How to engineer a mood
A study of sound in audiovisual contexts

Hur konstruerar man en stämning
En studie av ljud i audiovisuella sammanhang

Nynke Hoekstra
Abstract

“How to engineer a mood - A study of sound in audiovisual contexts” is a study on the perception of sound in audiovisual contexts, i.e. sound and the image combined. There seems to be a consensus among researchers in the field on the fact that music influences the interpretation of film. However, there does not seem to be a specific consensus in regards to other sounds than music, such as the human voice or sound effects.

Guided by a set of theories, an audiovisual text will be analyzed, in part through results of a quasi-experimental pilot survey and partly by the author herself. This paper would like to contribute in making the relationship of different sounds in conjunction with the image become slightly clearer. The outcome will help to show how a mood can be created with sounds. Although there is a limited number of theoretical models that have been developed in this field of research, the theories that are introduced will hopefully help to create a better understanding of sound in audiovisual texts. Finally the theories and results will come together in this thesis; this will lead to an analysis and produce ideas for further research on how sound in audiovisual contexts influence our emotions.

As mentioned before there has not been much research in this field, hopefully this thesis will be able to provide some clues and could serve as a first step for further research to find more definite answers on how emotions work in a audiovisual context. However, many questions will remain unanswered, and further research is needed to answer them.

Keywords: Sound, audiovisual, sound design, musical soundtrack, sound perception.
Sammanfattning

Denna c-uppsats handlar om uppfattning och varseblivning av ljud i det audiovisuella sammanhanget, dvs. ljud i kombination med bild. Det verkar finnas en enighet om faktumet att musik påverkar tolkningen av audiovisuella texter. Det verkar dock inte finnas en sådan enighet när det gäller andra ljud än musik, som till exempel rösten eller ljud effekter.

Med hjälp av teorier kommer en audiovisuell text att analyseras, dels genom att kolla på resultaten av enkäten och delvis skall författaren analysera det utvalda filmklippet. Genom att läsa den här uppsatsen borde förhållandet mellan olika ljud och bild bli lite tydligare och förhoppningsvis kan uppsatsen bidra vissa hur en stämning/emotion kan skapas med ljud. Även om det finns ett begränsat antal teorier som har utvecklats i detta undersökningsfält, de teorier som blir introducerade kommer förhoppningsvis att bidra till en bättre förståelse för ljud i audiovisuella texter. Slutligen ska teorier och resultaten komma tillsammans, detta kommer att leda till en analys samt idéer för framtida forskning om hur ljud i ett audiovisuellt sammanhang kan påverka våra känslor/emotioner.

Som sagt har det inte funnits mycket forskning i detta område, förhoppningsvis kommer denna uppsats ge några ledtrådar och kanske vara ett första steg för vidare forskning för att hitta mer definitiva svar på hur känslorna/emotioner fungerar i ett audiovisuellt sammanhang. Dock kommer många frågor förbli obesvarade, och ytterligare forskning behövs för att besvara dem.

Nyckelord: Ljud, audiovisuell, ljud design, ljudperception.
## Table of content

1. Introduction
   - 1.1 Background ................................................................. 6
   - 1.2 Purpose ......................................................................... 6
   - 1.2.1 Research Questions ...................................................... 8
   - 1.3 Disposition ................................................................. 8
2. Theory .................................................................................. 10
   - 2.1 History and Technology .................................................. 10
   - 2.2 Film studies .................................................................... 12
     - 2.2.1 Added value .............................................................. 12
     - 2.2.2 Synchresis .............................................................. 14
     - 2.2.3 Realism .................................................................... 14
     - 2.2.4 Immersion .............................................................. 16
   - 2.3 Psychology ....................................................................... 16
     - 2.3.1 Sound perception & hearing ......................................... 16
     - 2.3.2 Listening modes ........................................................... 18
     - 2.3.3 Gestalt principles ......................................................... 19
     - 2.3.4 Emotion sound ........................................................... 20
     - 2.3.5 Perception of time ......................................................... 20
2.4 Music in film theory ............................................................ 21
   - 2.4.1 Categories of relation between film music and motion pictures ......................................................... 21
   - 2.4.2 Aesthetics and adjustments .............................................. 22
   - 2.4.3 Associationism ............................................................. 23
   - 2.4.4 Multi-level congruence Associationist framework .............. 24
   - 2.4.5 Empirical Research ....................................................... 25
   - 2.4.6 Early research ............................................................. 25
   - 2.4.7 Qualitative methods ..................................................... 26
   - 2.4.8 Cognition and incidental learning .................................... 28
   - 2.4.9 Physiology ................................................................. 29
   - 2.4.10 Sound track as a whole ............................................... 30
   - 2.4.11 The role of music in cinema ........................................... 31
   - 2.4.12 Music functions ........................................................... 32
2.5 Other sounds ........................................................................ 33
   - 2.5.1 Sound effects ............................................................... 33
   - 2.5.2 Silence ........................................................................ 34
   - 2.5.3 Point of synchronization ............................................... 35
1. Introduction

1.1 Background

A film without sound remains a film, a film without the image however is no longer a film

— Michel Chion

The sense of hearing is the first sense to develop in the womb, (Sonnenschein 2001, p. 72) and unlike our eyes which we can easily close with our eyelids, we cannot close our ears; at least not without any external help from for instance our hands. Sound always surrounds us; we can hear around corners and perceive what is happening behind our backs. We even hear in our sleep, when most of our senses are turned off.

Our perception of the world relies a great deal upon sound. You can try to go about your regular day with earplugs and see how your world differs from the normal world full of sound. No voices, no music, and maybe more dangerously no traffic noise. If this practical experiment seems too drastic, try watching a TV show or a movie without sound and see what you think. It probably isn’t all that fun, right?

Even with all this experiential wisdom, sound is still viewed as an add-on in the film experience from both the ontological and historical points of view (Chion, 1994 p. 143). The reason for this is probably that film has been “silent” without synchronous sound for several decades before synchronized sound made its way into cinema.

But since there have been so many technical and scientific developments since the early days of film, can we really still consider sound as an add-on in this day and age where movies frequently rely heavily on sound, both musically and vocally?

It can be concluded that sound from a historical perspective is an add-on, because sound, at least synchronized sound, came later. The technology had not yet been sufficiently developed for a useable display of sound in its various aspects. More on this will be presented in the theory section. The fact is however that music was always present; can it therefore really be seen as an add-on?
This is what Walter Murch, a famous sound designer who won an Oscar for sound for his work on *Apocalypse now* (1979) and was awarded Oscars for editing and sound for *The English Patient* (1996), has to say about the relation of sound and the image:

"Whatever virtues sound brings to the film are largely perceived and appreciated by the audience in visual terms - the better the sound, the better the image." (Murch, as cited in Chandler, 2009)

Murch’s statement would indicate that, even when the sound is the source of the improvement, the audience experiences a visual difference. It is quite remarkable that a celebrated sound designer as Murch concludes that the sound is still subservient to the film. With this fact Murch confirms to some extent the argument of Chion (1994), as put in the quote that we started this chapter with, namely: “A film without sound remains a film, a film without the image however is no longer a film”. It seems to be an undeniable fact that the sound always is submissive to the image and maybe this implies that the sound designer therefore always is subservient to the director.

Walter Murch, because of his works as a sound designer is well versed with sound in audiovisual formats. Murch was the first person to receive the title sound designer, and he is responsible for quite a change in film sound. The recognition Murch received by the nominations and Oscar award may be considered as quite an achievement in the field (Stam, 2000).

So what is a sound designer? Well, sound designer is a profession that first came to be after the so-called second sound revolution in Hollywood. This revolution started during the crisis Hollywood was experiencing starting around the late 1960’s. Less people went to the theaters and smaller amounts of money came in. The concern about declining audiences and revenue lead to the improvement of the sound systems. In those days George Lucas created the THX system, which was a great improvement on the previous sound system. Movies like *The conversation* (1974) and *Apocalypse now* (1979) which Murch worked on focused greatly on the use of sound throughout the films, something that wasn’t all that common during those days. More and more sounds would be used in film from that point in time (Stam, 2000).

Design in the context of sound design refers to art, and thus goes far beyond the use of technical capabilities and experience. Sound design isn’t just editing and mixing the sounds, it is an art. The standard of sound we have in today’s films is thanks to Walter Murch, George Lucas and other modern sound pioneers from that generation. Although it wasn’t just due to the improved sound that people came back to the theaters, Lucas and others like him proved that the pioneer work
was taking place in many different areas. Lucas amongst others has played an important part in this revolution of the development of film sound (Stam, 2000).

Even with the fact that sound effects are very common in today’s cinema, it has not really formed a specific field of study on it. Film music on the other hand has been studied quite heavily (Cohen, 2001) and sometimes with remarkable outcomes, however, the question is: where do sound effects fit in to these types of sound studies? Since there aren’t any specific theories on sound effects I will try to implement these musical and other sound theories and see how they hold up.

1.2 Purpose

Over the years many disciplines such as psychology, film studies and music theory have looked at sound in audiovisual texts. However, all this information is in different places, such as books, articles etc. I would like to create a theoretical framework for sound in film where the information from all the different fields of study comes together in one place. This theoretical framework will be used to answer questions about sound.

There will be a short film clip that will be analyzed by looking at the material gathered in the theoretical frame work. The final goal is to see if a mood can be engineered with sound, and if so how one goes about creating a mood.

1.2.1 Research Questions

Above was described what the purpose of this essay is, namely to investigate if a mood can be engineered with sound, and if so how one goes about creating a mood. However we cannot answer this question without having to formulate certain sub-questions. The following questions were used to operationalize the main question:

1. How is the clip perceived in regards to sound in film?
2. What functions does sound serve in the clip?
3. How important is the sound in the clip?
4. Does the sound in the clip evoke emotions?
In question 1, we want to see how perception can differ under certain circumstances.

Question 2 is about functions of sound, it seems important to know what functions the sound has in order to be able to know how to create moods with the sound.

In question 3, the assumption is that we’re be able to demonstrate that sound in audiovisual texts in general, specifically film is significantly more important than often is assumed.

Question 4 is a very important question that will help answer the main question in regards to the emotion, if a mood is created through sound we can probably find out how a mood is created through sound.

1.3 Disposition

In the first chapter the topic is explained and introduced furthermore the research questions are revealed. The second chapter introduces different theories and methods used in previous research. Chapter three is about the method that is used in the analysis section of the thesis; there are for instance sections on the survey method, validity and reliability, and the audiovisual observation method. The fourth chapter shows the results that have been obtained in through the survey and shot breakdown.

In chapter 5 there is an analysis of an audiovisual text, i.e. the clip. Chapter 6 reveals the conclusion, here the research questions will be answered, and we will look at the answer to the question on “can a mood be engineered by sound”. The last chapter with significant information is chapter 7 which is about future research that could be performed based on the knowledge that was gained in this research. After these chapters you will be able to find a list of references, and three appendixes.
2. Theory

Like any field of study the research of sound in audiovisual texts has led to the development of various theories. This chapter is about different theories that have been used to explain sound in an audiovisual environment. Some of the theories provide a framework for understanding of sound in film, while others are more directly applicable in terms of analyzing an audiovisual text, in this case a video clip.

This chapter on theory begins with the history of synchronized sound, and how technology played a major factor in how sound eventually made its way into cinema. After that we move on to some theories in the field of film studies to explain certain phenomena that exist in this field of study.

Following the short history part will be the psychology of sound; this section mainly deals with perception. Perception is fundamentally a subjective sensation, therefore it is important to conduct field research on perception. It is also important to see whether similar results can be found in this perception. Later in the thesis it will be clarified what requirements will have to be met for a qualitative research. For this purpose different theoretical models will be studied. This approach also leads us to look at the way we as humans perceive sound. Subsequently we will take a look at music in film; this part will be quite extensive because music is probably the most researched type of sound in the field of audiovisual research. After music, as a specific form of sound in film we will take a look at other sound in the movies besides music.

2.1 History and Technology

People often refer to the early days of cinema as silent, which in one way is true. There was no synchronized sound that accompanied the image; however it was far from silent. In Japan there was the so-called Benshi, a narrator who explained the actions in the film. In the metropolitan theaters there were orchestras and in the smaller theaters there was the piano player who accompanied the image. (Elsaesser & Hagener, 2010)

Even though today’s audience probably doesn’t wonder why sound accompanies a film, it is not as obvious as one might think from a historical perspective. The average modern viewer thinks of sound as a natural part of film, and probably sees film without the sensation of sound as
incomplete. This feeling relates to our natural everyday experience of the world, where we also hear sounds that surrounds us, all the time. Earlier we saw that the development of sound in film has been quite problematic. This was mainly a technological issue in synchronizing the sound and the image (Ibid.).

Whether the expectations in the beginning were equivalent to still photography, where sound didn’t play any role, is only something one can only suspect. In the beginning of the moving image, it was just that; a set of “moving” images, or a rapid series of photographs. This is actually more or less what the kinetoscope was, more so than film, it presented moving pictures, which is a different experience from what we consider film to be today. Yet at first music was used to mask the sound of the projector. However, when the sound projectors became increasingly quieter, the music stayed, it had become a way of communicating the message. (Cohen, 2001)

Gradually, attempts were made to integrate auditory elements in addition to the visual aspects of a sensory experience of the film. Edison’s company created the kinetoscope, a predecessor of cinema to compliment his phonograph a forerunner of the record player. Edison said that the kinetoscope would do for the eye what the phonograph does for the ear. This shows that film initially the image was seen as a complement to sound, rather than the other way around. (Allen & Gomery, 1985)

The main limitation in the development of sound in film was of a technical nature. Edison’s system turned out to be flawed, synchronization turned out to be a rather large bump in the road. Many people would try to solve the problems with synchronization and amplification a decade later, without any luck (ibid.). When film without synchronized sound became popular, the view of the audiovisual relationship changed. Now the image became more important than the sound, which it remained even when synchronized sound eventually made its way into cinema. People therefore always refer to films in visual terms, e.g. watching a movie; no one ever says “let’s listen to a movie”.

Edison later announced that his laboratory had finally solved all of the issues that had existed in the system before. Four theaters in New York City got the Kinetophone installed. However, this did not last long, soon after the first showing the system was abandoned because of the problems with synchronization and amplification.

An important step in the development of synchronized sound was made by AT &T and Western Electrics, when they finally solved the problem after more than a decade of research. Together
they had developed and improved much of the technical gear such as loudspeakers, microphones and other technical devices (Ibid.).

AT & T was trying to make a profit from their new invention, therefore they decided to work together with Warner Bros. (1924) and later also with The Fox Film Cooperation. However, it would still take a while before sound cinema actually took off and became popular. In 1927 the moment had arrived, The Jazz singer (1927) and the accompanying vaudeville shorts became very popular. After this the demand for synchronized sound films became huge. All thanks to Warner bros. vitaphone system, Warner Bros. would become one of the big studios in Hollywood. The other studios would soon follow and started using synchronized sound in their films. Synchronized sound had been proven to be very lucrative and staying behind was not an option, with the growing demand for synchronized sound. After September of 1929 Hollywood would only produce these synchronized sound films, also called talkies. (Allen & Gomery, 1985)

2.2 Film studies

2.2.1 Added value

Added value is a theory within the phenomenon of audiovisual illusion, Michel Chion (1994) describes Added value as such:

"The expressive and informative value with which a sound enriches a given image so as to create the definite impression, in the immediate or remembered experience one has of it, that this information or expression "naturally" comes from what is seen, and is already contained in the image itself.” (Chion, 1994, p. 5)

This means that sound creates a value which enriches the image and creates a definite impression, and this can be an immediate or a remembered impression. The definite impression’s meaning comes naturally from the image itself. For example when you watch a video of three airplanes, and you would state: “Two airplanes are in front of the last airplane”, everybody would see it like that and the statement would seem redundant. Yet if you would say there are three small airplanes, everybody would see that and see this statement as redundant. The meaning can be seen in the image, but it is the sound that creates the specific meaning in the example. People often think because they can see that the two planes are in front of the other, that the sound in
In this case the human voice as a carrier of the information is unnecessary. However, the chances that the audience would come to that specific conclusion without the sound may be very slim (Chion, 1994).

In conclusion, the aspect of added in the term added value is why people feel that sound is not a necessity, and that its only function is to duplicate a meaning found in the image. The research questions specifically address presumptions such as these. However, in reality though sound gives its own meaning, either by itself or by incongruence of sound and the image. The latter one means that sound can give a meaning when the image and the sound are dissimilar. (Chion, 1994)

We can distinguish “value added by text” which roughly refers to the voice and “value added by music” which exists out of two types of music. These two concepts are explained more extensively below.

Value added by text as mentioned is mainly about the dialogue, the spoken voice in film. The voice in cinema is fairly dominant in today’s film; cinema usually privileges the voice over other sounds such as sound effects. When shooting a movie the voices are usually the main focus. Verbal expression is the most important aspect of the voice for film. Being able to fully understand every word is much more important than acoustical fidelity. Therefore sound in film is not just vococentric (focus on the voice), but also verbocentric (focus on the words). This emphasis on the word does however require a comprehension of the spoken language. The reason why we focus on the words in film is that humans in general naturally are verbocentric, we as human’s commonly first focus on voices in a given environment only after we know who is speaking and what is being said one might focus on other sounds. In conclusion we first seek meaning of the words, and when we are satisfied we can move on to other sound. Of course this only goes for languages we actually are able to understand (Chion, 1994).

Value added by music is divided in two different groups: empathetic and anempathetic music. These are the two ways for film music to create particular emotions in relation to the action seen on screen. Empathetic music participates in the feeling of the scene; it can take on the same rhythm, tone etc. This is culturally divined and can stand for sadness or happiness etc. (Chion, 1994)

Anempathetic music however is quite indifferent to a specific situation on screen; it progresses in a steady pace. This type of music intensifies emotion “by inscribing it on a cosmic background”
as Chion (1994) puts it. The anempathetic kind of music reinforces the character and the audience’s emotion, even though the music is quite indifferent to the action on screen.

There are cases where music is neither empathetic nor anempathetic, music can also have an abstract meaning or be there just for the sake of being there. One can tell whether this is the case if there is no exact emotional meaning (Chion, 1994).

### 2.2.2 Synchresis

Synchresis is, just as the previous section about Added value, a theory formulated by Chion (1994). The unusual term comes from the combination of the words synchronism and synthesis. By this Chion (1994, p. 63) refers to the “spontaneous and irresistible weld between a particular auditory phenomenon and visual phenomenon when they occur at the same time.”

Synchresis is the phenomenon that makes us believe that the sounds we hear actually belong to the object or person we see on screen. This phenomenon also makes sound effects and things as ADR possible, where Automatic Dialogue Replacement stands for dialogue that is recorded in a studio and replaces the dialogue recorded on set. Because of synchresis it is even possible to match different kinds of sounds to images that have nothing to do with each other, to a certain extent. For instance the example of sizzling bacon which has been used as a sound effect for rain, bacon in a frying pan has nothing to do with falling rain of course.

Nowadays sound effects in film are very normal, we have all been conditioned to perceive this as something obvious. Yet back in the 1920’s synchronized effects were experienced as astonishing. Being able to see an object on screen and at the exact same moment hear a noise made on screen was amazing to the audience at that time. Today we are completely used to this because we see it every day, on TV, in film or even in video games. This has led us to believe that this is a natural phenomenon, when it actually isn’t natural. (Chion, 1994)

### 2.2.3 Realism

Realism and natural sound may not always go together; a sound can be realistic but may not be natural. A natural recording would be a sound recording that is absolutely free of manipulation. The question is whether this is even possible. Is a recorded sound still the same reality? Can sounds be brought back in the exact way they were first heard? This is an interesting technical
and philosophical question, which remains outside the scope of this research. The conclusion, which entails that a realistic sound is just a sound that sounds real, will suffice in this context. An example for realistic versus cinematic realistic sound is a cinematic punch, a punch can seem realistic in film but it is actually far from natural (Chion, 1994).

Natural sounds are few and far between, even direct sound; the kind of sound that is recorded during filming is frequently manipulated in some way in post production. As mentioned before in connection to ADR, sound effects or other sounds are often added in the post production process. Usually sounds are also eliminated right on set by the placement of microphone or sound proofing and the like (Chion, 1994). This is mostly done to avoid unwanted noises, such as traffic sounds. The real or natural sounds are not always the best sounds to use in film. Recording techniques have moved into a specific aesthetic direction that may not be the most natural sounding. According to Walter Murch sounds can be meaningful on other levels than just reality; sound can be metaphoric for instance (Collins, 2008).

“This metaphoric use of sound is one of the most flexible and productive means of opening up a conceptual gap into which the fertile imagination of the audience will reflexively rush, eager (even if unconsciously so) to complete circles that are only suggested, to answers half posed” (Murch, cited in Collins 2008, p. 135) Therefore sound is as much an aesthetic choice as a reproduction of the imagined according to Murch. Reproduction and rendering are two notions that need to be distinguished. As mentioned before the audience sees sounds as truthful or appropriate, not because these sounds are exact reproductions of a real situation. What matters is what the sounds “render” i.e. to what extent they express the feelings associated with the particular situation. Again this is not a conscious action, at least not fully (Chion, 1994).

According to Chion (1994) rendering involves several sensory channels, not one in particular. When we perceive a cinematic punch or hear a body falling in film, the sound is much louder than it would be in a real life situation; the volume has to render more than in real life. The weight, violence and pain all need to be represented by the punch or fall in film. Rendering such events thus goes far beyond creating a sense of realism. Most of our sensory experiences rely on more than just what is perceived by one of our senses; they exist out of groups of gathered sensations. (Chion, 1994)
2.2.4 Immersion

Immersion is “characterized by the diminishing critical distance to what is shown and increasing emotional involvement in what is happening” (Grau, cited in Rosar 2008, p. 133)

The illusion of being immersed in a three-dimensional atmosphere is greatly enhanced by audio. Audio by nature is three dimensional and can therefore help to overcome the two dimensionality of the image. Spatial acoustics can also help in the immersion by creating an environment, the music, dialogue and sound effects symbolize and emphasize a sense of location whether it is a cultural, physical, social, or historical environment (Collins, 2008). There is a difference between communication of meaning through music and mood induction:

“Mood induction changes how one is feeling, while communication of meaning simply conveys information. One may receive information depicting sadness without him or herself feeling sad.” (Rosar, cited in Collins 2008, p. 133) The communication of meaning is a far more frequent occurrence.

The previously mentioned realism plays a big part in immersion because realism is needed to become immersed; it is not easy to get immersed in an audiovisual text that isn’t realistic. Realism plays an especially significant role in the construction of spaces which in turn drives toward immersion. (Collins, 2008)

2.3 Psychology

2.3.1 Sound perception & hearing

There are several ways to describe sound. Adding the word “perception” makes it even more complicated, because perception brings a subjective element to the meaning. In order to have a more concrete starting point we will take a look at some definitions. In the Encyclopedia Britannica (2012) perception is defined as such:
“Perception, in humans, the process whereby sensory stimulation is translated into organized experience. That experience, or percept, is the joint product of the stimulation and of the process itself.”

Or, from Latin: Perception (perceptio, percipio) is the organization, identification, and interpretation of sensory information in order to fabricate a mental representation through the process of transduction, which sensors in the body transform signals from the environment into encoded neural signals (Schacter, 2011)

Often people talk about different kinds of perception, as totally separate entities, even though there may be different kinds of perception, they are linked in some way. Visual and auditory perceptions for instance are much more linked than one might think.

An interesting example of this is an experiment on the McGurk effect. In the documentary “Is seeing believing?” this experiment is demonstrated and explained. In the documentary you see a close up of a man's face saying “ba” however when you close your eyes you hear “fa”. The scientists have recorded the “ba” sound and put it behind the “fa” clip. Our brains fool us into believing that the facial movement i.e. the visual is true. Even when you know it is some sort of illusion you still can’t help but hear “ba”. This shows that there is a very close connection between visual and aural perception. (Austin, 2010)

Chion (1994) confirms that there is a strong bond between the aural and the visual, he writes that visual and auditory perceptions influence each other by lending their properties by contamination and projection. In perception of sound a guideline is, the clearer the treble e.g. the high end of the frequency range, the faster the perception. This also leads to a stronger sensation of presence in the audiovisual text. When film sound progressively got better, the frequency range increased. Since vision and hearing rely a great deal upon one another high frequencies make for faster perception of on-screen events. Because of this relationship between sound and the image we have favored rhythm in the image, much as is found in music, over a continuous flow of events, common in silent film. The rhythm found in music seeped onto the image. Films would be cut with a certain sense of rhythm. In conclusion the rhythm and speed found in current cinema is thus thanks to the influence of sound.

Dolby stereo, an analogue stereo cinema sound format, has made it possible to use more noises in film, which ultimately lead to greatly improved reproduction of noises. These noises are of good quality, tailored for each film and no longer “signs” of sound effects. Because all of this
Dolby has greatly changed the balance of sounds, noises are no longer conventional signs for sound effects. (Chion, 1994)

The reason for the fact that the ear processes external information faster than the eye is because the eye has more to do. The eye has to explore and follow both time and space, the ear needs to follow only in time. Added value, as discussed before prevents us of noticing these different speeds of cognition, between the eye and the ear. Auditory punctuation is used as a device for this difference in perception. In an action film with many rapid movements, the spectator does not get confused because the movements are “spotted”. By the use of sound such as punches, shouts, whistles etc. certain points get marked and leave a strong combined memory, an audiovisual memory. This is mainly an unconscious process for the audience, but it can be used as a tool in the production of a film (Ibid.).

Sound can also direct our attention to a certain place within the image, for example when we hear the sound of a crying baby; we will most likely focus on a crib in the room, not to a sofa, that happens to be in the room. And sometimes sound even makes us believe we see things that aren’t really happening; an example of this is in the Star Wars film “The empire strikes back”. In the film an image of a shut door is shown, then a jump-cut to an open door. Hardly anyone notices this, because of the familiar “pssshh” sound that is heard. The spectators think that they just saw a door slide open. It is so convincing that we get fooled by our minds. (Chion, 1994)

2.3.2 Listening modes

Hearing and listening is not the same thing, according to Chion (1994) hearing is passive while listening is an active action. When hearing, one receives information through the ears, listening on the other hand occurs through filtering, selective focusing and responding to the sounds. There is not just one way of listening, there actually are different modes of listening. There appear to be different ways to understand what we hear around us. There are three of these listening modes, causal and semantic, reduced. The first two are familiar to most people, the latter asks for special attention.

Causal listening means listening to a sound in order to get information, which can point out the cause of the sound. This is how we usually listen, except for speech. For example a doorbell, we listen to the sound, figure out it’s the doorbell, and go open the door (Chion, 1994).
Semantic listening is what we do when we listen to spoken language and other code systems. If you listen causally and semantically at the same time you can understand much more than just linguistics. Not only who is talking, but you can listen to subtle sounds people make while talking. Reduced listening is very different from the first two modes. Reduced listening is real-time awareness of all sound quality parameters, such as pitch, timbre and so on. This is what one uses listen to sounds itself, not the source or meaning. This is something you need to learn, this is not a natural listening technique. (Chion, 1994)

2.3.3 Gestalt principles

The gestalt principles comprise several principles used in visual perception. Gestalt is a physiological term that means “unified whole”. This theory says that specific properties cannot be derived from a summation of its component parts. (Sonnenschein, 2001)

Sonnenschein (2001) writes that the principles can also be used outside the field of visual perception; one can find equivalents in aural (sound) perception too. Our minds use these methods to analyze information. The gestalt principles support Chion’s (1994) “Synchresis” theory mentioned in the beginning of this chapter. These principles are very important for sound designers because it shows the willingness of the subconscious minds of the audience to accept the points of synchresis, and to believe the audiovisual contract (Chadwick, 2010).

*Figure versus ground* in terms of sound means almost the same thing as with the visual. For example when a sound that you are interested in pops out of room filled with other noises. That noise becomes the figure against the ground of the background ambience.

*Completeness*, our mind likes complete patterns. That is the reason we feel tension or conflict whenever a sound does not finish the way it was intended.

*Good continuation*, sounds often change smoothly when coming from one source; if sounds change abruptly another source has been set in motion.

*Closure*, As with the visual lines the mind likes to connect in visual perception, the mind also wants this for sound. The mind tries to finish melodies without an ending or broken sentences. When somebody skips a word in a sentence, we usually still know what the person meant to say, our minds fill in the gap.
Proximity, in aural contexts proximity means close together in time. Sounds that follow each other tend to be grouped together as one sound object. In music notes become a melody for example.

Similarity, similar sounds get grouped together, even when they are not close together in time. For example a barking dog, there can be tons of city ambiences in between two barks, but they are seen as one sound.

Common fate is the phenomenon where two parts of an intricate sound go through the same sort of changes simultaneously and then get grouped together and perceived as coming from one source. These changes can be in volume, pitch etc.

Belongingness means that only one component of sound can be assigned to one source at a time, it belongs to that source. For example a whistle, if there are a bunch of birds in one scene, you can’t have a person whistling and expect people to understand that the person also is whistling, at least not without a visual of the person whistling. (Sonnenschein, 2001)

2.3.4 Emotion sound

Primary emotions are the emotions of the character we watch on screen, according to Sonnenschein (2001) sound can create emotions to the primary person/group or the secondary group, which is the audience. One can identify with what the character is feeling or what the audience should be feeling.

For example a woman is walking in the woods, crackling twigs are heard. If the woman turns around it would be a primary emotion. If on the other hand she doesn’t seem to hear the sound, it is a secondary emotion, meant for the audience only. (Sonnenschein, 2001)

This can also influence our reaction on what happens on screen, if you for instance see the woman walking through the woods and you hear footsteps fast approaching, but she doesn’t hear it. You may think: “look back; there is somebody right behind you!”

2.3.5 Perception of time

One of the most important effects of added value is on the perception of time in the image; sound can have a significant influence on time within the image. Sound affects time in images in three ways.
The temporal animation of the image, this means that sound can turn the perception of time within the image as, exact, detailed etc. or as vague, variable etc.

Temporal linearization means that there is a linear story in a shot because of synchronized sound, otherwise the voices would be backward or the like. This does not mean that the entire film needs to be linear.

Vectorization & dramatization, sound is also capable of orienting shots toward a goal and creates imminence; a shot is going somewhere (goal) and is oriented in time.

However there are some conditions for temporalizing images these are: How the sound is sustained, predictability of the progress, tempo and sound definition.

How a sound is sustained basically means that a continuous sound is less interesting than an uneven sound. While a predictable sound is less “animating” than an unpredictable one. The tempo is all about the aural flow and whether it is regular or irregular. Finally the sound definition is about the richness of a sound, especially regarding high frequencies.

In conclusion, sounds can only influence the visual temporally when several conditions are met. First of all the image must lend itself to it, and the image should contain several structural elements. (Chion, 1994)

2. 4 Music in film theory

2.4.1 Categories of relation between film music and motion pictures

One of the early researchers of music in film came up with a theory on music in film. He came up with three different ways music in film functions. This was one of the first theoretical frameworks in the field. The three different ways music can function in film are, paraphrasing, polarization and counterpoint. A more detailed explanation can be found below.

1. Paraphrasing: “the specific character of the music corresponds with the specific content of the picture; the effects are presumably additive.”
2. Polarization: “The specific character of the music moves the ambiguous or indefinite content of the picture towards the character of the music.”

3. Counterpoint: “The specific character of the music contradicts the specific content of the picture; thus, the music conveys irony or comments on the content of the picture in another way.”

These are the basic categories of relation between film music and the image according to Pauli (1976). In all these three instances the music shifts the meaning of the image. This might be interesting to take a look at in the analysis section, to see whether or not this framework from the late 70's is still applicable.

2.4.2 Aesthetics and adjustments

According to Zettl (2005) music is one of the most effective means to fulfill a variety of inner orientation functions. In these orientations it can set mood and add to the aesthetic energy of a scene. Zettl also writes that sounds probably are the most direct aesthetic device to set different moods. E.g. happy sounds would transfer that feeling to the film. Scary sounds could influence the same film, and turn the happy film into an eerie film. This is called the associative power of music. (Zettl, 2005)

We tend to adjust video to the audio. This is an interesting point because it seems like the result from empirical research done by Ellis & Simons (2005) indicates that visual is the primary source of emotion. This seems to be inconsistent because it seems strange to adjust the secondary source to the primary source. One could wonder what Zettl bases this on because he does not use any reference to support this.
2.4.3 Associationism

One of the main film/music researchers is Annabel J. Cohen. Cohen has written quite a lot about musical soundtracks and how they relate to film. In her earlier research Cohen (1993) writes that Associationism is a good foundation in understanding musical influence. This ties back in with what Zettl (2005) wrote about the associative power of music. Associationism is the direct transfer of meanings created by music to the visual context, setting the mood or disambiguating the plot. It has four components, sensationalistic, mechanistic, reductionistic, and connectionistic. (Cohen, 2001)

Sensationalism could easily be misunderstood. Especially in connection with films, in psychology or physiology it directs our attention on the senses or how we experience the sensations. Sensationalistic identifies the basic components of mental experience with sensory experience. This means that simple ideas are basic unstructured sensations such as emotions.

Mechanistic is the use of simple algebraic rules which predict the properties of complex configurations from the properties of the underlying simple ideas. When it comes to music in film the Mechanistic approach is useful because when you add music to film it adds a different meaning to the film one is watching. It provides quality to film, and music is a great means of expression to a visual. In this case the effect of combining music and film can be predicted by using the mechanistic approach.

Reductionism is the art of explaining actions of complex nature and finding a simplified understanding to it by breaking it down into smaller parts. For example: understanding film and music separately. (Ibid.)

Connectionism consists of ideas, sense data, memory nodes that are associated together through simultaneous or contiguous experience. A good example of this is leitmotiv; this is a musical theme that accompanies a specific character or activity. The leitmotiv can take on the meaning of the character, so that the music alone is enough to remind us of the character.

Music is assumed to have a direct path to emotions and sensations, musical meaning can be attained by contiguity with film meaning, and music creates its own meaning and introduces information about the film that the visual does not present. The associationist framework can help to provide the empirical data needed to fine tune this theory. (Cohen, 1993)
2.4.4 Multi-level congruence Associationist framework

In the last paragraph Cohen’s Associationism was discussed. However in a later article (Cohen, 2005) Cohen writes that the additivity of associations in fact can be insufficient, because it doesn’t account for congruency. Congruency means that attention is drawn to specific visual information through principles of grouping. In the example of film and sound tracks it means matching genres, for example scary music with a horror film. This was demonstrated in new psychological experiments described by Cohen (2005). A Multi-level congruence associationist framework should be used instead, because this accounts for both congruency factors and associationism.

Even though there are many examples where associations from music seem to combine directly with associations generated by the visual information, it’s not always that simple. Associationism and congruence should come together to explain this.

Cohen also says that not all musical information actually influences our interpretation; this depends on the fact that our working memory is limited. We cannot absorb all the information that comes to us. Mental processes must relate the incoming information to past experience in the long term memory. If music is not adding to the visual, thus incongruent then it is harder for people to focus, because there is only a limited amount of working memory. However if the music is helping to build the visual, thus congruent then it will be easier to learn and remember the information. In conclusion associations are applied to focus the attention that is under control of structural congruence. Therefore structural congruence should also be included in the framework; it gives a more complete context that can even include other types of audio such as speech and sound effects. The framework exists of five different channels, text, speech, visual narrative, music and sound effect. These are the five domains that contribute to multimedia presentation. They are separate but can interact with each other.

At the surface stage, our sensory organs, i.e. the ears and eyes receive the physical information. At the next stage the information gets analyzed into structural characteristics (e.g. accent patterns or contours) and meaning characteristics (associations created in the mind). The short term memory is the next stage, at this stage this new info gets processed but old memories also get pulled in at this stage. Some of the info moves up to the long term memory. In the long-term memory all our knowledge that we gained throughout our lives is stored. The new information in the long term memory will stay there and become a part of our knowledge and experience. (Cohen, 2005)
2.4.5 Empirical Research

There is quite a lot of empirical research on musical sound track in film. More so than any other type of sound in film, Chion (1994) writes that the reason for this is the fact that music is culturally divined and therefore the easiest subject.

2.4.6 Early research

Early research existed of 2-min abstract animation, containing three figures: a large triangle, a small triangle and a small circle. Participants were asked to judge on three categories: Evaluative (good or bad), potency (weak or strong), and activity (active or passive).

The researchers added adagio and allegro to the animation, there also was a control group who only listened to either the allegro or adagio (music only), and a group who saw the animation
without sound. (Cohen, Marshall, 1988) The overall ratings of the film changed when music was present. The small triangle was judged differently when music was present, which means that a soundtrack can change the meaning of film in a direct way. On an evaluative level the judgments of the films depend on the appropriateness of the pairing of film and music, a type of cognitive congruency (Cohen, 1988)

Further research that moved away from the abstract film, to more realistic film material two one minute shorts film music pieces were presented to over 50 students. They were rated on semantic differential scales. They also had two different film clips; these were judged by different groups. One of the films had a man chasing a woman; the other showed a fist fight between two men. So if music indeed does influence meaning of film then the different background music should alter the judgment of the film. The result was that it did in many instances but not all. This in turn led to questions about visual ambiguity and film music. In the case of the fist fight it remained aggressive even combined with completely different music, which was not at all aggressive. (Cohen, 1993) The results of this study are important for the research. This is a foundational research which creates a prelude to today’s research.

2.4.7 Qualitative methods

The research done by Bullerjahn and Güldenring (1994) showed that music influences the understanding of the plot, more than anything else. They did a very different sort of research than all of the before mentioned research. A ten minute long ambiguous film was created, several professional composers were asked to create a certain type of sound track. Crime, melodrama and another indefinite piece of music were used; Bullerjahn and Güldenring chose five sound tracks in total.

What was very different from earlier research was the fact that they also had open ended questions in their survey, before there were only closed category responses. Now however, there also were open ended questions such as, what will happen next? What is the relationship between the characters? These findings are of vital importance to our research, because of the mixture of qualitative and quantitative research was sought after.
In the closed category responses where the subjects judged the mood of the film on six unipolar rating scales, such as sad, mysterious, thrilling, sentimental, clear and vivid. They found that there were big differences all but the clear and vivid, depending on the background music.

However the results of the open ended questions are more interesting. It seems that each musical soundtrack creates its own type of film and plot. The relationship of the man and woman in the film, were clearly judged based on the music. And so was the question on how the movie would continue in the fantasy of the subjects in the survey. This seems logical since there were not many other clues than the music on which the subjects could base the judgment on.

Together quantitative and qualitative research show that film music alters and divides the emotional atmosphere and influences the understanding of the plot. (Bullerjahn & Güldenring, 1994)

Vitouch’s (2001) research is quite similar to the previously mentioned experiment by Bullerjahn & Güldenring, in the sense that they both included a qualitative section in the research. Vitouch proves that there is an effect when different music is played in the same clip; the effect is however not as strong as expected.

These results came from the quantitative section of the research where the subjects rated their experience on specific rating scales. This proves that qualitative research is needed to create a more complete understanding of why these expectations aren’t reached. In the qualitative research participants were asked to answer open ended questions, on what the film was about and what would happen next. In several of the participants music seemed to have a “coloring” effect on visual perception. They are “hearing” pictures, which means, seeing details that are not shown on screen, and can only be inferred from the music.

Another thing that the study showed was that music is individually experienced; some may say a song is sad while other describe the same song as romantic. This plays a part in the emotions created from watching a film with music. The participants’ momentary mood might also influence the error in variance. (Vitouch, 2001)
2.4.8 Cognition and incidental learning

Boltz (2004) did research on whether or not music is encoded into the cognitive system, while watching a movie. Would the visual and music be encoded as a whole or as separately? She also raises the question on how music is represented in memory relative to film, which is also interesting because there is not much reliable research on this either. There were two separate experiments in order to answer the questions that were mentioned in this paragraph.

Yet again this study by Boltz focuses on congruence; this seems to be a very important element in the field of study of sound and film, since it is a part of many different researches. The first experiment by Boltz consisted of a set film clips with either congruent or incongruent moods in relation to the film clip. This has been done before; however what was interesting is that the participants were asked to remember either the film, music or both. They administered the same memory tests on all of the participants, to see whether or not incidental learning takes place. Incidental learning happens when focusing on one task and learning something else in the process, for example focusing on music and still learning about the visual.

The results showed that both encoding and remembering of film and music is influenced by mood congruency. The mood congruent clips were better remembered than the mood incongruent ones; this is consistent with Cohen’s (2001) research.

Mood incongruence leads to independent encoding and it also leads to nonintegrated memory representation. Mood congruent clips lead to the opposite effect, which means that incidental learning does take place in mood congruent relationships.

The second experiment by Boltz was done to confirm the results from the first experiment and to examine the remembering of music and film pairs as a function of mood congruency more closely. Boltz did the same experiment as before, except for the fact that only tune recognition was tested, instead of both visual and tune recognition.

The result confirmed that memory representation of music and film pairs vary with mood congruency. Mood congruent music and film are encoded together, which leads to incidental learning. In short this means that the subjects can remember the music by viewing the film. As expected the mood incongruent did the opposite, incidental learning did not take place. (Boltz, 2004)
So mood congruence both increases the remembering of music/film. It also leads to a different kind of encoding than mood incongruent film/music. In theory these results can be interpreted within the “Congruence Associationist” framework developed by Cohen (2001).

2.4.9 Physiology

Nearly all of the studies that have investigated music and film have used cognition or scheme-based approaches, without researching physiological effects. (Ellis & Simons, 2005) The common factor in all the earlier mentioned research is that music adds to the emotional impact of film.

In a study by Ellis & Simons (2005) Physiology was studied, subjects were supposed to rate their emotional reaction on valance(negative/positive) and arousal (low/high), they also used heart monitoring devices, EMG recordings and skin conductance responses were tested, in order to see if music and film is additive of valence and arousal. The subjects viewed thirty-two 6 second long film clips with music playing in the background, which would change every once in a while. The participants were told to treat it as they would do with actual film music. The research confirmed that positive films would be viewed more positively than negative films, and that high arousal films would be more arousing than low arousal films. Ellis & Simons also found support for the fact that films that were showed with positive music would be viewed more positively than films with negative music. Positive music had higher valance ratings of both positive and negative films. High-arousal music also elicited higher ratings with positive music.

Films showed with high-arousal music would be more arousing than films with less arousing music. This hypothesis had mixed results. The research further showed that physiological testing and self-testing were not always getting the same results, which makes it unpredictable and complicated. This physiological oriented research proved once again that aggressive films were less influenced by friendly music than the other way around. (Ellis & Simons 2005)
2.4.10 Sound track as a whole

According to Lipscomb & Tolchinsky (2004) film music is not a separate type of sound, but a component of a larger spectrum of sound, which also includes the dialogue, ambient sound, sound effects and silence.

They also state that in the absence of music, other types of sound can take over the role of music and function in a similar way. “By providing dramatically shifting and structurally meaningful sound to propel the narrative forward” (Ibid). This is interesting in regards to emotional affect created by sounds other than music in film. Lipscomb & Tolchinsky (2004) go through some models, and how theoretical frameworks have changed over time. First there was a tripart communicational model that was made in 1980, by Campbell & Heller that consisted of:

Composer → Performer → listener

This model was later expanded by Kendall and Carterette (1990) into a model of music communication. Their model contains various states of coding, decoding and recoding.

Music is a “culturally divined perceptual artifact, existing in the minds of enculturated listeners.” For this communication to work, the listener needs to share the same implicit and explicit knowledge structures.

“Music has all the earmarks of a true symbolism, except one: the existence of an assigned connotation” and, though music is clearly a symbolic form, it remains an “unconsummated symbol”

- Langer (as cited in Lipscomb & Tolchinsky, 2004)
This means that there needs to be a balance in the emotional and expressive intent of the message between the composer and the listener.

### 2.4.11 The role of music in cinema

The reason why there is an almost total lack of voice in the classical score is because music is an unconsummated symbol while sound is a consummated symbol. The human voice moves the musical symbol a step closer to consummation. This may leave them thinking where the voice is coming from, which is not what a musical score is intended to do.

There must be a certain interaction between the verbal dialogue (consummated symbol), the cinematic images (consummated symbol) and the musical score (unconsummated symbol) in order to make the biggest impact.
Gorbman (1987) mentions three methods where music signifies in the context in film:

1. Purely musical: where one tone relates to another.
2. Absolute music: give meaning and organization to the musical sound through patterns of tension and release.
3. Cultural musical codes: music that we associate with a certain mood or state of mind. (Lipscomb & Tolchinsky, 2004)

2.4.12 Music functions

This section contains several methods in which a musical soundtrack can communicate to the audience. These methods also represent the different ways in which music can bring forth emotions or convey the dramatic purpose of the story. (Lipscomb & Tolchinsky, 2004)

Music can express the general mood of a film. Music can give clues about genre, is the film supposed to be scary, comedic, romantic etc. This greatly depends on the level of ambiguity in the image. The more ambiguous a scene, the more influence music has. A fight scene will still be perceived as violent even with a funny tune in the background, this because of the lack of ambiguity in the clip. This would be a case of audiovisual mismatch; this will be discussed later on. Music can convey the scope of a film. Music can communicate whether the film is an epic drama, such as “Star wars” (1977) or a film on a more personal scale.

Music can convey the quality and size of a space & placement in time. Music can for instance make small rooms or artificial spaces seem bigger and it can add realism. Music can also “authenticate” the era by using music from a specific period of time.

Music can express a sense of energy. The choice of music can influence the energy that is perceived. For example a slow classical piece in a chase scene can make the scene seem slower.

Music can communicate the overall perspective or message intended by the director, this can relate to both events and characters. An example of how this works it that the same action can be depicted in different ways only by changing the musical score.

Music can convey the internal life, thoughts, and feelings of a character. The unspoken thoughts can be communicated through music, it can convey character. Leitmotiv is a commonly used
example of this function, it is a theme that represents a person and is usually played whenever the character is on screen.

Music can also convey narrative structure. Music can clarify or even establish a sense of order through the perception of a formal structure.

Music can communicate messages about where in the frame the audience should focus attention. According to research, music either has a negative or positive connotation. When music with a certain connotation is combined with an ambiguous scene, the memory of visual objects is greatly affected by the music. Another focusing device is temporal coincidence, where specific moments in the music draw attention to specific moments that is simultaneously happening in the visual.

Music can readily convey pace. If patterns are created in the musical score using music, sound effects, silence etc. and later changed, a film can look like it is speeding up or slowing down, even though in reality it is not.

Music can mismatch conditions between the audio and visual components. As mentioned before mismatch is when a genres clash, there are certain instances when this is used. For example to convey irony it makes the audience think, consciously or unconsciously what the intended meaning is. Silence or the lack of sound can also create a mismatch, when sound is expected and it isn’t there it makes the audience reflect on why it isn’t there what is the meaning. (Lipscomb & Tolchinsky, 2004)

2.5 Other sounds

2.5.1 Sound effects

Noises and other natural sounds have been left behind for a long time. There aren’t many studies on this type of sound. There is a lot of research regarding music in film; music is culturally understood, and therefore much easier to study. Noises have been neglected in research because they are seen as part of the diegesis; they are supposed to be there. Ever since the advent of sound cinema, the focus has been on the voice and music. Noises received much less attention. In the old days noises would interfere with the intelligibility of the dialogue. Therefore directors
often decided to get rid of these noises and replace them with the familiar stylized sound effects, as heard in many films of that time (Chion, 1994).

Another reason to get rid of noises is because of cultural reasons; noise is part of the sensory world. However it is totally devalued aesthetically. We perceive noises around us all the time, but this is not how film sound works. Film sound needs to be clean of noises that have no specific reason to be there otherwise the audience might start wondering where the sounds are coming from. This could take away from the immersion in the film. An aesthetic choice that composers back in the early days of sound cinema often made was to reconstruct all the noises seen on screen through their musical score. That way there was no need for sound effects and the like. (Ibid.)

It wasn’t until the coming of Dolby that noises would be heard during dialogue or music. The Dolby system made it possible to make use of more natural sounds in film, because of the wider sound strip and the larger number of tracks. Before Dolby, there was only one track therefore you had to choose a type of sound that would be heard, otherwise the sound would probably become intelligible. Noises were at the bottom of the list of sound choices and often were neglected. Since Dolby many genres have embraced this new layer of sound, especially science fiction, action, adventure and fantasy did so. Others have not, many auteur films have not incorporated this element to its full potential, one could wonder why. It’s fairly cheap and just as capable of expression as lighting, framing and acting (Chion, 1994).

2.5.2 Silence

When silence in films occurs, it is hardly ever the total absence of sound. Every place is unique in terms of silence; therefore there are always a couple of seconds of silence recorded on every recording set. This unique silence is also called Room-tone, sound editors use this room-tone if sounds need to be cut away from a scene; if they would just cut away the all sound the audience would think that something technically went wrong. Therefore the room-tone is looped until a new section of sound begins (Chion, 1994).

Silence is almost never 0dB or complete emptiness, it is the negative of sound that we heard before the silence; contrast is what creates silence. Film does not merely use room-tone to present silence; it also uses symbolic sounds to represent silence. These sounds are often subtle such as an alarm clock ticking, these sounds represent silence because they are normally only
heard when the world around it is silent. Another example of sounds used to represent silence are far away animal sounds such as a dog barking in the distance or clocks in the next room etc. Or even adding some reverberation (echoing sounds) to a sound, to highlight the sensation of emptiness in the sound can represent silence. This works because we cannot perceive reverb when many sounds are present at the same time (Chion, 1994).

2.5.3 Point of synchronization

“A point of synchronization, or a synch point, is a salient moment of an audiovisual sequence during which a sound event and a visual event meet in synchrony.” (Chion, 1994) Synch points generally obey the gestalt laws, but there are some exceptions:

One of the exceptions is when an unexpected double break in both the audio and the image tracks; this is used for external logic. Another exception is when it’s used as a form of punctuation, where at the end of the sequence the tracks come back together after having been separate. The third exception is when a synch point exists purely because of its physical character, for example a sound in the sound track that is much louder than the rest. And lastly if a synch point has an affective or semantic character, a word can have a strong meaning and be spoken is a specific way. (Ibid.)

Almost everybody will recognize it instantly, the cinematic sound of a punch. In real life however a punch doesn’t have to make a noise at all, and if it does make a noise, it sounds nothing like the cinematic punch. Even though most of us know this, we still desire the sound, because without it the punches seem unrealistic even if the punch actually occurred. This sound is an immediate representation of the audiovisual synch point. The punch becomes the center of the film, around which time is constructed. Before the punch we await it, after it we feel the sound waves. It is the combination of image and sound that etches the punch into our brains. (Chion 1994)
3. Method

Based upon the material from the literary review, a couple of different methods of investigation, forthcoming from the various theories presented in this exposé, were chosen to analyze an audiovisual text, in this case a short film clip. The clip chosen was the official trailer for the movie *Inception* (2010). The intention of using the methods of investigation in this thesis was trying to identify whether this approach would lead to consistent results, or alternatively the outcomes could complement each other in some way. If the results would turn out to contradict each other there would still be conclusions to be drawn from this research.

3.1 Sampling

A pilot survey will be presented, in which ten subjects have analyzed the clip and answered the survey questions. There was made use of a specific observation method, namely the masking method (Chion, 1994). Every subject watched all three versions of the clips, in the same order on their own computer. The subjects were forced to follow the instructions in the survey, guiding them structurally through the process; these instructions can be found in Appendix II.

The subjects were acquaintances known through several different universities in different countries in Europe and North America; all subjects are white Caucasian from northern and western European descent. This was a convenient sampling. The subjects were asked through email if they were willing to participate in a survey. If the persons agreed they received the survey (Appendix II) through email. The author was not present during the surveys; however subjects could contact the author if they were in need of help, either through email or social media. The subjects were allowed to take as much time as they needed to answer the questions in the survey, however the subjects were not allowed to watch the clip more often than stated in the instructions. When finished the subjects sent the survey back by email.

All the subjects fall within the target group of the investigation, namely young adults (20-27) associated with a student population of university level. There are several reasons for this choice:

The first reason is that young adults visit movie theaters more often than some other age groups, as shown by e.g. the BFI Statistical yearbook (2011). The second reason is that the author has a direct access to this population. The third reason is that this target group has a fairly high chance of fully understanding the questions, both linguistically and regarding a level of abstract thinking that is probably required to answer the questions. These skills seem to be necessary for studies at
Universities and other similar institutions, an example of this can be found on Karlstad University’s website, the link can be found in the reference chapter.

The writer of this thesis has conducted a thorough analysis, also involving the observation method, but the author will go beyond that, and do a shot breakdown after that, to be able to see all the details.

3.2 Observation method

3.2.1. Masking method

The observation method that was used in this study is the so-called masking method (Chion, 1994). In this method the person that analyzed the audiovisual text (the person further referred to as: the subject), screened the audiovisual text several times. Ultimately the subjects watched and listened to both the audio and the visual, after initially twice judging the masked version in which either the audio or the visual was concealed; thus essentially cutting out one of the senses in this multisensory event. The ultimate purpose of the masking procedure in this study is to let the subject hear the sound as it actually is without the image contaminating the aural perception. (Chion, 1994).

3.2.2 Quasi-experiment without random assignment

The observation method, combined with the survey, could be called a quasi-experimental study. This quasi-experimental study could in the future be turned into a standard experimental research. In the current quasi-experiment, the subjects were not randomly assigned to a particular condition. There is a selection of an existing group and it is therefore not directly used for the manipulation of a variable. On the basis of the choice of a quasi-experiment for this study, there are no manipulated variables, such as in a normal experiment (Adelman 1991). Being a pilot study, we searched for the variables that were expected to lead to a causal relationship.
3.3 Mini-survey

This quasi-experimental research survey functions as a pilot study, because it is not possible in the context of this assignment to do a survey on a larger scale. The schedule for this research made no room for a full scale survey, it did however provide enough time for the pilot that was performed. The research is therefore at this point not meant to produce a generalized outcome.

The rating scale that was used was a scale from 0 to 10, where 0 is low and 10 is high. An example of this scale can be found in appendix II, in the full survey. SPSS Statistics 17.0 was used to generate the results in order to get an overview of the results.

The following questions were presented to the subjects in the course of the survey. This section also includes explanations in regard to why the questions were asked. Firstly the parallel questions on the three observation modes are shown and explained; then the remaining questions on the different modes follow.

*The following questions were asked three times, in all three observation modes.*

Question: The clip is a little under two and a half minutes, did you perceive it to be that long or did it feel longer or shorter?

This question may answer how the subjects perceived time, and whether it changed in the three different screenings. The first two clips were masked and the last time it was presented with both sound and image; this way performing a masking procedure may show whether the perception of sound could affect the perception of time.

Question: How do you feel listening to the clip? (This pertains to the clip not your mood in general)

This question is based on answering the research question about emotions and sound. By asking this question three times under different circumstances, one should be able to see whether or not the sound had any effect on the emotional state of the subject.

The author postulated that the sound of the clip would create a suspenseful mood in the subjects and that the sound would most likely be perceived as exciting and would trigger a level of excitement in the subjects. Whereas the clip without the sound would seem
somewhat boring and bland, therefore the author presumed it would most likely leave the subjects emotionally unaffected. The last clip containing both sound and visual should have an accumulating effect, as a level of synergy might be expected. This means that the last clip should be perceived as exciting, however on a higher level than the first screening; this is tested in the following question.

Question: Rate your level of emotion after hearing the sound:

This also pertains to the emotion question, however here the subject rated their emotion on a scale. This way we can see the level of emotion that may be created under the different screenings.

Question: How would you rate the level of cinematic realism? (If that was the real world, e.g. Star wars has cinematic realism but it isn’t realistic)

This question pertains to realism; the author believes there may be a link between realism and immersion, the latter may also be related to emotions.

Question: How would you rate your level of immersion? I.e., did the story engage you?

The level of immersion was tested because it could be interesting to see whether the level of immersion connects to the level of realism. Because this question was asked after all three screenings one could see whether the level of immersion changed depending on the presence or absence of sound. This may show the importance of sound.

Extra question for the Clip without sound

Question: What did you like better: the clip with just the sound or the clip with just the visual?

This may be interesting because we could see what had a bigger impact: the visual or the audio.
Extra questions for both sound and visual

Question: How important do you think the sound is in combination with the visual?

This question pertains to the query about the importance of sound in the clip. Here the subjects could rate how important they thought the sound was for their experience in this particular clip.

Question: Do you think the sound is compatible to the visual? Does it seem like the sound belongs to the clip?

This question should answer the congruence theory, and also the categories of relation between film music and motion pictures.

Question: What do you think is the function of the music in the clip?

Question: What do you think is the function of the sounds (other than music and voice) in the clip?

Question: What do you think the function of the dialogue is in the clip?

These questions pertain to the functions of film music theory; this answered what the subjects thought applied to this particular clip. The voice and other sounds were tested this way to see if this theory could cover more than just music.

For the aural gestalt rules (Sonnenschein, 2001) mentioned in the theory chapter, the overall results of the questionnaire will be examined.

3.4 Choice of Method

Chion (1994) writes that audiovisual analysis is performed to understand a sequence or film in the way it uses sound in combination with the image, which if what we are trying to accomplish. It would be nearly impossible to implement theories without any reference to an actual sequence.
clip was therefore chosen to be analyzed and to demonstrate how these theories and methods can be applied in a real audiovisual situation.

Sound and films for that matter are subjective, everybody may have a different opinion on sound in film, and therefore it may be difficult to conduct interviews or surveys. Hence the choice was made to perform a quasi-experimental pilot survey. This approach may have better result because the subject does not have to filter away the excess information in their minds. The method used, exercising the three different observation modes, itself filters away this excess information. Chion (1994) writes that this method of observation gives the opportunity to hear the sound as it is and not as the image transforms and disguises it. The same also goes for the image as it is. Therefore subjects can answer what they feel without having to think too much trying to free themselves from the bigger context, in this case filled by the visual images. The questions have been constructed in a way that actually is examined what should be researched, without leading the subjects to the answers the researcher wants, thus having a certain content validity (Ekström & Larsson, 2000).

3.5 Validity & Reliability

The goal of the research is to answer specific questions and check their applicability outside of the given situation. The research questions are tested through additional questions in the pilot survey. These additional questions were meant to cover the entire research, not just one or two of the research questions.

As mentioned earlier on in the method section, the clip will be screened three times by the subjects. Validity in this way was established empirically through this observation method, which demonstrates that meanings change given the absence or presence of sound for any given audiovisual text. In the next section it is shown how it works in more detail and with more examples.

Because the survey is a pilot study, with a small sample, it is somewhat difficult to show high reliability. However this survey will be used to generate a proper instrument and it is therefore a learning experience.
4. Results

4.1 Instrument

The instrument exists of a survey and a shot break-down. A sample of the shot breakdown can be found below, while the full version of the shot breakdown is found in Appendix I. The survey can be found in appendix II.

4.1.1 Shot breakdown

A shot breakdown was preformed to clarify which exact shot is being analyzed. Also it structures the elements of the clip and this makes it obvious for the author as well as to the reader what sounds can be found where, without watching the clip over and over. In the analysis you will find references to some of the shots, a small sample of this shot breakdown is presented here.

**Inception trailer**

(Orchestral music starts)

1. A shot of the ocean.

2. ECU of a man lying on the beach, while waves crash into him. (Male voice: “There is one thing you should know about me”.)

3. LS, Hallway, man walking. (“I specialize in a very specific”...

4. Fade to black.

5. CU, the man’s face covered in shadows. (“Type of security”…)

6. CU, a gun being loaded and cocked. (SFX: a click, then a sliding metal click)

7. Fade to black.

8. CU, two men talking. (“Subconscious security”…)

9. CU, hands opening a briefcase.
4.2 Survey results

The first clip the participants screened was clip 1; this was the sample of the movie trailer without the image. In the second screening of the clip namely clip 2, the participants only saw the visual. In the third screening, clip 3, the participants experienced both the sound and image.

Based on the TV and film consumption frequency we can conclude that there is no significant difference between high and low frequency TV-show/film consumers.

Eight out of the ten subjects felt that the sound clip was longer than it actually was; the two remaining subjects thought that the length of the clip seemed to be accurate. During the second presentation of the clip, the one that just contained the visual, the subjects were divided 50-50. Half of the subjects thought the length seemed accurate and the other half thought that the clip seemed longer. The clip containing both audio and visual was mostly seen as temporally accurate. In the last screening, where the sound and the visual were combined, 70% of the subjects felt that the length of the clip seemed accurate. The remaining 30% wrote that the clip felt shorter than it actually was.

During the screening of the first clip, the one without the visual, 80% of the subjects felt excited when listening to the sound. One of the subjects felt bored and the remaining subject felt emotionally unaffected. During the second screening, of just the visual, most of the subjects felt unaffected, namely 70% of them, while 20% stated they were bored, the remaining 10%, being a single person, felt scared. None of the subjects felt excited during this screening. In the last screening all the subjects reported that they felt excited.
The level of emotion was rated with a minimum of 4 and a maximum of 9, with a 6.7 mean during the first screening. The second screening had a minimum of 2 and a maximum of 8, with a 4.9 mean. The last screening had a minimum of 6 and a maximum of 10, with an 8.2 mean.

**Table 4:1 – Rated levels of emotion**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate your level of emotion after hearing the sound (clip 1)</td>
<td>10</td>
<td>4</td>
<td>9</td>
<td>6.7</td>
</tr>
<tr>
<td>Rate your level of emotion after watching the clip (clip 2)</td>
<td>10</td>
<td>2</td>
<td>8</td>
<td>4.9</td>
</tr>
<tr>
<td>Rate your level of emotion after watching the clip (clip 3)</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td>8.2</td>
</tr>
</tbody>
</table>

When asked about the perception level of cinematic realism the first screening was rated as a 6.6 mean. The second clip received a 5.8 mean, while the third clip got a fairly high 8.5 mean.

**Table 4:2 – Level of cinematic realism**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>How would you rate the level of cinematic realism? (clip 1)</td>
<td>10</td>
<td>3</td>
<td>9</td>
<td>6.6</td>
</tr>
<tr>
<td>How would you rate the level of cinematic realism? (clip 2)</td>
<td>10</td>
<td>4</td>
<td>7</td>
<td>5.8</td>
</tr>
<tr>
<td>How would you rate the level of cinematic realism? (clip 3)</td>
<td>10</td>
<td>7</td>
<td>10</td>
<td>8.5</td>
</tr>
</tbody>
</table>

The level of immersion was viewed as a 5.9 mean with just the sound, as 2.5 mean with just the visual. With both the sound and the visual it was rated as an 8 mean.
When asked which of the two first clips the subjects preferred 80% stated that they preferred listening to the clip. The remaining 20% liked watching the visual better. Several of the subjects stated that the clip was hard to understand without any sound.

The subjects found the combination of both sound and visual very important they rated the level of importance a 9.6 mean. All the subjects thought that the sound and the visual were compatible.

**The functions of music**

The functions of music are explained in more detail in Appendix II.

The following are the percentage of subjects that thought the functions were present in the clip:

- 60% Conveying pace
- 40% Convey where you should focus attention
- 20% Conveying the scope
- 80% Convey the general mood of a film
- 10% Convey the quality and size of a space & placement in time
- 10% Convey the overall perspective or message intended
- 20% Convey the narrative structure
- 20% Convey the internal life, thoughts and feelings of a character
- No one thought mismatch conditions between audio and visual components was a function.
The functions of sounds other than music were:

- 30% Convey pace
- 70% Convey where you should focus attention
- 30% Convey the scope
- 30% Convey the general mood
- 10% Convey the quality and size of a place and space
- 30% Convey the overall perspective or message intended
- 10% Convey the internal life, thoughts and feelings of a character.
- 20% Convey narrative structure
- 10% Mismatch condition between audio and visual components

The functions of the dialogue were:

- 10% Convey pace
- 10% Convey where you should focus attention
- 30% Convey the scope
- 40% Convey the general mood of a film
- 20% Convey the quality and size of a space and place
- 10% Convey a sense of energy
- 80% Convey the overall perspective or messages intended
- 50% Convey the internal life, thoughts and feelings of a character
- 50% Convey narrative structure
- 10% Mismatch conditions between audio and visual components
5. Analysis

In this part of the thesis the clip is analyzed from the research questions and the theories which are described in the theory section of this thesis. Also we will take a closer look at the subjects’ interpretation of the clip. The survey answers of the subjects will be held up against each other and against the author’s interpretation of the clip.

From looking at the answers given by the subjects from the pilot survey we conclude that most the subjects agree on a number of questions.

On the question about the perception of time, it is clear that the clip with just the sound was perceived to be longer than the other two clips. The last clip was sometimes even perceived as shorter than the actual time. There is thus a significant difference in the perception of time.

On the second question regarding which emotion was felt during the clip, most of the subjects answered that they were excited. However in the second clip the subjects were not as positive. None of the subjects were excited, most of them stated that they were emotionally unaffected by the clip. When both the sound and the visual were displayed every single subject felt excited.

The level of emotion during the first clip, containing only the sound, was higher than the level of emotion in the second clip, presenting only the image; this coherent with and therefore to be expected given the previous answer on which emotion was felt. If the clip leaves one emotionally unaffected, you will probably not feel a high level of emotion. One thing we can see from this data is that most of the subjects perceived a change in emotion, both on what kind of emotion they were feeling and what level of emotion they perceived. Just the sound may be similar to listening to the radio; there is no relation to a frame since the subjects minimized their web browsers. However the sound did seem to have an emotional affect on most of them.

The screening without sound was seen as boring or the subjects felt emotionally unaffected, one could say there is a lack of synchresis. This means that sound in this case seems to be the major contributor of emotion for the subjects. When sound and the visual were combined, the level of emotion went up in all subjects. Therefore we can conclude that the biggest emotional affect was made by combining the two. This may indicate that Added Value (Chion, 1994) occurs here or at least some form of synergy. These results as a whole are probably not very remarkable, what is somewhat surprising is that the difference between emotions in the masked versions with respectively sound and the visual were fairly large. The variance between the incited emotions in
the masked assessments may derive from the fact that we as humans are used to listening to things without an image for example the radio, yet we are not used to seeing without hearing anything. As mentioned in the beginning of this thesis film was from its genesis never really silent, there was always music present in the theater. Even though sound wasn’t meant to carry meaning at first, since it initially was used as a masking device for projector sounds and the like, it was still present in the theaters (Cohen, 2001).

When asked about the level of cinematic realism, the subjects described that the sound was reasonably realistic it was given 6.6 mean. The clip with just the visual was seen as less realistic by most subjects. This difference in rating is a very interesting piece of data that could be used in future research on a larger population. The last screening which contained both the sound and the visual was seen as very realistic and was given an 8.5 mean.

This data indicates that in this clip the realism is best achieved when combining the visual and the sound. When either the sound or the visual aren’t available something may seem like it is missing. This may be perceived on a conscious or an unconscious level. This feeling of absence may be due to the fact that our senses are very much connected, as the McGurk effect proves (Austin, 2010). Our senses are used to working together; our sensory experiences usually consist out of a bunch of gathered sensations (Chion, 1994).

The level of immersion during the sound clip was rated as very average, at a 5.9 mean. When the subjects were watching the clip without sound the level of immersion was minimal, the subjects rated it as a 2.5 mean. However during the last screening the subjects felt very immersed, they rated their level of immersion as an 8 mean.

The level of immersion during the visual clip was seen as minimal, while the sound clip was perceived as very average. In this context we can conclude that sound was easier to get immersed in than it was in the case of the visual. Here it can also be assumed that Added Value is an essential part of the theories of film sound (Chion, 1994). The level of realism and the level of immersion do seem to follow each other to some extent, if this occurs because the level of immersion depends on the level of realism is unclear though.

When asked about which clip the subjects preferred 80% indicated that they preferred the sound clip; therefore one may conclude that this is quite interesting since sound is usually seen as an add-on. The visual is preferred from both the historical and the ontological point of view (Chion, 1994). But, as mentioned several times, film never really was silent, so maybe the image wouldn’t be preferred if there was no sound at all.
In the case of this clip, the music seemed to be the carrier of the story, there are over 90 shots in the clip and many don’t follow each other in a linear sense. The sound and music in particular, follows a theme line, and is not connected to each of the 92 shots. Thus synchresis (Chion, 1994) doesn’t seem to have a determining role here. That can be a reason for the fact that the subjects liked the sound more, while the sound told the story while the images were somewhat scattered and maybe even confusing without the context that the sound gives.

The combination of the sound and the image was seen as crucial by the subjects, as they reported. They gave it a 9.6 mean. If you look at all the data in a row you can see why they found the combination very important. Clip three was overall very positive, scoring no lower than an 8 mean on the scale questions, while the other two screenings scored much lower.

All the subjects thought that the sound and the image were compatible. This is probably not surprising; the image seems to be cut to the music. There are many synch points where for example shots change to the rhythm of the music this can be seen in shots: 22 and 23.

The dialogue is mainly responsible for the story line, what the clip is about. When watching the clip without the sound some of the subjects indicated that it was very hard to understand the story; nearly impossible even, according to one subject. Knowing how to feel is also quite hard when there are no musical clues present, especially because the image itself seems to be somewhat ambiguous (Cohen, 2001). There are actually very few sound effects in this clip; the music takes over many functions of the sound effects. This can be heard at some of the explosions for example, the sound that can be heard is from a musical instrument not one of an actual explosion.

The sound effects found by the author were: A gun being loaded and cocked, a building falling into the ocean, gunshots and metallic sounds and a large gunshot at the end which seems to morph into music. For a sequence that is over two minutes long this is a small amount of sound effects, especially given the rapid change of the action shots, where selected for this clip.

When looking at Pauli’s functions (1976) and the content of the music in the clip one could say that the function of the music in this clip is Paraphrasing, because the subjects conclude that the music is compatible to the visual content. One of the subjects does say that it is difficult to understand the visual. From that one might draw the conclusion that the clips function actually is Polarization. Polarization means that the music moves the ambiguous content of image to the character of the music.
Since all of the subjects experienced the first and the second clip differently, one might conclude that polarization is a better fit and would be is the actual function. There seems to be very little emotion in the image clip, yet the subjects found emotion in the sound clip. Therefore the sound clip will most likely be the main contributor of emotion in the combined clip. The separate clips do not appear to have the same emotional impact, and therefore polarization may be the best fit. However this does not exclude the fact that the sound and the image fit well together.

The final questions in the survey about the functions of music (and other sounds) proved to be quite difficult for the subjects to assess.

There are many possible functions in a theoretical approach of film sound; this doesn’t mean that all of them have to be displayed in the clip. However there are quite a few differences between the individual subjects and their interpretation of which functions can be found in this clip. The fact is that the subjects aren’t trained in analyzing film; therefore they probably miss some of the functions in their analysis of the clip. This however applies to the majority of people who watch movies and TV shows. Whether the attempt to provide some definitions (See appendix II) adds any value to the research cannot be determined at this point. As for now we assume that a distinction must be made between a professional analysis and an audience analysis. This was anticipated and therefore the author has performed a different yet parallel analysis.

Here we will take a look at the functions of music (Lipscomb & Tolchinsky, 2004); first we will go through the function that the music in the clip shows. After that we take a look at the same functions again, yet this time they are related to sounds other than music. Lastly the voice is analyzed from the functions of music. This may seem repetitive but we are actually looking at different aspects of sound. The purpose of this is to see whether or not these functions can apply a larger spectrum of sound, one beyond just the functions of music.
The functions of music

The functions of music that the subjects found were all functions except for mismatching of the sound and the visual.

Together the subjects found nine out of ten functions of music in the musical aspects of the clip. So which functions did only a few of the subjects find, namely 20% of the subjects or less?

- Conveying the internal life, thoughts and feelings of a character
- Conveying the quality and size of a space or placement in time.
- Narrative structure
- Conveying the scope
- Conveying the overall perspective or message intended

The music does indeed not seem to convey the internal life or thoughts and feelings of a character; there is no leitmotiv or any other indications of music representing one of the characters. Mismatching doesn’t seem to be a function in this clip either, genres don’t clash, and every time music is expected it is there. This leads us to the scope of the clip. We could conclude that the scope is conveyed, the music gives the impression that the clip is meant to be an epic clip. The music builds up quite often and makes the clip as a whole seem grand.

The music in the clip does seem to give some narrative structure to the clip; the story becomes clear when both the visual and the sound are played together. One of the subjects remarked that it was hard to understand the story without any of the clip’s sound; this could also depend on the spoken language. In this case the spoken language really “tells” what the clip is about. However it seems that the music also contributes to the narrative, because there isn’t always dialogue and the narrative still seems stronger than it did without the sound.

Conveying the quality and size of a space or placement in time, the music does make everything seem grand, so in that aspect one could say that this would actually be a function of music in the clip. However this seems to be nearly the same thing as conveying the scope. It doesn’t seem that the music makes specific rooms or spaces look big. Also the music isn’t used to authenticate the era; it is not music from a specific period from the past. Therefore we may conclude that conveying the quality and size of a space or placement in time is not one of the functions of the music in the clip.
Overall perspective or message intended the music does create a certain perspective. The clip could probably create a different meaning, if the music was different. A good way to test this would be by using the method that Cohen (1988) applies, in this experiment the music is changed to see how that affects the emotions of the audience. Chion (1994) calls this technique “forced marriage” but he writes that it isn’t a recommended way of analyzing sound in film. However, this “forced marriage” technique seems as though it would be the best option to test if changing the music actually changes the perspective. However this kind of experiment falls outside the scope of this research, but it is definitely worth looking into in future research.

The most obvious function that exists in the clip seems to be: conveying a mood, which 80% of the subjects thought was a function. Music was concluded to be the main source of sound in the clip, and the subjects felt emotions because of the sound as shown on the previous pages. Therefore one may conclude that the music conveyed a mood in the subjects. The emotions felt in the clip as a result of the music are secondary emotions (Chion, 1994), as the characters do not hear the music, the music has been added during the elaborate post-production process, whereas the emotions that were conveyed through the dialogue were primary as the character also could feel these affections.

Conveying pace seems to be a function in the clip, because the build-up of the music seems to make pace faster and the fade outs appear to slow the clip down. The buildups and the fade outs also convey a sense of energy; the music is very energetic in the build ups. These categories appear quite similar, it seems as though they may be connected. When the energy is high the pace is also high and the same goes with low pace and low energy. Therefore the conclusion is that both energy and pace are functions in this clip.

Where to focus attention seems more as a function of sounds and dialogue in this clip, for this specific reason: The music is non-diegetic; this means that the music we hear cannot be heard by the characters on screen. The music has no source within the diegesis; therefore it seems unlikely that the music would focus attention to specific actions on screen. Sound effects and dialogue quite often do have on-screen sources, which allows them to focus attention to specific items or people.

Above we’ve looked at the functions of music and how they are applied to the music found in the clip. This theoretical framework is called the functions of music, but can these functions work beyond just music?
The functions of sounds other than music and speech

The functions of sounds other than music that 20% or less of the subjects found were:

- Convey the internal life and thoughts of a character
- Convey the quality and size of a space and place
- Mismatch conditions
- Convey narrative structure

In the section above we saw that the subjects had found nine out of ten functions when analyzing the music. Here it turns out that the subjects have found more functions, namely, all ten.

There does not seem to be a mismatch regarding these types of sound. All the sounds seem to have a source and fit together with the image, for example the gun fire sounds as gun fire is expected to sound. The internal life or thoughts of the characters don’t seem to be represented by any of the sounds. The quality of space and placement in time don’t seem to be functions either.

A narrative structure seems very unlikely in this clip, because there are so few sound effects in this clip. This means that there is not a continuous flow of sound effects, which should be present if the sound effects would be a determining factor of the storyline. The question is if any of the sounds that were found in the clip propel the story forward. There don’t seem to be any specific sounds that do that. Other than that, there is no real structure in this clip as far as sound effects go, therefore a narrative structure doesn’t seem to be a function of the sounds in this clip.

Where to focus attention would definitely be a function sounds can have. People usually listen causally as described by Chion (1994). Therefore it would seem natural for the subjects to look for the sound they heard on screen and focus on it. The locus of sound attracts the focus, one could conclude.

One could argue that all sound is energy, and therefore one would think sound should have to convey a sense of energy. However this is probably not really what the authors of this framework had in mind. The gunshots at the end of the selected clip do seem to give a sense of energy, especially because the music is fading out. As mentioned earlier in the thesis the contrast of sounds creates silence, so when the music fades out the gunshots create a contrast (Chion, 1994), which wouldn’t have been as strong with the music accompanying or even prevailing the
explosion from the gunshot. This makes it possible for the gunfire to convey more energy. In this case the sound does not seem to alter the pace as it did in the music, this may be because of the lack of sound effects.

It is very tricky to say something about the function “convey the general mood” and whether it can be found in this clip. The fact is that there are very few sounds which are neither music nor dialogue. There are types of sound effects that can probably be used to convey the general mood. For instance “Mickey Mousing” (which in film refers to comedic cartoonish sound effects as used in Mickey Mouse cartoons) one might be able to understand that it is meant to be comedic. However, this clip does not seem to have specific types of sound effects, meant to induce certain moods. And because the small number of sound effects used in this clip, the author is convinced this isn’t a function of the sound effects in this particular clip.

Conveying the message intended seems very unlikely here, what was the intended message the sounds conveyed? There were very few sounds in the clip that could have conveyed an intended message; therefore this does not seem to be a function of sound here. The underlying question here is of course what is meant by this trailer. The clip is not meant to tell a linear story. Rather, it is meant to make people curious about the story behind the trailer. The function of sound is in that case is to attract or even lure people into the theaters. This shows that the underlying objective of a clip needs to be taken into account. However there will be no further elaboration on this in the context of this research. It would stretch out the purpose of this study too much.

The functions of speech

At this point we have looked at the categories “functions of music” and “functions of sound other than music”. The next category is the dialogue; we’ll take a look how these functions initially meant for music could pertain to the dialogue in the clip.

These were the functions that less than 20% of the subjects found in the dialogue:

- Convey pace
- Convey where you should focus attention
- Convey the quality and size of a space and place
- Convey a sense of energy
- Mismatch conditions between audio and visual components
Mismatch is yet again not found in the clip. The voices all seem natural and coming from the right sources. The scope doesn’t seem likely either the characters aren’t talking in a way that may be seen as epic or the like.

A sense of energy may be found in the voices, one example could be in shot 66 again, where the two men are standing in front of the cab. There the energy is different than in other shots when the same man is talking in other instances in the clip. Pace could most definitely be a function; the way people speak has a pace. When a person is talking fast the pace can be perceived as faster, and this can also work the other way around of course.

Placement in time does not apply to this clip; however this could be a function of the voice. Think about films as period dramas, the actors quite often use certain accents and words from the specific time the film takes place. This seems like a logical function of the dialogue, however not in this clip.

Conveying where to focus attention seems like a logical function of dialogue. As mentioned we as humans listen primarily causally, we look for the origin of the sound, therefore it is logical for the subjects to focus on a person on screen when the person is talking. For instance in shot 26 when you see the two men, but the one you’ll most likely focus on is the one talking, at least according to the listening mode theory. The general mood of the film seems to be created by the music; however the internal life and feelings and so on, seem like they could be a function, more on this later.

Conveying the narrative structure in the clip seems logical. Subject one didn’t understand what the clip was about with just the visual. Could this depend on the lack of music or the lack of dialogue? If you just focus on the music of the clip, it does make the clip a lot more exciting to watch, but would you really understand what is going on? This may be doubted, because the dialogue explains what the story is about, and why a lot of the actions are happening. Conveying the narrative structure therefore seems to be a function of the dialogue in this clip.

The dialogue conveys the intended message very easily, the director can just let the characters say the intended message, in this way it is definitely a function of the dialogue. By changing the dialogue the director can also create a different meaning.

The human voice has more meaning to it than just the content of the spoken word. As mentioned earlier on in the thesis we can listen in different ways, that is: causally, semantically and reduced (Chion, 1994). To the human voice we listen primarily semantically, we listen for the
content of the words. However you can also listen causally to someone speaking. That way you can hear the meaning beyond the content of a word. So the internal life, thoughts and feelings of a character, can be heard from both the content and the way it’s said. For example in shot 66 when one man angrily says “This is your responsibility, we were not prepared for this.”

As mentioned in the section about the gestalt principles (Sonnenschein, 2001), we as humans use the principles to analyze the world around us, this occurs unconsciously. Analyzing according to the gestalt principles is something we do automatically and it is therefore very hard to access this information, our minds process the information even before we can think about what is happening and builds our perception of the world around us. It is very difficult to look at these principles from a conscious level and try to analyze an audiovisual text by those principles, even though it actually is a very simple task for our unconscious mind. In conclusion, the gestalt principles seem to be an important part of our perception and cognition, but not one that needs or maybe even can be done consciously. For conscious analysis other theories seem to be better suited.

We’ve looked at the theories and results from an analytic perspective, now we will move on the next section, where the findings from this section, namely the analysis will be clarified. There the research questions will be answered and other conclusion will be drawn.
6. Conclusion

The previous chapter makes it clear that in this case, sound in combination with the image definitely can create emotions. Music seems by far the easiest way to convey mood, to let the audience know how they should feel. We accept the conclusion that many aspects of sound, which is particularly obvious in the familiarity in relation to language, but certainly also in relation to music, are strongly culturally determined and of influence in mood induction. When the meaning of words aren’t understood an important part of the transfer of knowledge and feelings is lost. Music also appears to be understood only when the person listening, has enough knowledge of the culture in which the audiovisual text is created. This was already mentioned in the theory and the survey done in this thesis seems to confirm this.

It is much more difficult to see if and how sound effects convey emotions, I am fairly sure that sounds can convey emotion on some level, but more empirical research is required on this subject. On the basis of this research there was not enough concrete evidence in this clip to support any claims.

To answer the first research question, how is the clip perceived in regards to sound in film?

The sound in the clip was generally comprehended as exciting and important. All in all the sound was seen as positive by most subjects even when it did not relate to the visual. The perception of all the variables changed in the different screenings. This would indicate that sound alters the perception of audiovisual texts.

Then we took a look at the second question: How important is sound in the clip?

The answer to this question is: Very important, the subjects all found the sound extremely important for every aspect questioned. The sound in the clip was crucial to the visual: This confirms what was mentioned throughout the thesis, that sound in fact is a vital part in this clip.

The sound also turned out to be the big contributor to the understanding of the narrative.

Regarding the question about what functions sound has in the clip, the answer is that sound can have many functions shown in the analysis section. However these were just the functions that were thought music may have according to literature. There may be other functions for sound effects and the human voice or even for the music. However, quite a lot of functions seemed to be applicable to the other types of sound too. It is therefore a good base to start from.
To answer the question, stated in the purpose; from which the title is derived from and which this research is based on, namely: Can a mood be engineered with sound?

We can conclude that for the time being, because of the limited sample, in regards to both the use of only one clip and the fact that all the facets of the theoretical models couldn’t be answered with an indisputable “yes”. Therefore the conclusion is: Yes, from what we have seen in this particular clip, with this group of subjects and this type of quasi-experimental survey, a mood can be engineered with sound; music turns out to be the easiest way to engineer a mood. An experiment, going beyond a quasi-experiment, being an empirical evaluation method, could lead to more definite conclusions. Apart from this conclusion I will share some thoughts on further research in the next chapter.
7. Further research

There has been quite a lot of research on film music, even though that field of study also needs a lot more research, here I will focus on some other aspects for further research.

One thing that definitely could be done is a continued research on this exact subject; one could further develop the survey and get a larger sample. In previous sections the limitations of the results and the definitive interpretations of this study is indicated, that mainly relate to the fact that this is a pilot research. A summary of this will follow below.

In further research it could be good to choose a couple of minutes from a feature film to analyze; this will most likely contain more material to analyze. Trailers can quite often be heavy on the music part, from my experience. Some barely seem to have any other sounds. Choosing a clip that contains more sounds than just music will be able to give sound effects and ambience a closer look. It would definitely be interesting to see more research about sound effects and investigate the effect they can have on an audience. Lipscomb & Tolchinsky (2004) raise a good point about the fact that is strange to take one part of the sound track and researching it, since all the sounds are part of a larger whole. However, they also write that sound effects could take over the role of music, which means that sound effects could be studied separately.

I would also like to see some research done on just sound effects. Even if sound effects are part of a bigger sound track it would be interesting to investigate sound effects more specifically, much like the research done on music. By researching sound effects in this way one would be able to prove or disprove whether or not sound effects by themselves can affect emotions. Something that seems interesting to investigate is sounds that seem to be quite primal, and not culturally divined, and if so how it functions. Examples could be for instance a snake’s hiss or thunder. Even though snakes for example have a different place or meaning in distinct cultures. In general people often react with fear to those types of sound, do these sounds affect emotion on-screen or even off screen?

I would propose an empirical research, probably some sort of experiment, using observations and possibly interviews to see how people react and so they can describe their feelings. One person at a time watching clips involving sound effects and maybe ambient sounds, but no music included.

Alternatively there could be an experiment on the sound track as a whole, to see to what extend this works however the specific effect of the sound effects will be much harder to prove this way.
A good third option may be the observation method described in the method section with some changes. When using audio editing software, one uses different tracks also called stems. If you keep these separated you can make different versions. This basically means that you would be able to utilize more variables. You could for example have a version with just sound effects, just ambience or music, or combining several.

There seem to be many research possibilities within the field of film sound. This has everything to do with the limited amount of research that has been done in area of scientific investigation. When research is being done it is often in disciplines as psychology or neurology, therefore the emphasis is usually on the recipient of the information. And of course everything is ultimately in the eye of the beholder. Yet we discovered that even within our own discipline, on both practical and theoretical fields, research questions can be asked. These have been illustrated and applied as much as possible in this research. A first attempt has been made to involve our field of study in this ongoing research. Hopefully many will follow this example and we will be able to find out more about what exactly the affects of sound are.
8. References

8.1 Books


Elsaesser, T. & Hagener, M. (2010) *Film theory, an introduction through the senses*. Taylor & Francis 2010. (pp. 129-149)


8.2 Scientific articles


8.3 Other sources


http://www.kau.se/utbildning/anmalan-antagning/behörighet (June 15th 2012)

8.4 Quotations


Appendix I

Shot breakdown

(Orchestral music starts)

1. A shot of the ocean.

2. ECU of a man lying on the beach, while waves crash into him. (Male voice: “There is one thing you should know about me”.)

3. LS, Hallway, man walking. (“I specialize in a very specific”…)

4. Fade to black.

5. CU, the man’s face covered in shadows. (“Type of security”…)

6. CU, a gun being loaded and cocked. (SFX: a click, then a sliding metal click)

7. Fade to black.

8. CU, two men talking. (“Subconscious security”…)

9. CU, hands opening a briefcase.

10. Fade to black.

11. CU, two men talking from reverse angle. (Other man: “You’re talking about dreams”)

12. ECU, a hand holding a chain, sliding of an armrest. (Hypnotizing kind of sound that builds up)

13. Fade to black, text “From Christopher Nolan” fade out, text “The director of the dark knight” fade out. (Sound builds up to a “musical explosion” loud music is played. The shots are punctuated by sound)

14. MLS, Man seen from behind walking toward building. (Voice: “Mister Cobb”…)

15. MCU, A girl talking to two men. (Old man: “has a job offer”…)

16. CU, the two men, reversed shot. (”he would like to discuss with you”)
17. MCU, A girl talking to two men. (Girl: “You’ve got a work placement?)

18. LS, train in the middle of the street, colliding with cars.

19. MCU, man stands beside car, covers up his face.

20. LS, Taxi cab driving backwards in a slip.

21. MCU, man in a cab holding a gun.

22. LS, Man running through building, shot is covered by shelves. (Cut musically punctuated)

23. MLS, Two men fighting. (Cut musically punctuated)

24. MLS, one of the men falls, but sideways. (Cut musically punctuated)

25. Fade to black. (Music drops out)

26. CU, two men talking. (Man:”Not exactly”) 

27. Fade to black.

28. CU, girl looking up. (Man: “we create”…)

29. LS, buildings come up and almost fold over. (“The world of the dream”…)

30. Fade to black.

31. MLS, man seen from behind stands in front of door while it’s opening. (“You bring the subject into the dream”…)

32. CU, girls face. (“And they”…)

33. LS, a beach seen through elevator door. (“Fill it with their secrets”)

34. Fade to black.

35. MCU, Man and girl talking in the street. (Girl: “Then you break in and steal it”. Man: “Well”…)

36. MCU, Man holding a briefcase, seen from behind turns around walks to the camera. (“It’s not strictly”…) (Music drops out)

37. MCU, Person in suit opening briefcase, head is not visible. (“Speaking legal”) (Music starts again)
38. MLS, Men in suits floating in air in a room.

39. CU, Man’s face seen through a window.

40. LS, Three men on rooftop, one of the men is holding another man.

41. Fade to black.

42. LS, Several people lying in a dark room. (Man: “It’s called Inception”)

43. LS, Man floating in mid air in what looks like an elevator shaft.

44. Fade to black.

45. ECU, of man’s face. (Man: “I’m ready”) (Music build up)

45. LS, avalanche. (music changes, less base heavy, more melody)

46. LS, two people in white suits abseiling of a snowy cliff.

47. CU, spinning top, hand grabs the top, camera moves to focus on man’s face. (Man: “I think I’ve found a way”…)

48. MCU, Man and woman standing intimately outside by the water. (“Home”…)

49. ECU, Man and woman’s faces. (“And this”…)

50. ECU, Man looking into a scope of a rifle. (“Last job”…)

51. LS, snowy building explodes, person jumps down from building.

52. MCU, Man talking. (“That’s how I get there”)

53. Man in white suit falling.

54. Fade to black, text “Leonardo DiCaprio”. (Aurally punctuated)

55. MLS, Three men sitting in chairs.

56. CU, woman. (Man: “Dreams”…)

57. MLS, girl walking outside, bridge appears. (“Feel real while we are in them”…)

58. LS, man walking between two narrow walls.

59. Fade to black.
60. CU, Girl sitting with man across from her. (“It’s only when we wake up”…)

61. CU, shot reverse, man talking to girl. (“That we realize that something was actually strange”)

62. LS, two men seen from above in stairwell.

63. Moving animation, text “your mind”.

64. CU, man with eyes closed, water splashes on him.

65. LS, Explosion, person in white suit leaves scene on a snowmobile.

66. MCU, Two men in front of cab, camera turns around them. (Man: “This is your responsibility” …)

67. LS, building in surrounded by mountains in snow. (“We are not prepared for this”)

68. MCU, Man holding woman.

69. ECU, Woman’s face, tears running from her eye.

70. ECU, Man’s face.

71. Moving animation, text “Is the scene of the crime”.

72. LS, car exploding, seen from above through window.

73. MLS, Two men walking into room, gun drawn.

74. MCU, Three men sitting in a vehicle, water hits them from behind.

75. MCU, man sitting on the floor of an elevator. (Man: “The dream is”…)

76. CU, Man holding a gun rolling over a table, runs away. (“Collapsing”)

77. LS, building in snowy environment explodes and collapses.

78. MLS, Man flies/falls through hallway. (Music starts to fade out)

79. Fade to black.

80. CU, man’s face. (Man: “I have it under control”)

81. LS, Man and girl sitting outside a restaurant, stuff starts flying around in background. (Music begins very base heavy)
82. Fade to black.

83. MLS, person on motorcycle in traffic, pulls out gun.

84. MLS, people floating in air while one man puts cables on them.

85. MLS, man seen from behind standing in a room between two staircases, water penetrates through the walls.

86. LS, buildings in ruins beside the ocean, camera moves, a building collapses and falls into the ocean. (SFX: stone crumbling and falling into the water)

87. Fade to black. (Music fades out)

88. CU, a man’s face looking to another man whose face is off-screen. (Young man: “I’d hate to see out of control”)

89. Moving animation, text “Inception” letters looks like mazes. (Music heavy on base starts)(Gun shots)

90. Fade to black. (Turns into background music) (Gun shots)

91. A man holding a rifle, other man walks on-screen gets out his large rifle and shoots. (Other man: “You mustn’t be afraid to dream a little bigger, darling”)(music fades out) (Metal sounds) (Gunshot turns into music)

92. Black, text “July 16”. (Music fades out fast)
Appendix II

Survey

There will be minimal risks in taking this survey and you will remain anonymous to all future readers of the research. The results will be discarded after the research is finished.

Instructions:

Please read fill in the questions on the first page before proceeding.

Open your web browser and fill in: http://www.youtube.com/watch?v=66TuSJo4dZM

This link will lead you to the Inception trailer on YouTube.

Play the clip without watching the clip, you may close your eyes and focus on listening to the sound.

Once you are done you can look at the first set of questions and answer them, when you finish you may go back to the clip.

Now you click on the full screen button and mute the clip, watch carefully.

After you finish watching the clip you may answer the second section.

When you are done you may un-mute the clip and watch it again. When you are done you may answer the third section of questions.

If you don’t understand one of the terms you can find a sheet explaining them at the bottom of the document.
Age:
Gender:
Occupation:

How often do you watch movies or TV-shows?
   o Every day
   o 3-6 days
   o 1-2 times a week
   o 1-5 times a month
   o Never

Clip without the visual

The clip is a little under two and a half minutes, did you perceive it to be that long or did it feel longer or shorter?
   o The actual length seems accurate to my experience
   o It seemed longer
   o It seemed shorter

How do you feel listening to the clip? (This pertains to the clip not your mood in general)
Happy sad
Unaffected excited
Scared other:

Rate your level of emotion after hearing the sound:
[Low] 0 1 2 3 4 5 6 7 8 9 10 [High]

How would you rate the level of cinematic realism? (If that was that was the real world, e.g. Star wars has cinematic realism but it isn’t realistic)
[Low] 0 1 2 3 4 5 6 7 8 9 10 [High]

How would you rate your level of immersion? I.e. did the story engage you?
[Low] 0 1 2 3 4 5 6 7 8 9 10 [High]
Clip without sound

The clip is a little under two and a half minutes, did you perceive it to be that long or did it feel longer or shorter?
  o The actual length seems accurate to my experience
  o It seemed longer
  o It seemed shorter

How do you feel watching the clip? Was it different than the sound fragment before?
Happy  Sad
Unaffected  Excited
Scared  Other:

Rate your level of emotion after watching the clip:
[Low]  0  1  2  3  4  5  6  7  8  9  10  [High]

How would you rate the level of cinematic realism? (If that was that was the real world, e.g. Star wars has cinematic realism but it isn’t realistic)
[Low]  0  1  2  3  4  5  6  7  8  9  10  [High]

How would you rate your level of immersion? I.e. did the story engage you?
[Low]  0  1  2  3  4  5  6  7  8  9  10  [High]

What did you like better: the clip with just the sound or the clip with just the visual?
  o The clip with just the sound
  o The clip with just the visual

Further thoughts:
Both sound and visual

The clip is a little under two and a half minutes, did you perceive it to be that long or did it feel longer or shorter?
  - The actual length seems accurate to my experience
  - It seemed longer
  - It seemed shorter

How do you feel watching the clip? Was it different than before?

Happy  Sad

Unaffected  Excited

Scared  Other:

Rate your level of emotion after watching the clip:

[Low] 0 1 2 3 4 5 6 7 8 9 10 [High]

How important do you think the sound is in combination with the visual?

[Not at all] 0 1 2 3 4 5 6 7 8 9 10 [Very important]

How would you rate the level of cinematic realism? (If that was that was the real world, e.g. Star wars has cinematic realism but it isn’t realistic)

[Low] 0 1 2 3 4 5 6 7 8 9 10 [High]

How would you rate your level of immersion? I.e. did the story engage you?

[Low] 0 1 2 3 4 5 6 7 8 9 10 [High]

Do you think the sound is compatible to the visual? Does it seem like the sound belongs to the clip?
  - No, it was ambiguous
  - Yes, it is compatible
  - No, they contradict each other
What do you think is the function of the music in the clip? You can choose multiple answers

- convey pace
- convey where you should focus attention
- convey the scope of a film
- convey the general mood of a film
- convey the quality and size of a space & placement in time
- convey a sense of energy
- convey the overall perspective or message intended
- convey the internal life, thoughts, and feelings of a character
- convey narrative structure
- mismatch conditions between the audio and visual components

What do you think is the function of the sounds (other than music and voice) in the clip? You can choose multiple answers

- convey pace
- convey where you should focus attention
- convey the scope of a film
- convey the general mood of a film
- convey the quality and size of a space & placement in time
- convey a sense of energy
- convey the overall perspective or message intended
- convey the internal life, thoughts, and feelings of a character
- convey narrative structure
- mismatch conditions between the audio and visual components

What do you think the function of the dialogue is in the clip? You can choose multiple answers

- convey the general mood of a film
- convey pace
- convey where you should focus attention
- convey the scope of a film
- convey the quality and size of a space & placement in time
- convey a sense of energy
- convey the overall perspective or message intended
- convey the internal life, thoughts, and feelings of a character
- convey narrative
- mismatch conditions between the audio and visual components

Further thoughts:
Terms explained

1. Music can convey the general mood of a film

Music can give clues about genre, is the film supposed to be scary, comedic, romantic etc. This greatly depends on the level of ambiguity in the image. The more ambiguous a scene, the more influence music has.

A fight scene will still be perceived as violent even with a funny tune in the background, this because of the lack of ambiguity in the clip. This would be a case of audiovisual mismatch; this will be discussed later on.

2. Music can convey the scope of a film

Music can communicate whether the film is an epic drama, such as “Star wars” (1977) or a more film on a more personal scale.

3. Music can convey the quality and size of a space & placement in time

Music can make small rooms or artificial space seem bigger and it can add realism. Music can also “authenticate” the era by using music from a specific period of time.

4. Music can convey a sense of energy

The choice of music can influence the energy that is perceived. For example a slow classical piece in a chase scene can make the scene seem slower.

5. Music can convey the overall perspective or message intended by the director.

This can relate to both events and characters. An example of how this works it that the same action can be depicted in different ways only by changing the musical score.
6. Music can convey the internal life, thoughts, and feelings of a character

The unspoken thoughts can be communicated through music, it can convey character.

Leitmotiv is a commonly used example of this function, it is a theme that represents a person and is usually played whenever the character is on screen.

7. Music can convey narrative structure

Music can also clarify or even establish a sense of order through the perception of a formal structure.

8. Music can convey messages about where in the frame the audience should focus attention

According to research music either has a negative or positive connotation. When music with a certain connotation is combined with an ambiguous scene, the memory of visual objects is greatly affected by the music.

Another focusing device is “temporal coincidence” where specific moments in the music draw attention to specific moments that is simultaneously happening in the visual.

9. Music can readily convey pace

If patterns are created in the musical score using music, sound effects, silence etc. and later changed, a film can look like it is speeding up or slowing down. Even though in reality it isn’t.

10. Music can mismatch conditions between the audio and visual components

As mentioned before mismatch is when a genres clash, there are certain instances when this is used. For example to convey irony it makes the audience think, consciously or unconsciously what the intended meaning is.

Silence or the lack of sound can also create a mismatch, when sound is expected and it isn’t there it makes the audience reflect on why it isn’t there what is the meaning.
Changes made after the examination

After the examination a couple of adjustments were made, some relatively small and some quite large. When I received the list I went through all the items on the list piece by piece. When I identified the problem and found it in my text I marked it. After I had found all the problems I set up a plan. The first thing I did was work a little on the language, punctuation and grammar. The same day I found some people willing to participate in the survey and sent out the surveys to them. While I waited to get the surveys back I decided to work on some of the easier items on the list. I made sure to put in more references in the places that were missing some. I rewrote some sections that did not make sense, and clarified them. I wrote down what the moods were that I expected the subjects to feel and so on. Meanwhile I had received some emails from my examiner, I looked at those and implemented the suggested changes.

When I got all the surveys back I put all the information into SPSS and analyzed and wrote down the results. Because of the new results I needed to rewrite the result, analysis and sections, I also tied the analysis section in with the theories some more. After I was done with that I looked at the remaining list items and worked on specifying the research questions. Finally I read the entire thesis one more time and tried to fix all the misspelled words and confusing sentences.