

Interactions between music, sound design and dialogues in movies and series

Interactions entre la musique, le design sonore et les dialogues dans les films et les séries

Par
Jérôme Yven

Université de Montréal, faculté de musique

Mémoire présenté en vue de l'obtention du grade de maîtrise en composition à l'image du master InMICS
(International Master In Composition for Screen)

Mai 2022

© Jérôme Yven

Université de Montréal

Ce mémoire intitulé

Interactions between music, sound design and dialogues in movies and series

Interactions entre la musique, le design sonore et les dialogues dans les films et les séries

Présenté par
Jérôme Yven

A été évalué par un jury composé des personnes suivantes

Pierre Michaud
Président

François-Xavier-Dupas
Directeur de recherche

Francis Perron
Membre

Résumé/Summary

Ce mémoire s'intéresse aux interactions et aux relations dans les films et les séries entre les trois éléments sonores principaux que sont la musique, le design sonore et les dialogues. Au travers d'analyses et de réflexions, nous nous intéresseront aux particularités propres à ces domaines sonores, aux frontières floues entretenues entre eux, ainsi qu'à la façon dont ces éléments peuvent être rencontrés, imbriqués ou mélangés afin de servir l'histoire racontée par le média du divertissement audiovisuel.

This thesis is a reflexion about interactions between all different aspects of the sound in movies and series that are music, sound design and dialogues. Through analysis and hypothesis, we will be focusing on the particularities of each of these sonic categories, on the boundaries that they share and in the way they meet each others, they blend and they merge in order to serve the narration being told by the audiovisual media.

Mots clés/Key words

Musique – Design Sonore – Dialogues – Films – Séries – Audiovisuel - Media

Music – Sound Design – Dialogues – Movies - Series – Audiovisual – Media

Acknowledgments

I would like to thank my teachers of the InMICS Master, and especially my research director François-Xavier Dupas for helping me take a major turning point in my evolution as a composer by the discovery of his amazing work and his incredible teachings, Gilles Alonzo for the unbelievable richness of his teaching, Pierre Adenot for his kindness and his wonderful artistic personality, Thibault Cohade for his help in discovering new music territories, my orchestration teachers Luca Antignani and François-Hugues Leclair and my composition teachers Vincent Laurin Pratte and Tristan Capacchione.

I would also like to thank Isabelle Replumaz for creating this master with such energy and joy, and the partners from the film festival in Ghent (Thomas Van Parys), Aubagne (Samya Papasoff) and Krakow (Mickael Lyczek).

Finally, I would like to thank Roser Graell, Véronique Lefebvre, Andrada-Erika Badiu, and all my teammates of the InMICS program 2020-2022.

Table of contents

Résumé/Summary	III
Mots clés/Key words	III
Acknowledgments	IV
Table of contents	V
Introduction	VII
I. Music and sound design	IX
<i>A. Three listening modes</i>	<i>IX</i>
<i>B. Chernobyl</i>	<i>XI</i>
1. Presentation.....	<i>XI</i>
2. Composition Process	<i>XII</i>
<i>C. Analysis</i>	<i>XIV</i>
1. Scene 1	<i>XIV</i>
2. Scene 2	<i>XVII</i>
II. Interactions between music and dialogues	XXI
<i>A. Neurosciences researches about categorization of sounds</i>	<i>XXI</i>
<i>B. Musicals and released soundtrack</i>	<i>XXIV</i>
<i>C. Effects on dialogues</i>	<i>XXIX</i>
Conclusion	XXXIII
Bibliography	XXXV
Table of illustration	XXXVI
Abstract	XXXIX

Introduction

The study of sounds in movies and series is a wide field of research, as it can refer to a various number of aspects, each one of them being very different from the others: analysis of the audio of a movie can focus on either film scoring, sound designs or dialogues. Those three elements are studied in film schools all around the world, and there is a clear distinction made between them. Those studies lead up to three different fields of work as either composer, sound designer or voice editor, having within themselves specialties that can't be found in the other fields, as it exists a wide variety of film scoring and sound design techniques for example. Moreover, we can most of the time establish some kind of hierarchy within audio for movies and series, dialogues being the most important element to be heard correctly, followed by sound design and music composition.

However, there are movies and series that redefine the way we conceive and interpret audio for visual media. The tripartite separation between music, sound design and dialogues tend to blur in some works as we will see later. There could be for example a “blurred” zone between music and sound design, or between music and dialogues, and some treatments on the audio itself. All those aspects are adding new elements to what we are used to experience as a spectator.

We can ask ourselves: why does those interactions happen in audiovisual works, and what is the goal behind them? Pushing the limits of something established could be one of the reasons. Finally, they can emanate from a bold artistic intuition, a desire to create something new, in order to find new exciting ways of telling stories. 0

Today, crossing boundaries between those three audio elements have never been easier for creators: all of these interactions can be done on a single computer within the same digital audio workstation (DAW), therefore blurring a bit more the frontiers between music, sound design and dialogues.

The purpose of this essay is to study some exceptional interactions between music, sound design and dialogues in different movies and series, and how they are adding depth and powerful connections to the storytelling. To do so, we will first analyze two extracts from the miniseries Chernobyl. Then, we will explore the connections between music and language with the help of neuroscientific research focusing on the ways the brain categorizes sounds, music and language, before showing the similarities that exist between music and language by studying their interactions in musicals and released soundtracks. Finally, we will see how manipulating dialogues has an impact on the storytelling and the narration, and how it can be a powerful tool to tell stories.

I. Music and sound design

A. Three listening modes

When watching audiovisual content, the way we listen to it isn't linear. As in everyday life, we make a clear distinction between hearing and listening: I can be hearing roadworks in a street near me, while however focusing my attention on the music I'm listening to in my headphones. Our listening is constantly shifting between different levels, and it is interesting to understand why as it also appears to happen to the way we're listening while watching a movie. Michel Chion, a film theorist and composer of experimental music, has listed three listening modes in his book *Audio Vision: Sound on Screen*.

The first one is called causal listening. It consists of "listening to a sound in order to gather information about its cause (or source)"¹. If I see a man talking on screen in the street with a car passing by behind him, causal listening will allow me to assign sounds from their sources: the human voice will be associated with the man and the motor sound with the car. It is interesting to notice that causal listening can also be perceived in different levels within itself. We can recognize the general nature of a sound (a mechanic engine turning on for example) without really identifying which object it emanates from: "Even without identifying the source in the sense of the nature of the causal object, we can still follow with precision the causal history of the sound itself. For example, we can trace the evolution of a scraping noise (accelerating, rapid, slowing down, etc.) and sense changes in pressure, speed, and amplitude

¹ Michel Chion, "Audio-Vision: Sound on Screen", *The Three Listening Modes*, Columbia University Press, 1994, p. 25

without having any idea of what is scraping against what”². This different level within causal listening will be very helpful later, for our analysis of Chernobyl.

The second one is called semantic listening. It “refers to a code or a language to interpret a message: spoken language of course, as well as Morse and other such codes”³. This is a more complex listening mode as it may vary a lot between individuals because of the linguistic barrier. The author underlines the fact that one single sound can be heard both in causal and semantic listening: “we hear at once what someone says and how they say it. In a sense, causal listening to a voice is to listening to it semantically as perception of the handwriting of a written text is to reading it”⁴.

The last one is called reduced listening, which is an expression borrowed to Pierre Schaeffer who categorized it as “the listening mode that focuses on the traits of the sound itself, independent of its cause and of its meaning”⁵. It is no wonder that Michel Chion adds this listening mode to the film soundtrack analysis as he was a student of Schaeffer, who was, along with Pierre Henry, one of the most famous composer of concrete music, a musical genre that focuses on the nature of the sound itself rather than on its source: “Variations for a door and a sigh” (Pierre Henry, 1969) is one good example of a sound that we associate with an everyday object, but used in this context only for its acoustic characteristics.

We have seen that the difference between hearing and listening comes with an understanding of the way we categorize our listening between causal,

² *Ibid.* p. 27

³ *Ibid.* p. 28

⁴ *Ibid.* p. 28

⁵ *Ibid.* p. 29

semantic, and reduced modes. Far from being linear, our listening is a more complex mechanism that is shifting throughout our perceptions of the world, and by extension our experience of listening to an audiovisual content. In order to develop this aspect, we are going to analyze examples of the soundtrack from the miniseries *Chernobyl* (2019) which plays with the way we are listening to music and sound design.

B. Chernobyl

1. Presentation

When considering interactions between music and sound design, one can wonder: as there is a clear distinction between musical instruments-that were used during thousands of years for musical purposes- and sound (both sound design and foley recording) that is necessary for the comprehension of a movie and to give a sense of realism, why should we search for similarities between those two different aspects, both in their role and nature ? What can be the point of looking for similarities between those two audio elements when they are created by different professionals: sound designers and composers? How can these interactions be interesting to study, and what can their artistic and emotional impact on the spectators be?

The line that separates musical information from sound has been questioned and crossed by composers and sound designers throughout cinema history. The « Mickey Mousing » film scoring technique comes to mind, which was used extensively in early Walt Disney animated films from the 1930's and 1940's,

and in which the music acts as a sound design element: if a character falls down from the stairs, the music could then play a descending chromatic scale synchronized with the impact of the character on each stair, replacing the need for a more natural impact sound effect. One example of this technique can be found in the *The Skeleton Dance* from the *Silly Symphonies* (1929) in which every time we see a skeleton walking in the grass, we can hear strings playing a note in unison synchronized to the picture.

In order to analyze the interactions between music and sound design, we will be focusing on the miniseries *Chernobyl* (2019) and the composition process from the composer that used field recordings as the starting material for the score of the miniseries, thus making the distinction between intra and extra diegetic elements less noticeable.

This miniseries was produced by HBO in 2019 and officially released the 6th of May 2019. It was created and written by Craig Mazin who has been the director for *Hangover Part II* (2011) and *Identity Thief* (2013). He has been working on *Chernobyl* with director Johan Renck who worked as a director for television series *Breaking Bad* (season 3, 2010) and *Vikings* (season 1, 2013). Icelandic composer and academy award winner Hildur Gudnadottir was asked to compose the soundtrack for the miniseries. She is known for her work for *Sicario* (2018) and *Joker* (2019). The audio team also included score producer and field recording engineer Chris Watson, supervising sound editor Stefan Henrix and sound designer Joe Beal.

2. Composition Process

The composition process of Hildur Gudnadottir for *Chernobyl* is directly linked to the making of the miniseries. *Chernobyl* was shot in Lithuania, “in

a power plant that has been decommissioned”⁶, and “before [the film crew] went there to shoot, [the composer] went there to record sounds of the power plant”⁷. This is part of a pre-production aspect of the making of the score, Hildur Gudnadottir went to the place where the scenes were going to be shot in order to “experience what it feels to be inside a power plant”⁸, and going there with field recording engineer Chris Watson allowed them to capture this feeling on site. We can draw a line between what Hildur Gudnadottir and Chris Watson recorded in the power plant and the concept of reduced listening from Michel Chion: they listened to the sounds of the power plant for their acoustic nature and potentials for creating a musical score in production,. Once in production, the composer “wrote the score from those recordings, every single sound in the score is made from these power plant recordings”⁹. Therefore there is a porosity between what the composer is doing with the field recording of the power plant and the work of the sound designers that are using identical material for the audio design of the series. And finally, all those elements are combined in post-production during the editing process, and the blurred lines between the score produced from field recordings and the sound design are being explored during the editing process. This part of Chernobyl also gives a very important role to the editors Jinx Godfrey and Simon Smith in the way they choose to use or put emphasis on sound design or music score in different scenes and contexts.

⁶ *Score: The Podcast*, “Chernobyl composer created entire haunting score from real power plant sounds”, 2019, 65 min

⁷ *Ibid.*

⁸ *Ibid.*

⁹ *Ibid.*

C. Analysis

We will analyze two extracts from Chernobyl where the soundtrack is crossing the line between music and sound design. Because of the nature of the composition process from Hildur Gudnadottir, the interaction between sound and music is always part of the soundtrack, and we will see that this choice is justified by the narration and what the story is telling to the spectators.

1. Scene 1

The first chosen scene is from the first episode “1:23:45”. In this episode, we are following the events that took place during the night when the reactor number four of the nuclear power plant of Chernobyl exploded.

The scene begins with the explosion of one of the reactors of the power plant, which is seen at night through the windows of an apartment in Pripyat, the nearest city from Chernobyl. We are in the apartment of a woman named Lyudmila. As a spectator, we first see the explosion in the distance, and a few seconds later the woman is being surprised by the shockwave that makes the building shaking. At 00:07:34, after the shockwave, a soft airy pad is heard as the woman and her husband, a fireman who will be called later in Chernobyl to extinguish the fire, are trying to figure out what happened. This pad is constructed around the B flat note, with the addition of other noises as if someone with a metal pipe was scratching on a plaster wall. We don't know if this is what the composer recorded in the power plant in Lithuania, but we can identify the general nature of this sound: a pad made out from industrial sounds.



Figure 1: Analysis of the score structure from the first episode 1/2

On top of the airy pad in B flat, another music layer is heard quickly after, in the form of a second airy pad without added noises, and with a root of E. It is difficult to know if those root notes were already present in the original recordings of the power plant, or if they have been treated in postproduction in order to enhance some specific frequencies. This scene comes at the very beginning of the miniseries, and therefore we are introduced to the musical universe of the show when the power plant explodes by hearing musical sounds related to the place where everything happens, therefore enhancing the spectator's audio experience. Afterwards, a voice is heard in the distance, with reverberation and filter of the high frequencies, calling for "Camarad Diatlov". This voice makes the transition from Pripyat to the inside of the power plant.

At 00:08:02, we are changing location and we now see the inside of the control room of the nuclear power plant, in which workers from the night shift are trying to figure out what happened. The music stays the same, holding the pads that were heard before, and the voice is still calling for “Diatlov” with the effects on the sound. But now, we also hear the sound design of the control room consisting of the room tone, sound of dust falling from the ceiling, neon light bulbs buzzing, and a rumble in the low frequencies which we cannot define as either sound design of the place or musical element. This is the starting point of the “blurred” zone in the sound of Chernobyl: even if we can notice musical elements (airy pads in B and E), there are other sounds that we are not able to categorize, as both the music and sound design are originating from the same source. This blending of sound design and music in the miniseries is creating “blurred” sound zones in which those two elements can be perceived as one, as our causal listening is acting on different levels : hearing the sound effects from the scene, hearing the musical elements that are recognizable, and trying to categorize the other sounds or trying to focus on the sounds that were heard before but that are now blurred with the other elements.

At 00:08:12, the filter and the reverberation on the voice are turned off to symbolize the fact that the man called Diatlov isn’t shocked anymore and is able to hear correctly what his colleagues are saying. Once the effects are turned off, other sound elements are heard: we can distinguish three different alarms coming from the control room, while the music and the room tone are still present in the background. At this point we cannot hear very clearly the airy pads from the beginning as they seem to have been blurred with the sound effects of the scenes, and maybe disappearing in those. At this moment of the scene, our causal listening might prioritize what we hear as a whole ensemble of audio sources coming from what we see (the control room) and not focusing on the music elements anymore.

At 00:09:41, the scene ends with Diatlov leaving the control room and going into a hallway, the general ambiance of the control room made from music and sound effects fading out to leave space for another soundscape. This technique is frequently used in the series: when a character leaves a room (and so does the camera), the « blurred zone » fades out, enhancing the fact that the music belongs to the place we see on screen.

With this first scene, we saw that our causal listening is constantly challenged while listening to the soundtrack of this miniseries because of the composition process which is directly linked to the narration. We will now analyze a second extract from the first episode to study how the blending between music and sound design is used later in the show.

2. Scene 2

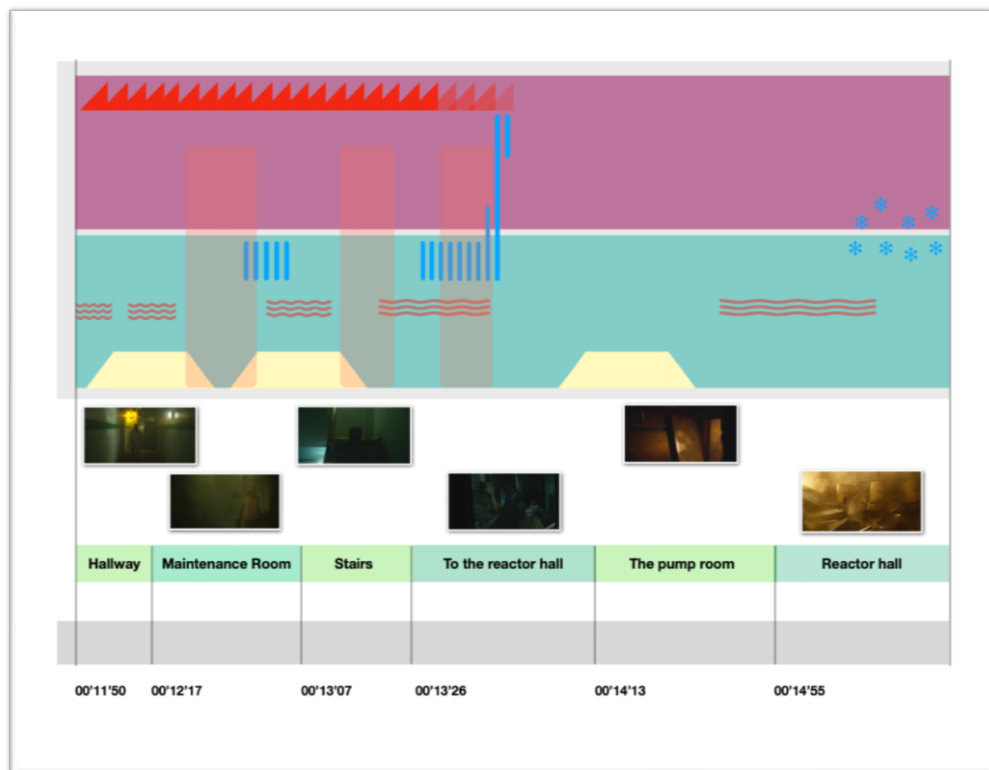


Figure 2: Analysis of the score structure from the first episode 2/2

When the scene starts at 00:11:50, we are following a man walking in a hallway a few minutes after the explosion. An alarm is being heard at the beginning of the scene until 00:13:14, accompanied with a low room tone. A high pitched rhythmic and saturated element can be heard during this scene. As we can assume that the composer manipulated the sounds from the recordings in order to create musical tones by enhancing frequencies, we can also assume that this effect of saturation on the sound is a deliberate choice from Hildur Gudnadottir to create a “nuclear” atmosphere for the score: the high level of radiation and the great danger that such a catastrophe is producing for human lives can lead to the use of an effect that adds harshness and a sense of brutality to the sound, hence the choice of the saturation. This sound is more likely to be categorized as a musical element by our causal listening because its regular pulse and intonation gives us a sense of a synthesized sound that is being used to add more tension to this scene rather than a sound coming from the power plant. Therefore, we can associate it to some kind of mechanical noise while at the same time being part of the music score. Another musical element is added, a high pitched pad shifting from the notes C to C#. This is an airy and saturated pad that sounds like steam blowing through an open pipe. This element is recurrent throughout the scene and can be perceived as a clear musical landmark because it is repeated almost identically in other places of the power plant. After that, once the man enters one of the maintenance room to ask his colleagues for a dosimeter, the high pad fades out of the general ambiance, in such a way that we can't define a precise moment when the music stops. Later in this scene, the previous elements that we heard come back again with the addition of a new musical sound: a low percussion, which comes in a regular way. Once again, we could wonder whether this sound is from the raw recording or if it's more the result of a musical processing, but its regular

rhythm fading in instead of being heard directly when the character enters the room leads to think that it is a choice from the composer.

Later, when the man is going from the maintenance room to the stairs, the airy pad is once again fading out very slowly, and even a precise listening of the transition cannot help us to define whether some musical elements are still present or not. This is the second « blurred zone » of the scene. Transitioning to a corridor leading to the reactor hall, the high pitched rhythmic and saturated element is coming back again in a variation. In this place, we are hearing a low impact coming on and off very regularly as we see the characters walking. We don't know where this sound is coming from, and its sound characteristics can make us think about the low percussion heard previously in the maintenance room, and it also adds tension to this new environment in which the characters are going. The first beats are heard with a new iteration of the saturated element. The first seven "beats" are heard very identically: the sound is filtered and repeated the same way. The eighth beat is repeated with a subtle delay, and its high frequencies are more opened. This is the moment when the character on screen is wondering where this sound might be coming from, he didn't seem to care about this information before. The two last beats are heard completely unfiltered, and we finally see where this sound comes from: one character was smashing a door with an extinguisher in order to open it. This is a moment when the blurred zone was both active for the people watching the episode and for the character on screen, this shift of the causal listening is happening both on screen and off screen (for the spectators) adding a strong sense of identification to the character from the viewers experiencing the scene. The alarm that we were hearing from the beginning fades out at this moment.

The next scene, happening in the pump room, is introduced by the sound of the airy pad on a high C, which will fade out halfway through the scene. As

the character nears the reactor room from where most of the radiations are emitted, the saturated element comes again with variations in its rhythm. Right before the character enters the final place of the scene where the reactor exploded, a final element is introduced: a rhythmic glitchy sound similar to crackles slowly fades in before being heard loudly as the man enters the room, where the radiations are the strongest. The sound is clearly linked to what we might see on screen, but this is not a sound that could be heard loudly by the characters, so it is a mixing choice that can make us interpret it as a musical element.

An analysis of the soundtrack from Chernobyl allows us to underline the value of putting the emphasis on the narration as a composer. Here, we are the witness of a nuclear catastrophe in Ukraine. This setting gave birth to a unique composition process: going to a nuclear power plant for a field recording trip in order to gather sounds that will be the starting material for the music score. The resulting score has a strong and powerful effect on the listener of the show, adding a great sense of immersion and identification that we certainly wouldn't have experienced if the score was recorded with a string ensemble, because the raw sound doesn't come from the same source (a defined instrument on one hand, on the other a precise sound directly linked to what we see on screen). The sense of immersion is also enhanced by the way our causal listening is challenged while watching the show : the blur between music and sound design is sometimes heard on two sides, both by the spectators and by the characters of the scene, thus making us feel the same way as the characters. Chernobyl is a strong and powerful example on how interactions between sound design and music can lead to exciting new experiences as a spectator. We are now going to see one another strong interactions that is interesting to study as a spectator, which is the one between music and language.

II. Interactions between music and dialogues

It is quite common to hear that music is a language like any other: it clearly has its own notation properties along with a unique vocabulary and complex nuances and expressions. But the linguistic value of the human language makes it unique: as music can be perceived as universal, the language can't, or it would be the result of a general convention. But those two elements are nonetheless sharing a lot in common, and we will first determine why with the help of neurosciences before coming back to specific examples from the audiovisual repertoire.

A. Neurosciences researches about categorization of sounds

The studies about understanding the brain mechanisms to process sound are quite new, and a lot is still to be discovered. In their article “Brain Optimized extraction of complex sound features that drive continuous auditory perception”, researchers Berezutskaya J, Freudenburg ZV, Gu“c,lu” U, van Gerven MAJ and Ramsey NF have established that “a lot remains unknowns regarding how the human brain processes sound in a naturalistic setting, for example when talking to a friend or watching a movie”¹⁰, and that “we still lack the comprehensive understanding of the brain mechanisms that support continuous auditory processing”¹¹. After establishing this observation, the goal of the researchers was to construct a data-driven neural model that would help define a hierarchy of perception in the brain, and to “identify the features that different cortical regions extract from the incoming sound

¹⁰ Julia Berezutskaya, “*Brain-optimized extraction of complex sound features that drive continuous auditory perception*”, PLOS Computational Biology, 2020, p. 1

¹¹

signal and to understand how they are transformed into high-level representations specific to sound type (speech, music, noise etc...)”¹¹. Their result “indicates a hierarchical processing structure in the brain during auditory perception that is generic across individuals and experimental material”¹². This model “can be used to test existing theoretical assumptions and generate new insight about the complex mechanisms of auditory perception in the brain”¹³. This experiment allows us to understand the correlation between what we could name a high level representation of sound which could be related to the three mode of listening of Michel Chion, understanding what we hear, and a low level representation which would be associated with neuroscientific researches regarding why we hear those sounds and how our brain categorizes them.

Another study will help us underline the importance of understanding how the sound is processed by the human brain, as it could be processing sound in the same way that it processes images. This discovery was made by Max Riesenhuber, professor of neuroscience at Georgetown University and coauthor of the article published in Neuron. Researchers funded by the National Science Foundation were trying to understand “how the brain categorizes visual objects and found that at least two distinct regions of the brain were involved. One region, in the visual cortex, encoded images, while a region in the prefrontal cortex signaled their category membership”¹⁴. After discovering this, researchers were wondering “whether the same processes underlie

¹¹ *Ibid.* p. 2

¹² *Ibid.* p. 3

¹³ *Ibid.* p. 17

¹⁴ National Science Foundation, “*How does the brain learn categorization for sound? The same way it does for images.*”, 2018

16

categorization of auditory cues?”¹⁶. In order to understand this, they made people listen to sounds without labeling them. This happened in “the lower level auditory cortex, which is responsible for representing sound but does not appear to give it any meaning or significance”¹⁵. Then, the same people were asked to judge which category the sounds belonged to, which happened in the prefrontal cortex, thus allowing them to understand the “similar process for visual and auditory categorization [that] promises important advances for how we understand learning”¹⁸. This study helps us making a neuroscientific junction with the modes of listening from Michel Chion: the causal and semantic listening would therefore be associated with the prefrontal cortex that adds meaning to a sound, whereas reduced listening would be associated with the lower level auditory cortex, which is not giving meaning or significance to the sound. With this information, we are moving one step further between the high and lower level representations of the sound by the brain, in a context where sound is acting like visual representation.

One last study clarifies the ways the brain is perceiving sounds. The article “The Brain in Singing and Language” by Valerie L. Trollinger, professor at Kutztown University of Pennsylvania is providing us with a good insight on the interactions between music and language in the brain. The author tells us that the available brain researches gives us information about “how the brain works while processing and producing singing and language”¹⁶, and that there is a “strong relationship between musical perception and processing with language”¹⁷. She underlines that “most of the research indicates that language

¹⁵ *Ibid.*

¹⁸ *Ibid.*

¹⁶ Valerie L. Trollinger, *The Brain in Singing and Language*, General Music Today, 2010, p.20

¹⁷ *Ibid.*

21

processors in our brains also help us process harmony in musical relationship (as) two language areas, Broca's area (generally associated with processing grammar for language) and Wernicke's area (generally associated with vocabulary), are also engaged in musical processing"²¹. The article "The Brain in Singing and Language" helps us identify a precise location in the brain to study interactions between language and music, and this information is good to bear in mind for our interest in language and music in audiovisual context, as we now know that the separation between those two elements is less clear, and that these information are treated in the same area of the human brain.

A lot is still to be discovered concerning how the brain processes sound, but we already know enough information to argue that music and language are linked together in the human brain and that we can analyze it through two different aspects: the high and low levels of representations of sound. Challenges of researches to come will deal with connecting even more those two levels, in order to finally explain how the brain is categorizing every sound and where it happens. Once we know that, we can add an audiovisual context to these information by focusing on the examples from the musicals and the released soundtracks.

B. Musicals and released soundtrack

One of most common interaction between music and dialogues in movies are found in musicals: there are constant jumps or cuts between scenes where dialogues are spoken by the actors on screen, and scenes where the music starts and the dialogues are becoming songs. This shift, an integral part of what makes a musical, could be analyzed as a jump between semantic and causal listening.

On one hand, words and sentences pronounced by the speakers are being heard by spectators on a semantic level, the main goal of it being the comprehension of each sentences to understand the story being told. On the other hand, the addition of music, that could be intra or extra diegetic, adds another level of listening (if the music is extra diegetic, the spectators have to listen to it as sound being part of the story but not playing from inside the scene), that makes the role of the dialog more complex by giving it another nature : a song.

This other nature of the dialogues comes with a shift during the listening of a movie or series. As we saw in the first part, the way we listen to an audiovisual content isn't linear, we are always moving from one listening mode to the other. Michel Chion is adding later in the *Audio Vision* more information about what he calls units of sounds that might change in nature during our listening of the movie. For the author, "such units-sentences, noises, musical themes, "cells" of sound- are exactly the same type as in everyday experience, and we identify them accordingly to criteria specific to the different types of sounds heard. If the scene has dialogue, our hearing analyzes the vocal flow into sentences, words- hence, linguistic units [...] In other words, we hear as usual, in units not specific to cinema, that depend entirely on the type of sound and the chosen level of listening (semantic, causal, reduced)"¹⁸. This means that while watching a movie, we can hear dialogues as a semantic unit at one moment of the movie, and we can consider it as another unit, a musical one, later. It is likely to think that this shift might happen quite easily in the brain as we saw that it is processed in the same place.

There is one example of musical that definitely played with this notion of level of listening, and this musical is *Les Parapluies de Cherbourg* (1964), directed by Jacques Demy- who is also known for directing other musicals such as *Les*

¹⁸ Michel Chion, "Audio-Vision: Sound on Screen", *The Three Listening Modes*, Columbia University Press, 1994, p. 45

Demoiselles de Rochefort (1967) and Peau d'Ane (1970)- and with an original score by Michel Legrand. The particularity of this musical is that every single dialogue is sung by the characters, as if the whole usual script for a nonmusical movie was transposed into a musical. Hence, the jumps between semantic and causal levels of listening transform into one single level where the spectator is invited to follow the whole story through a music spectrum. It is interesting to notice that with this particular semantic approach, there is no clear hierarchy between musical elements within themselves, but the classical organization for audio in cinema where music is subordinate to the picture is reversed here, the music is guiding the way the movie is shot. Within the musical elements, it is interesting to notice that sometimes the sentences rimes, and some other times there are no clear junctions in term of rhythm or melody between them. But we still have moments where a classical song organization is used: verse and chorus alternating, rimes, melody that comes back frequently.

This musical is unique with its use of dialogues in a linear way during the whole movie in opposition to other musicals where there is a non-linear following of the dialogue because of the cuts in the narration by the addition of songs. Moreover, the prosodic way of singing the everyday life makes the junction between a “standard” use of dialogue and music.

The example of the musicals allows us to consider dialogues as units, segments of sounds, that can evolve in the way we listen to them throughout a movie. By shifting between spoken language and songs, dialogues are playing with the way we interpret the words that are being said on screen in a context of semantic listening that stays every time we hear words (as it is the first information that spectators want to hear in order to gather information about

what is happening) in relation to causal listening (concerning the fact of knowing where the music is coming from, and who is singing the song).

Another interesting interaction between music and dialogue can be found in the released albums from movie soundtracks, especially from musicals. Those albums usually feature most of the music and songs present in the movie. On some tracks, we can encounter dialogues directly taken from the movie that don't actually belong to the music score, as they can help understanding the context of the beginning of the song or introducing the arrival of the song. We will focus on two examples that used dialogues in their soundtracks.

The first one is from the released soundtrack of the movie *Moulin Rouge* (2001) directed by Baz Luhrmann- who also directed *Romeo And Juliet* (1996) and *The Great Gatsby* (2013)- and with an original score by Craig Armstrong. The first track of the album "Nature Boy" starts with strings in F minor fading in, from where we hear the voice of Ewan McGregor (who plays Christian, a writer living in Paris) narrating the following sentence with a trembling voice: "This story is about love. The woman I loved is dead". The song starts right after that sentence, as David Bowie is singing the first words of the song "There was a boy...". In this example we can perceive two different units: the first one being the narrator speaking and including the sound of the voice to a dialog line, and the second one being the voice of the singer of the song. Although being used for another goal (setting the context of the story), the narrator becomes part of the musical content. It is a good example of the importance allocated to dialogues even in this purely musical context as those words are the first heard in the soundtrack, there are therefore a lot of emphasis put into them. Furthermore, it is also a dialog line that closes the soundtrack: the last track is a variation of the first one, also called "Nature Boy", but produced by the English band Massive Attack. At the very end of the song, when David Bowie stops singing and the music slowly fades out, we can hear the whisper

from Nicole Kidman (who plays Satine, one of the dancers of the Moulin Rouge) saying “I love you”, thus making the junction with the very beginning of the soundtrack.

The second example is from the movie *La La Land* (2016) directed by Damien Chazelle -also known for *Whiplash* (2014) and *First Man* (2018)- with an original score by Justin Horwitz. The song that we will be focusing on is the twelfth one on the released album, titled “Audition (The Fools Who Dream)”. This song appears when the character Mia (interpreted by Emma Stone) is telling a story during a movie audition. The beginning of the song in the soundtrack starts the same way as in the movie, with Mia telling the story about her aunt that lived in Paris. The dialog begins by “My aunt used to live in Paris. I remember, she used to come home and tell us these stories about being abroad. And I remember she told us that she jumped into the river once, barefoot. She smiled.”. Then, Mia starts singing the rest of the story with the next sentence “Leapt, without looking...”, all by herself at the beginning, and then joined by the piano and the whole orchestra. In this extract, there is only one “unit” heard concerning the voice as there is only one person talking and singing, a fact our brain easily recognizes. Once again, this dialog part at the beginning of the song is used for setting the context, in this case the story about the aunt of the character. This shift between spoken language and song brings the listener to another level of listening, as the background music that accompanies the singer isn’t an intra diegetic element, so we move from a realistic point of view (the audition) to a space where we don’t know what is real or not.

Musicals are a great example to illustrate how words can be carrying different informations and being interpreted in various manners according to the

context. In a non-linear way of listening to a movie, we interpret dialogues as units that can change nature by becoming songs, a shift that is easily made by our brain as they are familiar information. We saw that dialogues are taking more space in the soundtracks by being recognized as an element that could fit in a movie released soundtrack. We are now going to see how effects applied to the dialogues can add to the storytelling.

C. Effects on dialogues

When watching and listening to a movie, the most common approach about handling dialogues is the what could be called a “naturalistic treatment” of the voice : as words are the easiest way to tell stories with audio, it is important that the spectator could hear everything clearly in order to understand what is happening. The semantic listening described by Michel Chion is intrinsically linked to language and spectators expect dialogues to be clear and to add information about the story being told. In that sense, the intelligibility of the words is crucial for dialogues to be used effectively. Even with movies shot in other countries with other languages, subtitles are here to give us the necessary information.

In this regard, playing with the perception of dialogue is a way to play with the story itself and the narration, therefore impacting the movie experience of the spectator. As there seems to be a distinct hierarchy in audio for movies, considering dialogue as the element that should be audible and understandable at all times, adding effects on dialogues, lowering their volume, let alone completely deleting them can act as a powerful resource for creating emotional scenes, shifting the narration to another level or setting a feeling linked to a character. We will analyze two scenes that will help us understand this aspect.

The first one is again from the miniseries Chernobyl, in the second episode “Please Remain Calm”. The day after the explosion, the decision is made by the authorities to evacuate Prypiat, the nearest city from the Chernobyl nuclear power plant. As we are following a car with a megaphone on the roof diffusing an audio message warning people about the situation, the camera is filming different parts of the city where families are taking their luggage in direction of busses that will help them evacuate. The music of this scene starts at the same time as the diffusion of the message, the music being at a lower volume than the speaker. We can also hear some elements of sound design (people walking, children crying, discussions...). But as the scene moves forward, the sound design completely vanishes while the music gets louder. The voice from the speaker diminishes in volume, so that we mostly hear the music being played while we are watching an old man talking to an agent before going in the bus. This dialogue is utterly silent. The audio hierarchy is inverted here: we can't hear anything from this scene except the music and the speakers in the distance. The choice of not giving the spectators the information conveyed by this dialogue highlights how critic the situation is and gives a dramatic turn to the story. We are experiencing a catastrophe that will have a huge impact on human lives, and therefore the dialogue between the old man and the agent is simply irrelevant to support this idea. The music is taking care of adding this dimension to the story telling, more is said by removing the dialogues. Here, the absence of dialogues is more powerful than if they were in the scene. The information is carried by the music, which is taking care of giving the spectator the mood of the scene, to make us listen to something that words could never express.

The second one is from the movie Sleepers (1996) directed by Barry Levinson -who also directed Rain Man (1988) and Good Morning Vietnam (1987)- with an original score by John Williams. At the end of the movie, Lorenzo Carcaterra (interpreted by Jason Patric) is visiting his friend, Father Bobby

(Robert de Niro) to tell him the truth about his childhood at Wilkinson home for boys, and the physical and psychological abuses of which he was the victim with his friends. Lorenzo is asking Father Bobby to be a witness at his trial, because he killed a man who was a guard at Wilkinson's, and in order to do so he's telling him the truth about the guards at this place. In this scene, the music starts just before Lorenzo begins telling his story. After the first sentence "He was a guard at Wilkinson's, you know", the voice of Lorenzo is heard on another level, as a voice over. As the character in the scene starts his long enumeration, the voice over is summarizing what is being told in the background. The music and the voice over are becoming the loudest audio elements heard in this scene, and the dialog on screen is lowering down in volume while being filtered and reverberated in different ways. This choice of concentrating mostly on the music and the voice over is coherent with the directing of the scene: the camera is set on the face of Father Bobby and on his facial reactions. The focus isn't on the speaker, but on the listener; therefore, the music emphasizes the silent listening of Father Bobby.

Here, adding different units of dialogues is used in order to add another layer on the narration, the story we know as a spectator is bypassed by the voice over which is telling us more about the scene we are watching. We don't need to hear this story again, so we focus our attention on something else: the reactions of Father Bobby. Applying effects on the voice is allowing us to put the emphasis on something in particular, here the deep feelings of the priest who is shocked by the brutal truth that he's discovering, an interior feeling that takes over the dialogue. This is helped by the addition of filters and reverberation on Lorenzo's voice.

Regarding dialogues, we can argue that sometimes less is more. Our hear is used to movies and series where the voice is always used in the same way, in a naturalistic approach. Playing with the perception of the voice with the help

of effects such as reverberation, filters, echo... can add a lot to the narration and help the story to be perceived on another level by the spectators.

Conclusion

Throughout our researches, we found that interactions between music, sound design and dialogues are directly linked to the way we listen to movies and series. The three modes of listening of Michel Chion have helped us interpreting sounds by explaining that our listening isn't linear, it is an acoustical journey in our brain that lives and evolves during our experience of a movie or a series. The analysis of the soundtrack from Chernobyl helped us put the emphasis on the importance of the composition process in the making of a music score, as we can find interesting ways to interact with the audio of a series by following a guideline from the narration, in this particular case using sounds from a nuclear power plant to create the music. We have also underlined the interesting aspects of the "blurred zone" in which our listening is always shifting between what we categorize as sound design or music, both for spectators and for the characters on screen, thus adding depth to the storytelling. Then, we showed that in order to understand the interactions between music and language we needed to find out how the brain categorizes sounds with the help of neurosciences that allowed us to see that language and music are connected in our brain. These connections have been analyzed in the context of musicals where we noticed that the shift from language to music happens in units within the audio in movies and that this porosity between music and language becomes very clear as we saw dialogues appearing in released soundtracks from movies. Finally, we underlined that the absence of dialogues or the addition of audio treatments on them with the help of effects (reverberation, filters...) might have a strong impact on the storytelling as it redefines the way our listening is guided throughout the audiovisual experience.

It appears important to notice that this field of experimentation about audio interactions is a great way to explore new ways of telling stories, but it shouldn't by any means aim to become a standard, as it should stay a useful resource to help the narration in specific situations. But the curiosity that drives those researches is a great tool to keep as a creator, for audio has so much to bring to the creation of a story. As Michel Chion explains in *Audio Vision*: "Film and video makers, scholars and technicians can get to know their medium better as a result of this experience and gain mastery over it [...] So just as directors and cinematographers -even those who will never make abstract films- have everything to gain by refining their knowledge of visual material and textures, we can similarly benefit from disciplined attention to the inherent qualities of sound."¹⁹ Considering audio as an organic ensemble that interacts within itself is a great opportunity to explore new acoustic territories in order to feel new emotions and hear new stories. As neuroscientific researches will continue to study how the brain processes sounds, it will be interesting to see how music, sound design and dialogues are interacting together in other fields of art such as audio fictions, video games, virtual reality, or performances. Such mediums are the perfect place to experiment with the way we listen.

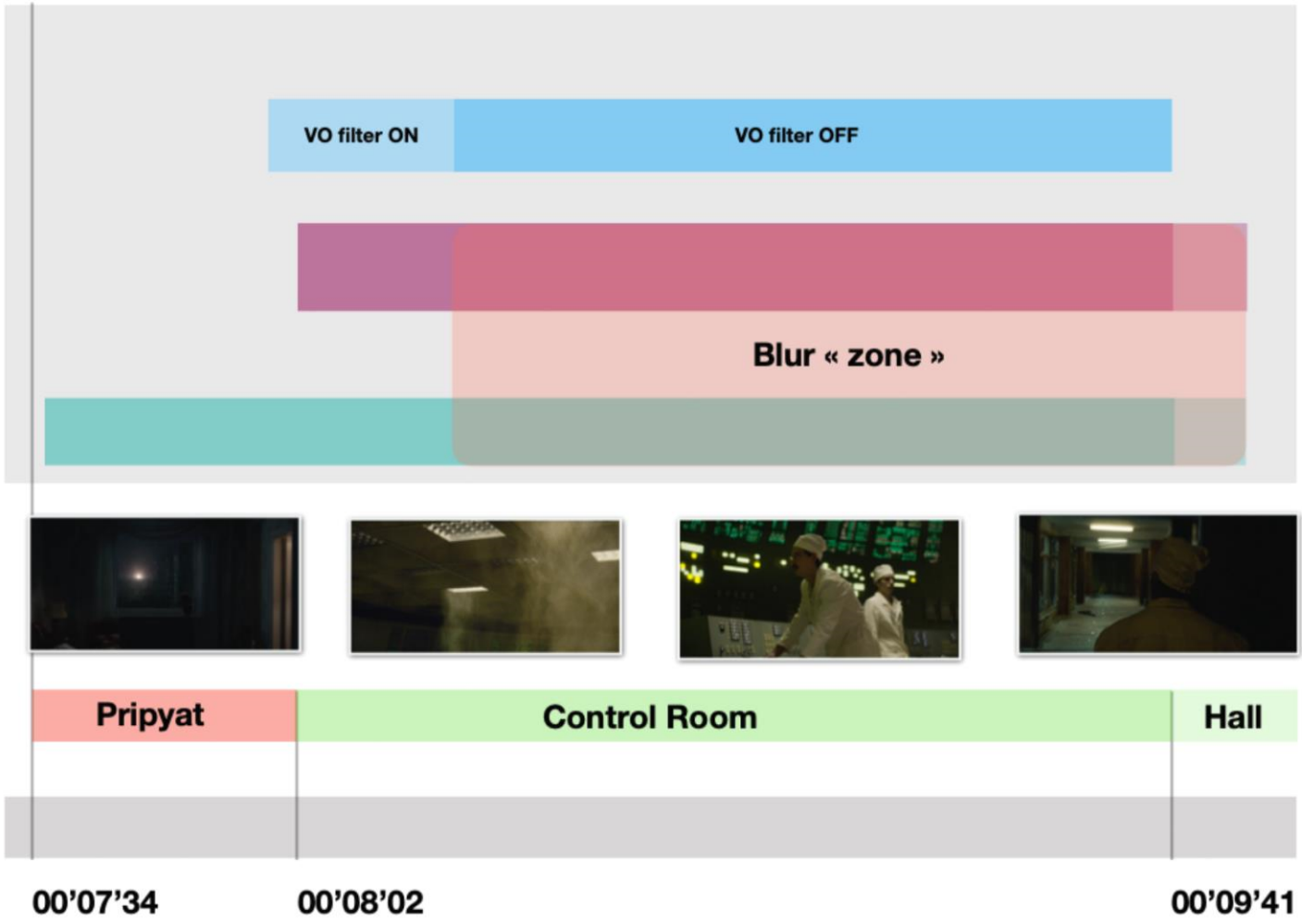
¹⁹ Michel Chion, "Audio-Vision: Sound on Screen", *The Three Listening Modes*, Columbia University Press, 1994, p. 31

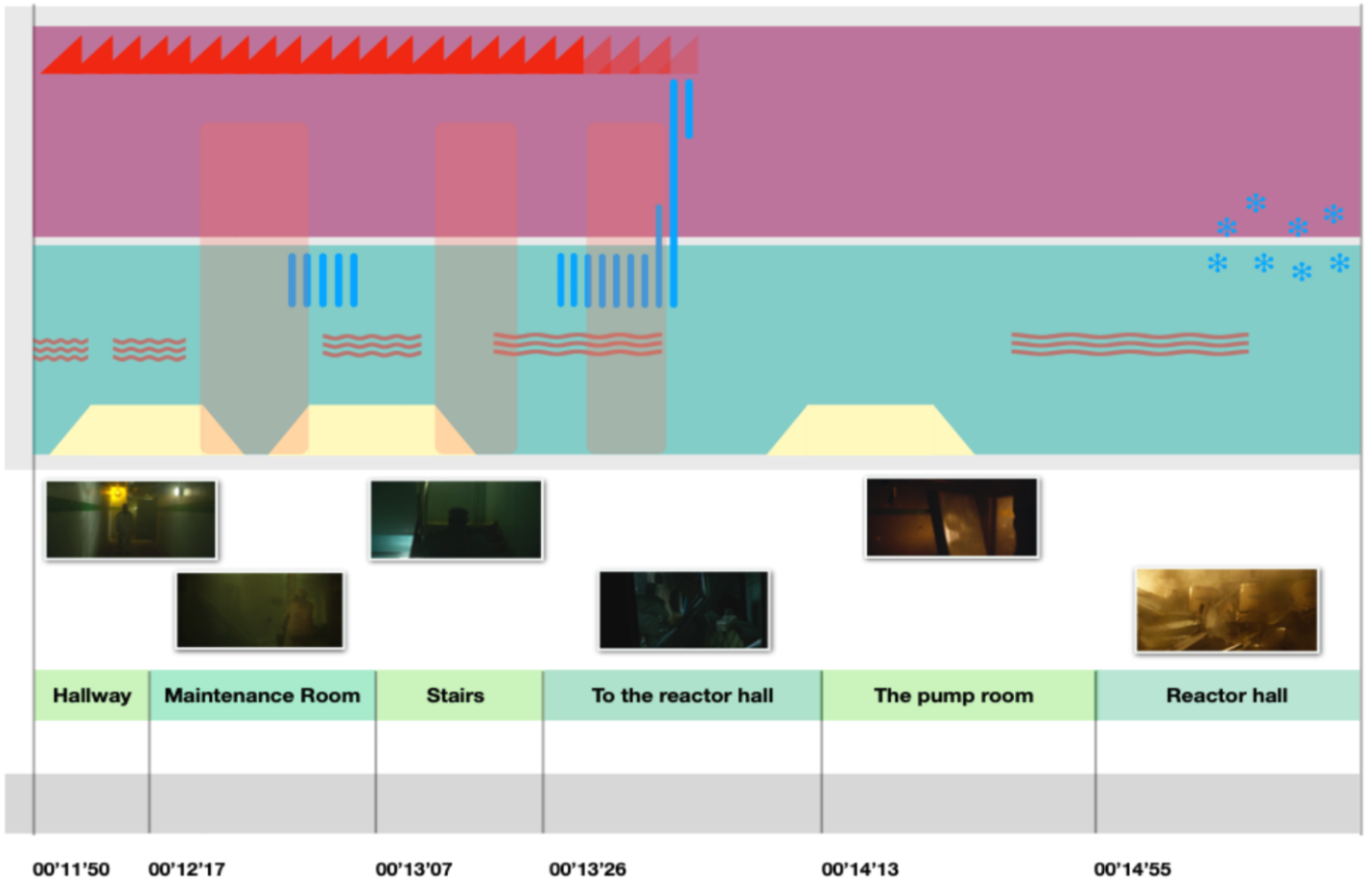
Bibliography

- Michel Chion, “Audio-Vision : Sound on Screen”, *The Three Listening Modes*, Columbia University Press, 1994
- Michel Chion, *Music in cinema*, Columbia University Press, 2021
- *Score : The Podcast*, “Chernobyl composer created entire haunting score from real power plant sounds”, 2019, 65 min
- Julia Berezutskaya, “Brain-optimized extraction of complex sound features that drive continuous auditory perception”, *PLOS Computational Biology*, 2020
- National Science Foundation, “How does the brain learn categorization for sound? The same way it does for images.”, 2018
- Valerie L. Trollinger, *The Brain in Singing and Language*, *General Music Today*, 2010
- Liz Green, *The Palgrave Handbook of sound design and music in screen media*, Palgrave Macmillan, 2016

Table of illustration

Figure 1: Analysis of the score structure from the first episode 1/2	15
Figure 2: Analysis of the score structure from the first episode 2/2	18





Thesis by: Jérôme YVEN

2nd cycle (Master), 2020-2022

Interactions between music, sound design and dialogues in movies and series

Abstract

This thesis focuses on the different interactions that can exist between music, sound design and dialogues. Through the analysis of soundtracks and neuroscientific researches about the way our brain reacts to sounds, the goal is to demonstrate that crossing the lines between those three audio elements can add new interesting layers for the narration of the stories that we see and hear in audiovisual contexts.