even an intentional action” (Currie 161), but that it “must be done unintentionally, mostly at a subconscious level” (ibid.). I agree with Currie here, since based on my own experience of watching the film, I found that I automatically projected myself into that situation, questioning myself how I would behave if I were Stu.

Through imagination, film enables the audience to make inferences about how the characters’ minds function. By means of another example, I will illustrate the simulation process as proposed by Currie in terms of acousmatic sound and shed light on how listeners can be deceived in terms of perception. In *Stranger Than Fiction*, the main character’s situation is this: Harold hears a disembodied voice narrating his everyday life and habits and is totally distracted on an emotional level. As a condition for the simulation process, it is helpful in this case that Harold tries to figure out where the voice comes from [00:04:31-00:04:41 or 00:04:47-00:04:54] but is not successful in finding the voice’s source. According to Ivan Leudar and Philip Thomas, these attempts of pursuing the voice can be called “reality-testing” (203), which is a process that helps the “voice hearers [to] know that only they hear their voices” (ibid.). Commonly, ‘reality-testing’ is performed through “establish[ing] the

source of the voice" (ibid.) and through "seek[ing] social consensus" (ibid.). The former is done by Harold in the example scene I just mentioned. When he fails to identify the voice's source, he is applying the latter and repeatedly asks a strange woman at the bus stop, "I'm sorry, did you hear that? The voice, did you hear it? Harold just thought it was a Wednesday?" [00:06:08-00:06:24]. He hopes to find approval, but her answer makes clear that she does not know what he is talking about, which illustrates that she does not hear the voice [00:06:18-00:06:30]. Harold tries to "seek social consensus" one more time at work, when he reports to his friend Dave that he is being followed by a voice. He is desperately searching for confirmation and says, "Listen" [00:08:12], while filing and looking expectantly towards his colleague. When he asks, "Did you hear that?" [00:08:34] and Dave denies having heard anything, Harold is faced with a puzzling experience: a voice that is clearly heard by him alone and that seems to know everything about him [00:08:42-00:08:46]. Dave's and also the woman's reaction both indicate that hearing a voice is something considered as abnormal in society. This implies that hearing a voice is – definitely in this case – a private experience, which can even be considered schizophrenic, since no one in Harold's environment is able to hear that same voice. The confirmation that this voice is only audible to our main character, encourages the simulation process as proposed by Currie. When we spoke about *Stranger Than Fiction* and the main action concerning the protagonist and the voice, I asked two of my interlocutors (focus group 1) how they perceived the acousmatic voice and whether they at some point imagined being in Harold's position, without speaking of Currie's imagination theory. Their first impression of the voice seems to be unremarkable, as they perceived it as an ordinary omniscient narrator depicting the main character and his everyday life. One of them reported that her impression changed when it became apparent that Harold can hear the voice, which she mistakenly assumed to be nondiegetic [00:04:28]. She could not further describe her feeling other than she found it odd and felt somehow torn between perceiving the sound as a narrator's voice outside of the story world or through Harold's point of view. I would argue that this is clear evidence that a listener's perception is manipulated through acousmatic sound. This person seems to be torn between perceiving it through her own sense or through the character's perspective, which is caused by the fact of the voice being acousmatic. This uncertainty causes different emotions, which I will discuss in a moment. Furthermore, both my interlocutors reported that as soon as they realised this voice is audible only to Harold, they unconsciously started to imagine themselves being in Harold's position.

What would they do? How would they react or behave? These were the questions they asked themselves in the course of the film. I claim that this coincides with how the simulation hypothesis works according to Currie. Here, both persons project themselves to be in the character's position to respond to the situation in imagination. Both my interlocutors seemed to be able to imagine thoughts and feelings as if they were in that situation. They referred to specific situations, as for instance Harold's emotional outburst after the voice tells that he is going to die [00:21:37-00:22:42]. One person said that she was explicitly imagining how she would feel and what it would do to her when an acousmatic voice who seems to know every detail about you, predicts your own death. She said that she would feel desperate and frightened, probably similar to Harold who is devastated in that moment. The other interlocutor confirmed these simulated beliefs but stated that these changed as soon as the narrator's voice was deacousmatized to the audience [00:13:24-00:13:35]. She reviewed that she once again adopted the point of view from the outside, like in the beginning when the audience did not know that Harold could hear the narrator. This is insightful in terms of perception, which apparently is influenced by an acousmatic voice in a different way than by an embodied voice. Based on Currie's approach, their mental processes happen 'off-line', which would mean that their imagined emotions are disconnected from real beliefs, since she is not actually in that situation. I agree with this point since emotions and decisions are imaginary as long as she is not in fact in that situation. However, I contend that our real 'belief-desire system' cannot generally be assumed to be strictly separated from that of a film character for example. I would rather say that a character's emotions are actually capable to influence our sensations in the real world and can have a lasting impact on our mental processes.¹¹ Currie himself admits in one of his examples that "[i]magining yourself in danger causes you *really* to feel disturbed" (150; emphasis added). This implies that mental processes to understand fiction and those in reality are connected indeed. As I highlighted, I endorse this approach and Currie's statement about the dangers of imagining, namely that imagination might become a real belief and thus, blur the boundary between fiction and reality (162). Nevertheless, there is a slight disagreement in Currie's theory with regard to the disconnection between the imagined and the real mental processes. However, the simulation process as a special type of cognitive processing method is illuminating in terms of perceptual manipulation. I have demonstrated that through the process of imagination, sensations are

¹¹ I will further expand on possible implications in the concluding chapter.

influenced by acousmatic sound, since the simulation hypothesis works accordingly successful or not. I will amplify repercussions of the analysis's result in the following concluding chapter.

3.4 A Cognitive Understanding of Emotions

With respect to different cognitive processing strategies, an exploration of how emotions are represented in film and experienced by the audience with regard to acousmatic sound, is essential. I will illustrate the relation between emotions and cognition in terms of mental processing and show how it contributes to a possible manipulation of perception. Emotional empathy plays a crucial role in this regard and is defined as “the ability to be affectively triggered by the emotions of others, commonly known as emotional contagion or empathetic arousal” (Schumacher Ratti/Fuentes Bravo 396). I will give examples of how emotions on the listener's side coincide with those of the characters or how they differ with reference to the acousmatic sound's influence on the processing of these sentiments. Examining emotions is of such great significance as they influence beliefs, which are considered by Nico Frijda and Batja Mesquita as “a proposition that a person considers to be true” (3). According to the theorists, emotions affect beliefs in two possible ways: “they may give rise to beliefs where none existed, or change existing beliefs” (Frijda/Mesquita 3) or “they may enhance or decrease the strength with which a belief is held” (ibid.). In a broader sense, this is important in terms of the impact that a manipulated perception can have on a human's real belief system. I will further reflect on the implications of this in the following concluding chapter.

Concerning film, I have analysed in the previous chapters in how far an acousmatic listening experience constitutes an individual activity based on diverse factors, which influences the audience's perception. I agree with Frijda that the emotional experience “is likely to differ according to the circumstances under which it occurs or is examined” (*Emotion experience* 475), suggesting that the degree of sentimental involvement plays a major role in terms of perception. I detected this for example in different scenes from *Iron Man*, in which we can hear the electronic and acousmatic voice of J.A.R.V.I.S., Tony's artificially intelligent system.

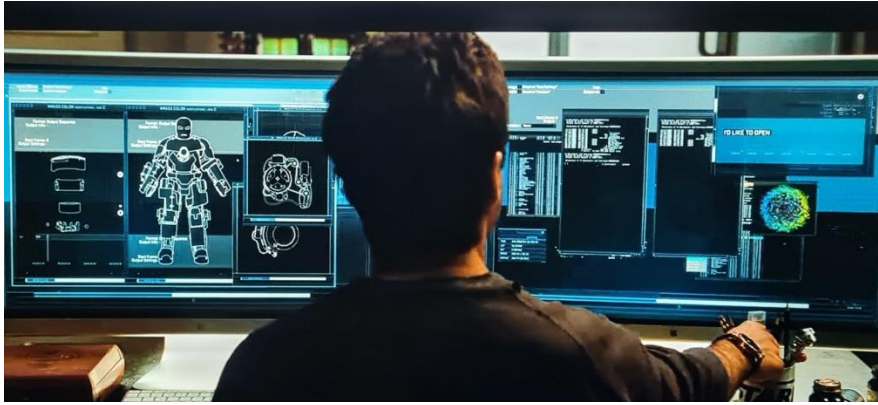


Figure 8. Tony and J.A.R.V.I.S. talking to each other [54:52].

When I asked my interlocutors which effect this voice's sound has on them, both respondents stated that J.A.R.V.I.S. did not considerably affect their emotions. I suppose this is the case because he is not a human being and is not even seen by the audience. An emotional connection might have been established, if he was visualised in any way as a human being. In his theory of emotions, Frijda explains that "[e]motions occur when a situation is relevant for an individual's concerns" (Tan/Frijda 51), what has been proven not to be the case here. Neither what J.A.R.V.I.S. is saying, nor his voice itself seem to awaken an emotional response in the listener, as proposed by my interlocutors. I would argue that according to Tan and Frijda, this response can be termed "*nonempathetic* emotion" (52), since the voice is perceived regardless of its influence on the characters and thus, without invoking an emotional reaction from the listener. This means, the voice being acousmatic influences the auscultator's perception of that same voice in terms of an emotionless interpretation.

Essentially, there are various instances in which acousmatic sound manipulates the audience's perception in terms of their emotions. This is particularly true of situations, in which the characters' emotions are passed on to the audience, sometimes on purpose and at times unconsciously. Tan and Frijda claim that "[f]ilm-elicited emotion [...] consists largely of *witness emotions*" (52), which implies that the factors which shape the emotional experience while watching a film coincide with the ones in reality. When we, for example, observe another person, with whom we have a relation and who is emotionally involved in a situation, we are usually not able to intervene, but we care about their destiny as witnesses (ibid.). An extreme example would be a tragic event in a friend's life, such as the loss of a family member. In this case, a person would feel the friend's pain and grief because he or she deeply cares and worries about them, but the person cannot take the pain away. Such a situation is similar to

engaging with fiction, since “viewers of films are led to imagine themselves as invisible witnesses that are physically present in the fictional world” (ibid.). Tan and Frijda call these feelings that arise as response to occurrences in a film “*F emotions*” (ibid.), which are referred to as “*empathetic emotion[s]*” (ibid.). What constitutes these emotions to a great extent is the importance of a situation for the character since this is meaningful to the audience’s sentiments (ibid.). Thus, in contrast to the perception of J.A.R.V.I.S.’s acousmatic voice, there are others that influence how we perceive the character’s emotions and that evoke real feelings. In terms of voices, I can clearly detect these so-called ‘*F emotions*’ in *Stranger Than Fiction* and *Phone Booth*. The most important and common feature of both films is that one acousmatic voice, though deacousmatized at different times to the audience, is audible to the respective main character. Even if the events in both films could not unfold more differently – apart from the threat of an impending death – there are several situations in which the audience is emotionally engaged. I have already demonstrated in chapter 3.3 in how far the audience is able to identify with Harold in *Stranger Than Fiction* according to the simulation process. The result is that listeners become instinctively involved and experience empathetic anger and despair when Harold feels this way, for example when he leaves the bakery and shouts angrily at the voice, while the camera zooms out upwards [0:13:13-0:13:22].



Figure 9. Harold outside of the bakery [0:13:14].



Figure 10. Harold from a bird's eye perspective, desperately shouting at the voice [0:13:20].

The camera movement is interesting in this scene, as I would argue that it emphasises the acousmatic voice's impact and the power it has over Harold, since he is seen from above in the last frame of this scene. Up to this point, Karen Eiffel's voice is what I introduced as acousmètre, since she has complete control over the character's thoughts, highlighted by the camera angle through which Harold appears to be smaller and smaller.¹² As audience, we empathise with him and understand his feelings, which is mainly because of the voice being acousmatic to the character and to us alike in this scene. This means, our perception is manipulated by the acousmatic voice in the sense that we do not have another reference – in this case the visual embodiment – that could equally contribute to how we perceive the voice and hence, the character's emotions. On the contrary, the fact that the sound is acousmatic intensifies our 'empathetic emotions', as we are in the same position as Harold in the beginning in terms of the relation to the voice. Whereas in this case, the listeners' and character's emotions seem to correspond, they might differ after the voice has been deacousmatized to the audience [0:13:24-0:13:35]. This is indicated by the ambivalence reported by one of my interlocutors, which she felt between the narrator's and Harold's perspective, to which I referred to in the previous subchapter. Although I claim that the audience's emotional engagement is stronger due to a shared perception of acousmatic sound

¹² In the shots of the following scene, the camera movement complies with the previous one. We see the upper body of a woman from the back, standing on a roof top smoking while the camera pans over her head. It then stops in order to emphasise her point of view, which we as audience share: she looks down on the streets of Chicago, where Harold is standing in the previous scene [0:13:23-0:13:32]. This does not only reflect Karen Eiffel's narrational omniscience, but at the same time constitutes the beginning of her deacousmatization [0:13:23-0:13:35].

with a film character, their empathetic feelings are still present after the voice has been deacousmatized, as for instance in the scene of Harold's outburst [0:21:36-0:22:55]. Although this situation is characterised by "some knowledge discrepancy between the viewers and the protagonists" (Tan/Frijda 52), which consists in the fact that the audience is acquainted with the voice's visual source by now, 'witness emotions' are still present. The audience emphasises with the character, but Tan and Frijda propose that "the viewer's emotion does not parallel the one observed in the protagonists" (53). I agree in the sense that sentiments cannot be completely equivalent since I maintain that the exact same emotions can only emerge if the same situation has been experienced. Probably, not many of the listeners have heard voices that only they could hear. Nevertheless, the example of Harold illustrates in how far the audience's emotional experience can be shaped by acousmatic sound.

In *Phone Booth*, emotional empathy on the part of the auscultators is indispensable to make the narrative work in the view of the filmmakers. This is mainly because the protagonist's feelings with regard to the acousmatic voice is the major component of the film. Stu's and the audience's crucial point in common is that both cannot see the source of the sound – in this case the unknown caller's voice – almost until the end of the film. Therefore, I argue that the listeners' emotional involvement is even stronger as in the previous example from *Stranger Than Fiction* in terms of empathy. What further contributes to this assumption is the degree of authenticity that can be assigned to Stu's situation. By this, I do not mean that it is very realistic for the audience to enter a phone booth in New York City and become a possible victim of a sniper in reality. My point is that the idea of being threatened by a stranger, who cannot be seen, and the associated emotions such as fear and despair that arise are more likely to be experienced than hearing a voice in your head that is not audible to anyone else. My interlocutor's reviews on how they perceived the acousmatic voice support my argument that 'F emotions' are present with regard to emotional empathy for Stu. Although I still agree with Tan and Frijda that the audience's emotions do not parallel that of the characters (53), it becomes clear in this example that emotions of both are extremely close to each other. Person A and B reported that they experienced feelings of anxiety that felt real and when I asked whether they felt sorry for the main character, they affirmed this. I assume that the majority of the audience feels this way for Stu, which is why the role of empathy is of such importance here. Another explanation can be found in Torben Grodal's account of the interplay of emotions with perception. He approaches film viewing via a "mental flow" (132),

which is an essential frame for comprehending the film's emotional impact. The series of mental processes he presents, largely coincide with what I presented in the previous chapters with regard to schemata and different associations such as past experiences. A new aspect in this approach is his ecological interpretation of emotional perception, as he illustrates that "our brain and body have evolved as tools for survival" (ibid.). This is why he indicates that films that are "related to survival or to human bonding, power, or sexuality" (ibid.), evoke a "stronger emotional impact" (ibid.). I claim that this is what happens in *Phone Booth*, in which Stu is desperately trying to survive and to rescue his relationship with Kelly. Moreover, as I have indicated above, Stu's situation has a powerful emotional effect on the audience. According to Grodal, these strong reactions are linked to the interconnectedness of the perceived emotions and possible actions (ibid.). I agree and would like to emphasise that in this case, it is not sight, but acousmatic sound that activates associations and possible reactions, which are "either mental tendencies toward possible future actions or concrete actions" (ibid.). The effect that this change in sense-perception has, is that the respective stimuli, such as a fear-inducing stimuli in Stu's case, enhance the auditory perception. The improvement in perception allows for a more powerful response to emotional experiences. 'Concrete action' is represented by this immediate answer when watching a film, which coincides with what Worthington and Fitch-Hauser coin "[a]ctive-empathetic listening" (82). They assume that this type of listening is connected to personal empathy and characterised by "cognitive and behavioral elements" (ibid.), which enable the listeners to sense, process, and respond to the "emotions of the people involved in the interaction" (ibid.), thus of the characters in film in this case. Both approaches support my claim that film-elicited emotions have a huge impact on the audience's real belief system. In *Phone Booth*, I have illustrated in how far the audience's emotional experience can be manipulated by acousmatic sound. The voice, which is not visualised until the end of the film, determines to a great extent how listeners perceive and process the situation in terms of emotions. This is significant with regard to the impact that a manipulated perception can have on real emotions, which have potentially important implications for beliefs and mental health.

Conclusion

As has been analysed in the previous chapters, acousmatic sound in its various modes of appearance can contribute significantly to manipulating both the characters' perceptions of experiences within the story world, and the audience's perception when attending to a film. Especially acousmatic music and voice emerged as categories of sound that are typically marked by manipulative aspects. I have demonstrated how sounds without a visual origin affect mental processing by means of an analysis of the characters' experiences and by investigating the cognitive state of the listener and the respective processes. In the following, I will discuss my findings according to the precedent chapters. The aim is to illustrate this innovative mode of engagement with fiction as related to sound, and also to debate parallels between reality and fiction in regard to human perception. Having focussed on acousmatic sound in film, I will consider the degrees of manipulation that come along with processing the auditory phenomena in my research objects. Finally, I will reflect upon possible consequences of film viewing in relation to auditory experiences. Perceptual manipulation through acousmatic sound in film induces certain effects on human cognitive and emotional processes, which in turn have an influence on mental tendencies concerning emotional reactions and actions. I will illuminate how these effects make an impact on individual and social life and critically engage with the question whether the health-related quality of life is influenced by film on behalf of audiovisual media in a harmful way.

In light of the importance of sounds' influential and deceptive powers, this comparison of findings serves for a better understanding of the mental phenomena found and its ramifications for perceptual processes in reality. In *Iron Man*, two main categories of sound are repeatedly displayed: nondiegetic sound in the form of music, and that of an acousmatic voice, represented by J.A.R.V.I.S.. Of all films, *Iron Man* is the most common sample of how nondiegetic music is applied to influence the audience's perception. It further is a representative work of the action genre, which lives by sound from outside of the story world and in which mostly there is a discrepancy between the characters' and audience's processing of sound, since it is only acousmatic to the latter. At least, this is the case for nondiegetic music, through which a coordination among sound and image is achieved or not. On the one hand, I have analysed examples in which the nondiegetic music does not match the images, especially in terms of rhythm, intensity, and time [e.g., 00:01:11-00:01:32 and 00:12:43-

00:12:57]. On the other hand, I have investigated a scene in which the gentle music playing in the background matches the suspenseful creation and sustains the level of tension [01:30:44-01:31:57]. What became clear is that the degree of correspondence significantly affects perception. A disparity between the images and the sound is what emphasises the “underdetermination and uncertainty in acousmatic listening” (Kane 148). I agree with Kane and argue that listeners are not only uncertain and confused by this kind of listening, since they cannot see a visual source, but also manipulated in various ways and to different degrees, as I have demonstrated that acousmatic sound is experienced individually. Furthermore, I observed that it could have a profound emotional impact, which I claim originates from the “structural gap” (149) that Kane identifies as “the being of acousmatic sound” (ibid.). I will illustrate in the following how this ‘gap’ is significant in its connection between acousmatic sound and the listeners’ real world sensations.

Although there are numerous nondiegetic pieces of music in *Lost in Translation* that reflect how the audience should perceive and interpret a scene, I mainly focused on the voices that are audible via phone calls for my analysis. This type of acousmatic sound is more complex in that it provides the audience with the characters’ perceptions and confronts them with their own interpretations, as I have analysed in chapter 3.2. My findings suggest that the degree of the emotions that the film audience assigns to the voice, is a determining factor on how the acousmatic sound is experienced and thus, how manipulative it can be. With regard to these phone calls, for example, one of my interlocutors stated that he actually perceived the conversations between Bob and Lydia, but that he instantly transferred the upcoming feelings onto his own relationship with his partner. He said that he worried about becoming like them someday, in the sense of being in a relationship in which both partners live totally separate lives and can only talk about trivial things. I consider his feelings as authentic and argue that they can be ascribed to the ‘gap’ that opens up through acousmatic sound, which is – often unconsciously – filled with emotions. This is in line with what composer and sound artist Thembi Soddell describes as the “internal, reflective quality of acousmatic experience” (347). She proposes that within the ‘gap’, imagination unfolds, and the mind is able to connect the sound with events of the real world in respect to emotions that affect the subconscious: “more abstract concepts, thoughts, memories [...] or lived experiences” (ibid.). I share Soddell’s understanding of acousmatic sound, as my example above suggests these deeper insights of the inner world of a person based on an acousmatic listening experience in film.

Despite the different types of representation of the disembodied voice according to their distinct genres, filmmakers share similar understandings of applying it in order to manipulate human perception in *Phone Booth* and *Stranger Than Fiction*. One of the key differences compared to *Lost in Translation* is in terms of emotional empathy since I have demonstrated that in both films, listeners build a bond with the main characters in terms of the acousmatic experience. As audience, we share perception with the main characters for as long as the voice is acousmatic. My analysis has shown how the protagonists are manipulated by the acousmatic sound on the one hand, and how individual listeners interpret the same auditory events primarily according to cognitive processing strategies. The close investigation of the unsettling nature of Karen Eiffel's voice in the comedy and that of the unknown caller in the thriller, support my argument that acousmatic sound manipulates perception in view of the listeners' responses and more intense mental processes at work. Soddell suggests that this mental activity in the form of imagination is "inspired by acousmatic listening [and] can be a pathway to reason and self-understanding" (351). Being aware of the importance of my twofold approach in terms of the fictional world in film, represented by the characters, and the real world, portrayed by listeners, is crucial to further explore the audience's mental state and possible consequences of acousmatic listening.

The implications for listeners of a detachment of sight and sound by way of acousmatic listening are profound. I claim that sound as I treat it in this thesis, is capable of manipulating our perception and thus, of influencing real-world sensations. One of the reasons why I was interested in the analysis of film sound in terms of perception in the first place, is because this medium represents the interface between two components: fiction and reality. There is nothing more intriguing than to investigate how perception, that is manipulated through acousmatic sound in film, can affect our real life on various levels. French electronic music composer Éliane Radigue proposed in an interview in 2012, that "[t]hrough music you can reflect anything. Sounds always reflect something from the mind" (*Telekom Electronic Beats* n.p.). This assumption is very plausible, since I have illustrated in the previous chapters that responses to sound are linked to other levels of reality, such as memories or anxieties. I would argue that acousmatic listening reflects a way for us as the audience to detect these different layers and to offer ourselves the opportunity to bring to the surface these underlying feelings that would otherwise remain on a subconscious level. This might be achievable on completely different levels. I recall, for example, my interlocutor's negative connotation of the unknown

caller's voice in *Phone Booth* due to her recurring experiences with a certain voice in her childhood. At first it seems that this type of acousmatic sound has a harmful impact on her, as listening might intensify the unfavourable association and deepen her anxiety. Not seeing the visual source in form of an embodied person could trigger fearful illusions within her. However, engaging with sound this way could provide a chance for her to confront herself with the burdens from the past and the related emotions. For another person, for instance, acousmatic listening means connecting to a level of creativity in a more positive way by using the released feelings for example in terms of imagination. How listeners respond to acousmatic sound in film is an individual and subjective endeavour, through which humans uncover diverse meanings and manifestations in the real world.

It is of interest to connect the listeners' acousmatic experience to how meaning is constructed in everyday life in order to illuminate how life's quality might be influenced. With regard to electroacoustic music, composer Gary S. Kendall is concerned with a perspective on listening, that focuses on the correlation between the environment and the listener's perception of it (63), which refers back to the ecological approach of perception according to Clarke. Kendall points out that "without our conscious awareness" (ibid.), meaning is constantly produced by the mind in an ongoing active process of perception. In terms of auditory experiences, listeners tend to refer to various domains of perception in their "network of mental spaces" (Kendall 73), to be able to understand and organise their environment (Kendall 64). Thus, he underlines that meaning making in acousmatic listening works accordingly to that in everyday life (Kendall 73). Kendall's approach is important with respect to my analysis in terms of the way that films are engaged with. I have based my argument on the assumption that the kind of films that I have examined, are watched without preconceptions. This means, they are not usually viewed as objects of analysis, for example in relation to unconscious cognitive processes. This is why, in light of an average audience, the way of experiencing meaning in film is well transferable to everyday listening. Although I have highlighted that it is rarely possible to identify objective consequences, I maintain that each individual's inner state is manipulated through an acousmatic listening experience in film.

The notion that our perception of film sound works equally in everyday life, substantiates my claim, and provides space for the question whether mental health is influenced in a harmful way in respect to acousmatic sound. I can only try to answer this with

reference to my analysis. It is necessary to notice and consider the negative affects – I call them uncertainties – that acousmatic listening induces. These include for example the negative associations with voices, or the alienated feelings that appear when listening to nondiegetic music in film. I admit that these uncertainties are clearly manipulative in terms of perception and can even be harmful at times. Again, the entire processing system is very individual and depends on the mental state of each person. Nevertheless, I claim that there are positive effects of acousmatic listening to the same extent. For example, we develop a more precise and thorough understanding of a character's emotional world. Although characters originate from a fictional realm, it must be remembered that they are still represented by real persons, who are able to experience the same emotions as the audience, which refers back to how meaning is processed in everyday life. Hence, I would argue that the ability of empathy and compassion is significantly enhanced by acousmatic listening. Additionally, sound of this type opens up possibilities for a profound self-awareness, which can lead to redefining for example past experiences, behaviours, or values in general. These insights indicate that a change in perception caused by an acousmatic listening experience in film, result in a transformed cognition in reality. Therefore, acousmatic sound can have a lasting impact on our mental activity, but not necessarily in a harmful way, as I have shown in the preceding chapters.

I have demonstrated, in this thesis, how acousmatic sound in film can be a means of perceptual manipulation in various ways. Whereas previous studies have mainly dealt with how visual scenes in fiction affect human mental processes, my analysis has been concerned with illustrating that auditory aspects are equally important with respect to cognitive perception. Hereby, I focused on acousmatic sound that is applied in different ways in my objects of study, to highlight the significance of the detachment of sound from sight. Four films of different genres, *Lost in Translation* (2003), directed by Sofia Coppola, *Iron Man* (2008), directed by Jon Favreau, *Phone Booth* (2002), directed by Joel Schumacher, and *Stranger Than Fiction* (2007), directed by Marc Forster, helped me to prove that acousmatic sound and its various manifestations influence the characters' and the film audience's perception. Concepts from Sound Studies, Film Theory and Cognitive Theory have helped to shed light on the mental dimension and profundity of acousmatic sound. Starting from this interdisciplinary approach, I supposed in my introduction that the investigation of sound would illustrate its uniqueness in terms of the notion of the acousmatic.

To understand what roles sound can play in film, I first of all examined its qualities and how it serves as a guidance for humans in their acoustic environment. Afterwards, I have provided the reader with a description of the development in sound technology to understand its application in film. In investigating the ways in which the auditory is made use of in film, my analysis has given an account of the complex acousmatic phenomena in regard to sound. These different phenomena, especially acousmatic zones, music, and the voice in the films revealed the intricacy of sound layers in film and how they contribute to perceptual manipulation. In the analysis, it became clear that listening in various forms is fundamental to the degree of possible influences that acousmatic sound can have on the audience's mental activity. The results indicated how much meaning sound, which is detached from a visual source, conveys, and that it is essential to different strategies of cognitive processes. Theoretical concepts such as Currie's notion of 'imagination' and the 'simulation hypothesis', for instance, were extremely beneficial in explaining how acousmatic sounds in the films are perceived and mentally processed by listeners.

Methodologically, a combination of a qualitative in-depth study of the research objects and reviews from my interlocutors provided a powerful tool to investigate the inner state of the characters and the role of the audience's mental processes. Various occurrences of music and different voices in particular, are examples of manipulating instruments used in film that determine how the mode of listening can be influenced and how acousmatic sound is cognitively processed. By illuminating how the human mind works according to different strategies, I discovered that certain modes such as information-processing, imagination, individual listening styles, and emotional empathy function as relevant features of perceiving sound. I have obtained comprehensive results showing that listeners process acousmatic sound differently from sound that is attached to its visual source, and that this can have multifaceted mental consequences. These observations have several implications for research into the mental state of the listeners. My work has revealed how these effects might look like by providing a framework of both negative and positive examples with regard to a manipulated perception. Finally, my study has offered an approach to meaning-making within perception, which happens in a space that I have argued as a merging area of the real- and fictional world. Primarily, most personal experiences of establishing meaning can be ascribed to how characters respond in the film.

This thesis has attempted to contribute towards enhancing our understanding of the manipulative powers of acousmatic sound in terms of mental processes at work. While engaging with fiction in form of contemporary films, the present study has highlighted the importance of the interconnectedness of sound and our minds with regard to our belief-system and values. In the future, it would be extremely interesting to conduct an additional long-term study in order to be able to identify long-term effects that sound in fiction have on the audience. My present findings seek to help to establish a framework for a further investigation of consequences which could help to prove or disprove the impairments I have analysed with regard to mental health issues.

Bibliography

Primary Sources

Iron Man. Dir. Jon Favreau. Perf. Robert Downey Jr., Terrence Howard, Jeff Bridges, and Gwyneth Paltrow. Paramount Pictures, 2008. Netflix.

Lost in Translation. Dir. Sofia Coppola. Perf. Bill Murray and Scarlett Johansson. Focus Features, 2003. DVD.

Phone Booth. Dir. Joel Schumacher. Perf. Colin Farrell, Kiefer Sutherland, Katie Holmes, and Forest Whitaker. 20th Century Fox, 2002. Online Access via Amazon Video.

Stranger Than Fiction. Dir. Marc Forster. Perf. Will Ferrell, Emma Thompson, and Dustin Hoffman. Sony Pictures Home Entertainment, 2007. Online Access via Amazon Video.

Secondary Sources

Alleyne, Mike. "Sounds Reel: Tracking the Cultural History of Film Sound Technology." *Sound and Music in Film and Visual Media: A Critical Overview*. Ed. Graeme Harper. New York: Bloomsbury Publishing Plc, 2009. 15-41.

Atkinson, Simon. "Interpretation and Musical Signification in Acousmatic Listening." *Organised Sound* 12.2 (2007): 113-122.

Bal, Mieke. *Narratology: Introduction to the Theory of Narrative*. 3rd ed. Toronto: University of Toronto Press, 2009.

Balazs, Bela. "Theory of the Film: Sound." *Film Sound: Theory and Practice*. Ed. Elisabeth Weis and John Belton. New York: Columbia University Press, 1985. 116-125.

Belton, John, and Elisabeth Weis. *Film Sound: Theory and Practice*. New York: Columbia University Press, 1985.

Bijsterveld, Karin. "Beyond Echoic Memory: Introduction to the Special Issue on Auditory History." *The Public Historian* 37.4 (2015): 7-13.

Bordwell, David. *Narration in the Fiction Film*. Wisconsin: The University of Wisconsin Press, 1985.

Bordwell, David, and Kristin Thompson. *Film Art: An Introduction*. New York: McGraw-Hill, 2010.

Bull, Michael, and Les Back. *The Auditory Culture Reader*. New York: Berg, 2003.

Chion, Michel. *Audio-Vision: Sound on Screen*. 2nd ed. Ed. and transl. Claudia Gorbman. New York: Columbia University Press, 2019.

---. *Sound: An Acoulogical Treatise*. Trans. and Introd. James A. Steintrager. Durham and London: Duke University Press, 2016.

Chion, Michel. *The Voice in Cinema*. Ed. and transl. Claudia Gorbman. New York: Columbia University Press, 1999.

Clarke, Eric F. *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. Oxford: Oxford University Press, 2005.

---. "The Impact of Recording on Listening." *Twentieth-Century Music* 4.1 (2007): 47-70.

---. "What's Going On: Music, Psychology, and Ecological Theory." *The Cultural Study of Music: A Critical Introduction*. 2nd ed. Ed. Martin Clayton, Trevor Herbert and Richard Middleton. New York: Routledge, 2012. 333-342.

Colman, Andrew M. *A Dictionary of Psychology*. 4th ed. New York: Oxford University Press, 2015.

Cowan, Nelson, and István Winkler. "From Sensory to Long-Term Memory: Evidence from Auditory Memory Reactivation Studies." *Experimental Psychology* (2005): 3-20.

Cuddy-Keane, Melba. "Virginia Woolf, Sound Technologies, and the New Aurality." *Virginia Woolf in the Age of Mechanical Reproduction*. Ed. Pamela L. Caughie. New York: Garland Publishing, Inc., 2000. 69-96.

---. "Modernist Soundscapes and the Intelligent Ear." *A Companion to Narrative Theory*. Ed. James Phelan and Peter J. Rabinowitz. Oxford: Blackwell Publishing Ltd, 2005. 382-398.

Currie, Gregory. *Image and Mind: Film, Philosophy, and Cognitive Science*. New York: Cambridge University Press, 2008.

Deutsche Grammophon - DG. „Interview with Max Richter on The Blue Notebooks (2018).” *YouTube*, April 25, 2018. Accessed August 8, 2021.
<https://www.youtube.com/watch?v=Jn2oGZT62Ps&t=1s>.

Doane, Mary Ann. "The Voice in the Cinema: The Articulation of Body and Space." *Film Sound: Theory and Practice*. Ed. Elisabeth Weis and John Belton. New York: Columbia University Press, 1985. 162-176.

Telekom Electronic Beats. "Eliane Radigue: An Interview." Deutsche Telekom AG, 2012. Accessed January 6, 2022. <https://www.electronicbeats.net/eliane-radigue-an-interview>.

Finnegan, Ruth. "Music, Experience, and the Anthropology of Emotion." *The Cultural Study of Music: A Critical Introduction*. 2nd ed. Ed. Martin Clayton, Trevor Herbert and Richard Middleton. New York: Routledge, 2012. 353-363.

Fletcher, Harvey, and Munson, Wilden A. "Loudness, Its Definition, Measurement and Calculation." *The Journal of the Acoustical Society of America* 5 (1933): 82-108.

Flueckiger, Barbara. "Sound Effects: Strategies for Sound Effects in Film." *Sound and Music in Film and Visual Media: A Critical Overview*. Ed. Graeme Harper. New York: Bloomsbury Publishing Plc, 2009. 151-179.

Frijda, Nico H. "Emotion experience." *Cognition and Emotion* 19.4 (2005): 473-497.

--- and Mesquita, Batja. "Beliefs through emotions." *Emotions and Beliefs: How Feelings Influence Thoughts*. Ed. Nico H. Frijda, Antony S.R. Manstead and Sacha Bem. Cambridge: Cambridge University Press, 2000.

Gatt, Michael. "Memory, Expectation and the Temporal Flux of Acousmatic Music." *Organised Sound: An International Journal of Music Technology* 25.2 (2020): 205-213.

Gaut, Berys. *A Philosophy of Cinematic Art*. Cambridge: Cambridge University Press, 2010.

Gold, Ben, Morgan, Nelson, and Ellis, Dan. *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*. 2nd ed. New Jersey: John Wiley & Sons, Inc., 2011.

Goldsmith, Mike. *Sound: A Very Short Introduction*. Oxford: Oxford University Press, 2015.

Gresham-Lancaster, Scot, and Sinclair, Peter. "Sonification and Acoustic Environments." *Leonardo Music Journal* 22 (2012): 67-71.

Grodal, Torben. "Emotions, Cognitions, and Narrative Patterns in Film." *Passionate Views: Film, Cognition, and Emotion*. Ed. Carl Plantinga and Greg M. Smith. London: The Johns Hopkins University Press, 1999. 127-145.

Gunther, Leon. *The Physics of Music and Color: Sound and Light*. 2nd ed. Medford: Springer Nature Switzerland AG, 2019.

Hodgkinson, Tim. „An Interview with Pierre Schaeffer – Pioneer of Musique concrète." *ReR Quarterly Magazine* 2.1 (1987): 1-9.

Holman, Tomlinson. *Sound for Film and Television*. 2nd ed. Burlington: Elsevier Science & Technology, 2014.

Ihde, Don. *Listening and Voice: Phenomenologies of Sound*. 2nd ed. Albany: State University of New York Press, 2007.

Kane, Brian. *Sound Unseen: Acousmatic Sound in Theory and Practice*. Oxford: Oxford University Press, 2016.

James, William. *The Principles of Psychology: In Two Volumes*. New York: Henry Holt and Co, 1892.

Kendall, Gary S. "Meaning in Electroacoustic Music and the Everyday Mind". *Organised Sound: An International Journal of Music Technology* 15.1 (2010): 63-74.

Kozloff, Sarah. *Invisible Storytellers: Voice-Over Narration in American Fiction Film*. Berkeley: University of California Press, 1988.

Leudar, Ivan, and Philip Thomas. *Voices of Reason, Voices of Insanity: Studies of Verbal Hallucinations*. London: Routledge, 2000.

Lewis, A. 'LEXICON' – *Behind the Curtain*. eContact! 2014. 18 May 2021.
<https://econtact.ca/15_4/lewis_lexicon.html>

Lewis, Andrew. "Past and present perceptions surrounding mission education: A historical-metabetical overview." DEd diss. University of Stellenbosch, 1999.

- Lindborg, PerMagnus. "Psychoacoustic, physical, and perceptual features of restaurants: A field survey in Singapore." *Applied Acoustics* 92 (2015): 47-60.
- Lindsay, R. Bruce. "Lindsay's Wheel of Acoustics." *Journal of the Acoustical Society of America* 36 (1964): 2242.
- Maeder, Christoph. "Analysing Sounds." *The SAGE Handbook of Qualitative Data Analysis*. Ed. Uwe Flick. London: SAGE Publications, 2013. 424-434.
- Menner, Albert L. *A Pocket Guide to the Ear*. New York: Thieme, 2003.
- Metz, Christian. *Film Language: A Semiotics of the Cinema*. Trans. Michael Taylor. Chicago: The University of Chicago Press, 1991.
- . "Aural Objects." *Yale French Studies* 60 (1980): 24-32.
- Mikos, Lothar. "Analysis of Film." *The SAGE Handbook of Qualitative Data Analysis*. Ed. Uwe Flick. London: SAGE Publications, 2013. 409-423.
- Murch, Walter. Foreword. *Audio-Vision: Sound on Screen*. By Michel Chion. 2nd ed. Ed. and transl. Claudia Gorbman. New York: Columbia University Press, 2019. 7-20.
- Neisser, Ulric. *Cognition and Reality: Principles and Implications of Cognitive Psychology*. New York: W.H. Freeman and Company, 1976.
- Percheron, Daniel, and Marcia Butzel. "Sound in Cinema and its Relationship to Image and Diegesis." *Yale French Studies* 60 (1980): 16-23.
- Schaeffer, Pierre. *Treatise on Musical Objects: An Essay Across Disciplines*. Trans. Christine North and John Dack. Introd. Christine North. California: University of California Press, 2017.
- Schnell, Jim. "Effective Listening: More Than Just Hearing." *Viewpoints* 120 (1995): 1-7.
- Schumacher Ratti, Federico and Fuentes Bravo, Claudio. "Space-Emotion in Acousmatic Music." *Organised Sound: An International Journal of Music Technology* 22.3 (2017): 394-405.
- Sergi, Gianluca. *The Dolby era: Film sound in contemporary Hollywood*. Manchester: Manchester University Press, 2004.
- Snyder, Bob. *Music and Memory: An Introduction*. Cambridge, MA: MIT Press, 2000.
- Soddell, Thembi. „The Acousmatic Gap as a Flexile Path to Self-Understanding: A case for experiential listening." *Organised Sound: An International Journal of Music Technology* 25.3 (2020): 344-352.
- Sonnenschein, David. *Sound Design: The Expressive Power of Music, Voice, and Sound Effects in Cinema*. Los Angeles: McNaughton & Gunn, 2001.
- Tan, Ed S.H. and Nico H. Frijda. "Sentiment in Film Viewing." *Passionate Views: Film, Cognition, and Emotion*. Ed. Carl Plantinga and Greg M. Smith. London: The Johns Hopkins University Press, 1999. 48-64.

Tankel, Jonathan D. "The Impact of The Jazz Singer on the Conversion to Sound." *Journal of the University Film Association* 30 (1978): 21-25.

Windsor, Luke. "Through and Around the Acousmatic: The Interpretation of Electroacoustic Sounds." *Music, Electronic Media and Culture*. Ed. Simon Emmerson. New York: Routledge, 2016. 7-35.

Winters, Ben. "THE NON-DIEGETIC FALLACY: FILM, MUSIC, AND NARRATIVE SPACE." *Music & Letters* 91.2 (2010): 224-244.

Worthington, Debra L. and Margaret E. Fitch-Hauser. *Listening: Processes, Functions, and Competency*. 2nd ed. New York: Routledge, 2018.

Appendix A: Reviews

Focus Group 1 consisting of Person A + Person B (females) + me

Focus Group 2 consisting of Person C + Person D (males) + me

Conversation in Focus Group 1

Me: I asked the two of you to watch *Phone Booth* and to pay special attention to sound in the film. Can you tell me how you perceived the omniscient narrator speaking at the beginning?¹³

B: Well, as a normal narrator?! I did not really think about him, to be honest...

Me: Okay, and you, A?

A: Same. I don't remember that I perceived the voice in any special sense. He's just there... in the background, giving the viewers facts on New York City life and the telephone connections.

B: Yeah, he is introducing the story. Like it's often done in films.

Me: Alright. Then let's move to the unknown caller's voice. I don't want to give too much away, so can you please portray how you perceived this voice?

B: Puh, difficult question. Hearing the voice made me feel very uncomfortable...

Me: You can be completely honest with me. The more I know about your emotional experience, the better. Why did it make you feel uncomfortable?

B: Well, you know, my parents are divorced. This is a long time ago, but before that happened, I was used to endless arguments... and harsh conversations at home. Often, when my parents thought I would not listen, my father became very loud with my mother. I have heard him do this many times, and it frightened me! You know, I couldn't quite place what this meant as a child. But... I am very sensitive to loud voices ever since in a negative way.

¹³ The exact reference can be found in the body text of my thesis. This applies to all following examples.

Me: Thank you for bringing this up! So, does this mean that the louder a voice gets, the more intimidated you feel? Is that the right word?

B: Totally. That's a good way to describe it. Because I'm not completely anxious or so simply by hearing a loud voice. I think it's more the fact when it's in the context of an argument, so when someone is personally attacked.

Me: To get back to the unknown caller's voice in *Phone Booth*. Can you summarise your perception of that voice?

B: Erm, well yeah, basically I just didn't feel comfortable watching this film. I was really tense throughout the film and the scenes in which the caller raised his voice against the main character were hard for me to watch. I think it was the threatening that frightened me most.

Me: Do you think seeing the person who calls Stu would have changed your perception? Your emotions?

B: Good question... I think so, yes. I mean I don't think it would have taken away my fear completely. But I think it would have made me calmer in a way. It felt very unfamiliar not to be able to see the person speaking for almost the entire time.

Me: Thank you, B! What do you think, A? Do you agree? Or how did you perceive the voice?

A: First of all, I have to say that I find B's perspective really interesting! This is not similar to how I perceived the voice, but I find your thoughts extremely insightful, B. Thank you for sharing that. As for me, I have to say that I don't have any negative connotations with loud voices, or voices in general.

Me: So do you feel you were unbiased when watching the film?

A: Totally. I rather found the voice very interesting. You know, I listen to a lot of podcasts and stuff like that. So maybe that helped to not be afraid of the voice, could be. But I was very much interested in the voice itself, it sounded so calm and mysterious and very intelligent to me as well. That was an interesting mix of a voice that made me want to know more about the character behind it! That's why I agree with B here, it was strange not to see the person earlier. And when I think about what B said, that is probably also in general what makes the experience a bit uncanny!?

Me: You mean not being able to see the voice's visual source?

A: Yes! Because of this there's lots of room for speculations... at least I was very curious about who this person might be, how it looked like from the outside and the inside. What did this person think? These were the questions that puzzled me.

Me: Thank you for your impressions, they are very insightful! One more question concerning Stu, do you empathise with him?

B: Absolutely! I remember the first few minutes of the film, where he was presented as an arrogant busybody... but, although obviously he was in a way forced to, he undergoes a positive change in the course of the film.

A: I agree. And also, I don't even want to imagine being in his position for real! It would freak me out! And although Stu probably needed a little learning – this is a tough one! I didn't feel he deserved to be in agony. I felt very sorry for him.

B: Oh, no! Me neither! The anxiety already felt very real.

Me: Thank you! Let us now have a closer look at your perception of *Stranger Than Fiction*. You will have noticed that again, there is an acousmatic¹⁴ voice in the foreground. How did you perceive this voice?

A: Well, at first it was similar to how I perceived the voice at the beginning of *Phone Booth*. It seems to be a narrator outside of the story world that is introducing us to the story and telling us some facts. But it soon became clear that they are not very similar actually.

Me: How?

A: Because the voice seems to know every little detail about Harold. That was weird.

B: I agree! And especially as I realised that Harold is able to hear the voice, I somehow perceived it differently. I was confused... The voice which I thought was nondiegetic in fact wasn't. Harold could hear the voice and I found it very hard to decide which perspective to choose.

Me: Can you elaborate on that?

¹⁴ I have explained the term to my interlocutors during our conversation.

B: Well, what I mean is that we only have a voice to trust, at least unless the character of Karen Eiffel is introduced to us. And so I found it very hard to decide whether I wanted to team up with the voice, which I thought did not belong to the story world, or with Harold, who was able to hear the voice that drove him mad.

Me: I see your point. Did any of you at a time imagine to be in Harold's position?

A: For sure! I can only speak for myself, but I have never seen a film, in which a character is hearing a voice, before. I asked myself a few times what I would do if I were Harold and how I would behave. Probably I would try the same things Harold did... trying to reassure that the voice is audible, seeing a psychologist... things like that. I don't know, it must be really tough.

B: I'm with you. Especially later, when Harold is already really desperate and has an emotional outburst in his room, I felt so sorry for him. I imagined what it must be like to have a voice in your head telling you that you're going to die very soon? How terrifying is this?! I would be so frightened and desperate, too.

A: Right. But what I have to say is that my view changed a bit after I got to see Karen Eiffel and realised that it was her voice. I don't know, I felt that I rather adopted the point of view from the beginning, where I clearly felt to be on the same level with the narrator.

Me: Thank you so much!

Conversation in Focus Group 2

Me: C and D, I have asked you to watch the film *Iron Man* and to focus on your perception of sound. Concerning the opening scene, I have asked you to watch it without sound first, and with sound afterwards. Can you share your thoughts on how you perceived the images without accompanying sound?

C: Sure. Well, I have heard of the film before, but as you know I am not the biggest Marvel fan and I have never watched it before. So, when I saw this first scene, I saw a sequence of dangerous shots... you could read that these armoured cars are in Afghanistan, so my first connotation was war! I expected to see a war scene next.

D: I agree! These images made me feel very uncomfortable. This “car” convoy seemed threatening, and I was also not expecting something positive to happen, especially because of the machine guns on top.

Me: Alright, interesting to see that you agree here. Accordingly, I asked you to watch the exact same scene with the original sound. How did you perceive the scene? Did you perceive it differently than before?

D: Yes! Honestly the music makes all the difference here. I admit I was a bit confused because it was not what I expected... I don’t know why but I feel that this kind of music made the scene even more spectacular.

Me: Do you know the song in the background?

D: No...

C: It’s ACDC!

Me: Correct.

C: I agree with D that the music not quite matches the images... or, at least that this is not the kind of music you would expect to be in the background of such a scene. I feel almost as if this song in a sense ridicules the scene, which makes it less... I don’t know, plausible?! It seems to me that it is less authentic with a rock song in the background. But what I have to say is that it underlines the strong will of the military to start an attack.

D: Yeah, I’m with you there.

Me: Okay, thank you. Let's stick with nondiegetic music for a moment. There is another scene which I asked you to pay attention to, namely the scene in which Pepper is on her secret mission to save important data from Tony's computer in his office. Can you just briefly report how you perceived this scene?

D: Yes. Let me say first that I loved Pepper as a character! She is a likeable person you empathise with and so did I do in this scene. This situation is of course a big deal for her and very risky... you can literally feel the tension.

C: Absolutely! The tension is very well transferred to the audience.

Me: None of you mentioned the music in this scene yet. Can you share your impression of the music?

C: True. Probably... the very calm background music contributes to the creation of suspense in the scene. It was instrumental music, wasn't it?

D: Yes, it was. And I agree, the music has a great effect here. It makes the scene much more dramatic, and the tension was hardly bearable!

C: True, this scene definitely wouldn't be the same without the accompanying music.

Me: Thank you. Different question: Which effect did J.A.R.V.I.S.'s voice have on you?

C: Erm. J.A.R.V.I.S. is smart and funny, but I think that it is rather contributing to the viewer's entertainment than having a specific effect on the viewer. At least it didn't have on me...

Me: So, there is no emotional effect?

D: No, I agree with B. J.A.R.V.I.S. did not cause any emotional feelings.

Me: Regarding all other films I have asked you to watch, which voice did have an emotional effect on you?

D: Good question. I guess my immediate answer would be... Bob's wife in *Lost in Translation*.

Me: Okay and why is that?

D: Well, I feel that although the encounter between Charlotte and Bob brings the audience joy and light-heartedness at times, the film also has this sad and bitter taste that constantly resonates. I think that this is very much reflected in the phone calls, especially in Bob calls with his wife. Hearing Lydia's voice made me feel sorry for her. Clearly, their marriage isn't working anymore, but Lydia somehow tries to maintain it, if only with unimportant issues. I don't know, their situation made me feel melancholic...

Me: Why melancholic?

D: I think mainly because I saw myself and my partner before me, wondering what life would be like in separate spaces, just hearing each other's voice via phone calls... but living parallel to each other in fact. Hearing what Lydia has to say made me think about my own relationship a lot, and also about the fear that we could end up like this one day. I know, this is deep haha, I'm sorry.

Me: No no, I see your point. Thank you for your honesty! Thank you both for sharing your thoughts with me.

Appendix B: Abstract

English version

This master's thesis investigates the phenomenon of acousmatic sound in light of its role in fictional films as a means of perceptual manipulation. It provides insight into the innovative ways in which sound that is detached from its visual source can have an impact on perception, and which consequences this has in respect to cognitive processes. With the help of a qualitative data analysis according to Lothar Mikos and Christoph Maeder, four films of different genres, *Lost in Translation* (2003), directed by Sofia Coppola, *Iron Man* (2008), directed by Jon Favreau, *Phone Booth* (2002), directed by Joel Schumacher, and *Stranger Than Fiction* (2007), directed by Marc Forster, are analysed. As these research objects offer a valuable insight into how acousmatic sound is used in film, this thesis investigates in which way acousmatic sound, with special emphasis on music and the disembodied voice, constitutes a means of manipulation in film. Furthermore, I will examine the audience's perception by means of multiple strategies of processing on a cognitive level and scrutinise possible repercussions. Offering an interdisciplinary approach, the theoretical framework will include important elements drawn from Sound Studies and Film Theory, combined with Cognitive Theory, which serve to examine how the analysed film scenes correspond to my propositions. My central claim in this thesis is that through the creation of acousmatic experiences in film, the conditions of listening are not only altered significantly, but that a manipulation of the affected characters' and the spectators' perception is elicited, which can lead to multifaceted mental consequences. The objective of this master's thesis is to contribute to the closure of the research gap in the field of the processing of acousmatic sound in the cognitive sense. Moreover, it attempts to demonstrate how the various manifestations of the phenomenon of sound shed light on the mental dimension of emotions and how this can have serious consequences beyond the fictional world.

Deutsche Version

Diese Masterarbeit untersucht das Phänomen des akusmatischen Klangs als Mittel zur Manipulation der Wahrnehmung im Hinblick auf seine Rolle in fiktionalen Filmen. Sie gewährt einen Einblick in innovative Arten, in denen Klang, der von seiner visuellen Quelle losgelöst ist, einen Einfluss auf die Wahrnehmung haben kann, und welche Konsequenzen dies in Bezug auf kognitive Prozesse hat. Vier Filme verschiedener Genres, *Lost in Translation* (2003) unter der Regie von Sofia Coppola, *Iron Man* (2008) von Jon Favreau, *Phone Booth* (2002) von Joel Schumacher, und *Stranger Than Fiction* (2007) von Marc Forster, werden mithilfe einer qualitativen Datenanalyse untersucht. Da diese Forschungsobjekte einen wertvollen Einblick in die Verwendung von Klang in Filmen bieten, durchleuchtet diese Arbeit die Arten in denen akusmatischer Klang, mit besonderem Schwerpunkt auf Musik und körperloser Stimme, eine Möglichkeit der Manipulation darstellt. In interdisziplinärer Ausrichtung umfasst der theoretische Rahmen wichtige Elemente aus den Klangstudien und der Filmtheorie kombiniert mit Aspekten der Erkenntnistheorie, welche zu der Untersuchung beitragen, inwiefern die analysierten Filmszenen meinen Behauptungen entsprechen. Die zentrale These dieser Arbeit ist, dass durch die Erzeugung von akusmatischen Erfahrungen in Filmen nicht nur die Bedingungen des Hörens entscheidend verändert werden, sondern auch eine Manipulation der Wahrnehmung der betroffenen Charaktere und der Zuschauer erzeugt wird. Diese Erlebnisse können vielfältige geistige Folgen nach sich ziehen. Das Ziel der Arbeit ist es, die Forschungslücke im Bereich der kognitiven Verarbeitung von akusmatischem Klang zu schließen. Des Weiteren versucht sie aufzuzeigen, wie die verschiedenen Erscheinungsformen des Klangphänomens in Filmen Aufschluss über die geistige Dimension von Empfindungen geben kann und welche gravierenden Folgen jenseits der fiktiven Welt entstehen können.

Appendix C: Anti-Plagiarism Statement

I hereby declare that this research paper was written by me. I have used my own ideas, except for quotations from published or unpublished sources. All of these quotations are clearly indicated and referenced according to the academic “Style Sheet for Literary and Cultural Studies” of the English and American Department of Vienna.

15th March 2022