KILLING THE JOKE

SOUND DESIGN PRACTICES IN HORROR AND COMEDY FILM AS CONTEXT-INDEPENDENT TECHNIQUES



Master Thesis in Sound Design

Leonardo Guerra / +41 79 884 45 44

Mentorship – Manuel Gerber

Zurich University of the Arts – Zurich

September 2026

hdk

Zürcher Hochschule der Künste
Zurich University of the Arts

Abstract

Genre films such as horror and comedy have developed a rich vocabulary of sound design practices, defined by both their semantic significance and their psychoacoustic effects. This paper argues that by analyzing and deconstructing these practices to understand the relationship they have to both semantic and psychoacoustic significance, and how they influence each other, they can then be applied to realistic filmmaking to achieve the intended affective state, i.e., fear and laughter, by isolating the context-independent aspects from their context-dependent (semantic) counterparts.

Practical application of these practices to two short films, one horror and the other a comedy, reveals that horror is more suited for context-independent applications of sound, due to reflexive responses that bypass cognition. Comedy, on the other hand, is more deeply tied to these semantics and semantic subversion, as the psychological mechanism of laughter necessitates cognitive understanding of said semantics to spot incongruities.

What this means for sound design practices is that there are some possibilities of applying sound design practices as context-independent, but they remain, however, forever bound to the respective genre semantics, as it always necessitates a degree of mitigation to achieve the intended affective state.

Further research might enlighten the degree to which these context-independent practices could be applied to other genres, and to what degree comedy relies on the subversion of overall film semantics that are not genre exclusive.

Table of content

1. Introduction 4
2. Genre-specific Sound Design 5
2.1. Genre Similarities 6
2.2. Subjectivity and Priming....9
3. Horror 10
3.1. Role of sound design....12
3.2. Jump Scares 13
3.2.1. Setup 13
3.2.2. Payoff....16
3.3. Foleys....17
3.4. Ambiances....19
3.5. Quantitative Analysis 20
3.6. Conclusion 22
4. Comedy....22
4.1. Slapstick....24
4.2. Sound Gags 26
4.3. Silence....28
4.4. Dialogue and performance....29
4.5. Conclusion 29
5. Final Considerations: Horror vs Comedy....30
6. Realistic films....31
6.1. First issue 31
6.2. Perceptual Realism....32
7. Practical Applications....34
7.1. Introduction....34
7.2. Talion 34
7.2.1. Approach....35
7.2.2. Noé gets shot in the face and dies....35
7.2.3. Where's that hunter, and why is he so heavy? 36
7.2.4. Why so many crickets? 37

7.2.5. Frequency distribution 38
7.2.6. Outcome 38
7.3. Kloster 38
7.3.1. Approach 39
7.3.2. Extension, substitutions, exaggeration, and other absurdities 40
7.3.3. R.I.P. little bird (2026-2026) 41
7.3.4. Let the face do the talking 42
7.3.5. Loud fly on the wall 42
7.3.6. Outcome 42
8. Conclusion 43
Outlook 44
Sources 46
Filmography 47
Video Links 49
Figures 50

1. Introduction

Genre cinema has developed a rich and immediate vocabulary of sounds. These sounds are recognized and interpreted with absolute immediacy, and can carry a whole world on their shoulder. From the red-tailed hawk screech in spaghetti westerns like *The Good, the Bad and the Ugly* (Leone, 1966), to the stingers in horror films like *Halloween* (Carpenter, 1978), Ben Burtt's sonic palette of lightsabers, blasters, and technological beeping in *Star Wars* (Lucas, 1977), and the descending trombones from *The Gold Rush* (Chaplin, 1925).

They are inarguably deeply woven into the fabric of their genres and their identities. In the case of spaghetti western and science fiction, the sonic palette places the viewer in this constructed world. Specific periods, locations, characters, and ideas are aided by sound design. The sound is therefore drenched in semantics related to the genre in question, and it's dependent on the context.

Genres like comedy and horror are somewhat different because although certain sounds are tightly woven into the sound aesthetics of the genre, they also serve an emotional and dramaturgical function.

What this means is that sound design practices for these two genres can be seen as both context-dependent and context-independent.

Context-dependent sound design refers to the semantics of the genre. How audiences relate to, understand, and expect these semantic elements according to their cultural understanding. Context-independent sound design constitutes a psychoacoustic effect on the audience, independently of the meaning of the sound.

This research aims to understand to what degree sound design practices can be understood as context-dependent and context-independent. Unveil how semantic elements can be analyzed and understood in relation to the context-independent psychoacoustic effect of said practices, and how the need for producing specific affective states is related to the semantics of their respective genres.

This will be accomplished by analyzing excerpts of films pertinent to the subject, paired with scientific and cultural research.

The understanding of the relationship between semantics and psychoacoustics can aid in practice, unraveling to what extent these elements can be applied as independent modes, or if it's even possible.

A practical case will be presented, where I attempted to apply many of the analyzed practices as context-independent sound design to two short films: *Talion*, a horror-thriller directed by Micha Straub, and *Kloster*, a comedy directed by Salomé Käsemodel.

As these films strive to be realistic, the objective of the sound design process is to engage as little genre semantics as possible, since those semantics, as we will later see, tend to be disconnected from any notion of realism.

2. Genre-specific Sound Design

In this chapter, we will lay the groundwork to establish the commonly used sound design practices in horror and comedy films. Said practices can be described as recurring techniques intended to underline and reinforce the desired emotional impact of the film.
The reason behind the choice of analysis of these two genres in particular can be traced to Rick Altman's theory on genres. He postulates two modes of separation: semantic and syntactic genres. Predominantly semantic genres are defined by traits, attitudes, characters, locations, and settings. Such genres are tipically world-building-centered: e.g., science fiction, fantasy, western, samurai, and so on. Predominantly syntactic genres are, on the other hand, are structured around the relationship between the film and the audience. They are more focused on emotional communication and the intended mood of a given film (Altman, 1984).
These two categories are not mutually exclusive, but rather work in tandem, influencing each other as syntactic convention can become semantic and vice versa.
Horror and Comedy have strong syntactical backing for the way they operate, and have in time developed their own semantics. When applied to sound design, these semantics are identifiable as recurring practices, tropes or clichés.

The aim of this paper is not only to illustrate some common techniques and practices, but also to analyse these techniques, breaking them down into their essential components and functions, i.e., reaching a deeper understanding of their effect on the viewer's emotion, independently of the film's possible setting and "genre identity". The issue at stake here is to understand the degree to which components of sound design are context-dependent and context-independent, and how they can influence each other as causality.
Context dependency would imply that the effectiveness of sound design is entirely related to the semantic qualities of the film, therefore its meaning, emotional impact, and narrative function. This semantic (contextual) layer is what the audience understands and even anticipates as overarching genre conventions. This would, however, imply that if sound design were entirely context-dependent, the semantics of a given genre are exclusively context-defined.
The term 'semantics', as used in linguistics, essentially refers to the meaning of words, utterances, or discourse. When applied to sound design, it refers to what a sound *means* to a listener. The issue at hand is to establish how much of this understanding is based on context, convention, and cultural learning (Barbosa & Dizon, 2020), i.e., communal common ground. The notion of 'common ground' in conversation was first theorised by Robert Stalnaker (1978) and further specified in Clark (1996), who distinguishes between *personal* and *communal common ground*. The first is based on mutual personal acquaintance between the interlocutors, whereas the second (relevant for our purpose) refers to a shared cultural knowledge (Clark 1996: 100-112). This shared cultural knowledge is mutually assumed by the interlocutors and implicitly taken for granted. What this means for sound design is that there's an assumed implicit understanding between filmmaker and audience as to the semantics of sound. This would allow for seamless emotional communication between filmmakers and audiences.
Context-independency, on the other hand, would imply the exact opposite: that sound design is unbound from its meaning and can be understood by its psychoacoustic effects. This would imply that sounds do not

convey their meaning through their associated semantics, but rather through their intrinsic acoustic properties. What this entails is that sound design is not dictated by its cultural context and the theory of common ground, but rather by an innate way of reacting to and interpreting the sounds' distinctive qualities. The listener's understanding of sound is thus innate rather than cultural. This idea is supported by crossmodal correspondences, an empirical phenomenon by which it is possible to map specific features, or dimensions, of experience across sensory modalities. This phenomenon points to a universal understanding of certain sound qualities in relationship to a corresponding congruent or incongruous visual stimulus (Spence & Parise, 2012). What this suggests is that peculiarity in sounds such as pitch and loudness can have an intrinsic meaning regardless of semantics.

Another clue pointing to the context-independent aspects of sound design is involuntary physical responses to acoustic stimuli. An example of this that will be discussed in the following chapters is the so-called *startle response.* A loud and sudden noise can engage primordial fight-or-flight responses, solely based on the acoustic qualities of the startle.

The opposition between context dependency and independence is complementary: the audience's understanding of sound can be in part context-dependent and in part context-independent. The aim of this research is not to argue in favour of one or the other of these paradigms. By pinpointing and deconstructing the psychoacoustic as well as the deeply cultural and semantic qualities and effects of sound design practice, the aim is to gain a deeper practical understanding of how context-dependent and -independent sound design interact and complement each other in the comedy and horror genres.

2.1. Genre Similarities

Horror and comedy, at first glance, might feel like complete opposites that could not be more different from one another. This might be due to the perceived difference in physical and emotional reaction they elicit. However, the two genres have much more in common than what appears at first sight. As Robert Bloch, author of *Psycho,* affirms, "comedy and horror are opposite sides of the same coin" (A Quote by Robert Bloch, 2026); he implies that beneath the surface, the same definitions and mechanisms are mutually applicable. This is possible by virtue of both genres' handling of the unexpected and of grotesque situations.

Manuel (2013) expands on this further, supporting the idea that even the physical reactions are not always different. According to her, since audience reactions are not mutually exclusive, defining the genre solely by the effect of its content on the audience is not a valid criterion for categorisation. Therefore, due to the varied reactions that might arise, it is sub-optimal to use the intended physical response as the benchmark for qualifying the film as a failure or success. A horror film might be amusing, while a comedy could be stressful to certain audiences. Ultimately, Manuel's research points to semantic differences between the genres, most importantly in how the characters react to chaotic situations. Comedy engages in denial of the chaos or fights it off through a positive reaction, whereas horror occurs when the characters give in to chaos or are consumed by it.

What is meant by chaos relates to a core principle of dramaturgy for both genres, that is, the handling and subversion of expectations. Chaos originates in incongruity, where someone or something is out of place or

in an unexpected context. Incongruity in itself does not generate the chaos, but the possibility of unexpected events arising from it. Incongruity can take many different forms and can be portrayed as threatening, relatable, or realistic. Despite the different effects that might arise from incongruities, they are a good starting point for understanding sound design for both genres.

Horror and comedy are film genres that thrive best on a first viewing. Action, adventure, romance, and so on can at times be predictable in the story beats: good triumphing over evil, the hero gets the girl, the adventurers realize that the real McGuffin was the friends they made along the way.

Syntactic genres, on the other hand, are more often than not built around expectation subversion. The incongruences give the viewer the hint about which element of the story or sequence to pay attention to, and then subvert those expectations to generate a humor/fright reaction.

Such subversions can happen in many ways and can happen on many layers. The aforementioned incongruities and chaos factor are intratextual, therefore dictated by the narrative of a film. Subversion can also happen at an intertextual level. This can be the example of Marion Crane's death in *Psycho* (Hitchcock, 1960). The film establishes her as the protagonist of the film, but then Hitchcock flips everything over its head by killing off the protagonist well before the halfway point of the film. Such a subversion is intertextual, as for the audiences at that time, it was inconceivable that the established protagonist would die in the first half of the film.

Subversion is always presented in a two-step format: setup and delivery, or payoff. This is observable in both genres as the setup opens the stage for the subversion to occur. Although the structure for laughs and scares operates under the same principle, they tend to have their idiosyncrasies. In the case of horror, the setup is a much longer sequence. Once the payoff arrives, in the form of a massively loud startle and a grotesque monster appearing out of nowhere, the audience is relieved for a brief moment. This is why it's always in a horror film's best interest to drag out the setup as long as possible. In this moment, the audience is anxious due to the unknown nature of the upcoming subversion, which in turn makes the startle much more effective once it hits (if it even hits).

Comedy follows the opposite structure the setup is expedited as much as possible to allow for more deliveries, or so-called punchlines. This is the reasoning behind historical forms of theatrical comedy. If the characters are defined and the situation is established, then the comedy can begin with a series of punchlines. This is also why sitcoms are so prominent in on-screen comedy. This is also possible in cinema, with the example of parody movies being very prominent in the 20th century. One must only think of Mel Brooks' filmography. Films like *Blazing Saddles* (Brooks, 1974a), *Spaceballs* (Brooks, 1987), *Robin Hood: Men in Tights* (Brooks, 1993), or even *Young Frankenstein* (Brooks, 1974b). These parodies make use of established genre tropes, or intellectual properties, as a general setup for comedy. They presuppose an audience's cultural and social understanding the respective references for the subversion to be understood. The pre-established contexts of genre tropes, characters, and sometimes even storyline (like Robin Hood or Frankenstein) remove the need for setup, thus allowing for immediate deliveries and subversions (Hutcheon, 1991).

It's important to note that these films were made at least one decade after the golden age of their respective parodied genres. After they've been embedded in the cultural fabric, ready for deconstruction.

Sound design allows for a wide variety of reinforcements and subversions. It's used to great effect to strengthen payoffs, like in the case of jump scares in horror movies, or in slapstick moments in comedies, where the impact of the frightening or funny action is strengthened sonically to achieve a greater effect.

The soundtrack can be used to craft a specific mood: scary, serious, ominous, peaceful, and so on. The subversion can then be achieved by way of a sudden change to these moods.

This is widely prominent in comedies, where it will borrow semantic and syntactic elements from other genres to either subvert expectations or satirize the borrowed genre directly.

Sound design can also be used in both genres as conveyors of off-screen information. By giving audio cues to the viewers, their imagination is engaged, extending the action beyond the space of the scene. The imagination of the viewer is used in different ways to achieve different effects. In the case of horror, the information given is left vague on purpose; e.g. alerting of an unknown presence behind the wall raises the stakes. An example would be the final scene from Hitchcock's *Rear Window* (1954), where the protagonist, alone in his apartment, hears ominous footsteps approaching his apartment door. The camera does, however, never leave the room, leaving the identity of the approaching figure to our imagination.

In the case of comedy, sound can be used to strengthen and reiterate a pre-established gag, for example the "Frau Blücher" joke from *Young Frankenstein*, where the utterance of the character Frau Blücher causes thunder to roar and horses to startle neighing. This is established visually the first time as she introduces herself upon Viktor Frankenstein's arrival in a carriage. But once inside, the joke is continued as the horses neigh again at the utterance of the character's name, and can be heard from inside the castle.

The core difference in the two approaches is in the communication of information. On one side, the sounds create mystery, intrigue, and suspense, pointing to an unknown presence; an identity purposefully kept away from the viewers. And on the other side, we have the strengthening of an absurd idea; it is imperative that the audience makes the connection that the woman's name causes animals to be scared, and the sonic repetition of this joke only enhances the absurdity of the idea.

In horror and comedy respectively, which, as prevalently syntactic genres, are more concerned with tone, mood, and emotional impact, the need for realism is all but rejected. This is not to say that logic and believability are completely sidelined, but to achieve the catalysing incongruity, some detours from reality are all to be expected. Sound effects and foleys are great tools to achieve incongruity, e.g. through the lack of fidelity in the sounds. Fidelity in sound for films refers to how faithful the sound we hear is to the source of the sound. This doesn't necessarily mean that the sound must come directly from the source, but that it has to match the source in a congruent way based on the viewer's experience ("What Is Fidelity," 2021). As mentioned before, these two pre-eminently syntactic genres are built on the premise of incongruity. Therefore, it is very common for non-faithful sound to achieve the desired effect, e.g., by means of various forms of exaggeration of intensity, pitch, length, or perhaps a completely wrong sound altogether.

To extend on this idea, it is also possible to layer already incongruous sound with non-diegetic sounds to further exaggerate a given action or object. A non-diegetic sound is any form of sound intended only for the audience to hear, and is not heard by the characters in the story. This means that the usage of these sounds is intended purely for enhancing the intended affective state of the audience and yet has no influence on the characters and story. These sounds can widely range from eerie drones in horror films to over-the-top

sweeteners in comedies. This points to a very utilitarian usage of sound design, rather than logical or realistic.

2.2. Subjectivity and Priming

Scary and humorous responses in audiences are unpredictable, to say the least. While there are some neurological explanations for such responses, especially relating to fear, a given stimulus does not necessarily constitute the intended response.

Fear's primary purpose is to prepare the body to react to impending physical danger; it manifests itself with a wide array of physical symptoms, such as higher blood pressure, dry mouth, butterflies in the stomach, goosebumps, dilated pupils, and so on. Some stimuli, most notably sudden loud noises and sudden movements, produce such responses quite universally. This is the so-called startle response, an innate physiological response to startling stimuli, which is further amplified if the subject is already in a fearful state.

Another factor for possible scares is prepared fear (reference). Prepared fear is a theory that postulates that certain environmental fear factors are genetically transmitted, such as snakes, spiders, blood, and the dark. Fear is a survival mechanism that heightens our awareness and prepares the body for a fight-or-flight response. Therefore, it stands to reason that it is triggered by environmental factors that pose a potential threat to our survival.

Prepared fears are observable in the very fabric of horror films; most on-screen monsters present features reminiscent of threatening animals. Blood and darkness are also used to alert our survival instinct and instil fear (Hye-Knudsen et al., 2024).

Nevertheless, responses to the viewing of a horror film can vary widely. Some people might cover their eyes in terror for the whole run-time of the film, while others might just laugh all the way through. Research on this phenomenon, although it proposes some compelling theories, does not provide a definitive answer.

One possible explanation for laughter is disengaging submission, i.e., laughing at a scary situation as an evolutionary means of communicating a desire to avoid conflict. Another explanation points in the opposite direction: rather than mitigating a situation, laughter serves as a complete denial of it. We're signalling to ourselves, and to others, that what is being observed does not pose a threat.

As horror and comedy share a common root in incongruity, a seemingly "wrong" reaction subverts expectations and social norms. The incongruity can also appear on the meta layer. E.g., in low-quality horror films, most members of the audience tend to laugh exactly because of the incongruity between the desired outcome and its dreadful execution.

Most theories argue against mutual exclusivity, since many situations portrayed in films could be both funny and scary at the same time (Green, 2015).

All in all, the takeaway is that reactions of fear, laughter, or neither of the two are mostly dependent on the viewer's individual predispositions. Ultimately, fear, very much like laughter, cannot be triggered at command, and to a large extent it is the individual who determines whether a stimulus is perceived as scary or funny.

Comedy is widely regarded as extremely subjective. Although there is a social element to both laughter and scares, where a social viewing can enhance the experience, a joke in a vacuum experienced by an individual alone will elicit a laugh based on what the person deems funny.
Unlike fear, which is rooted in our deeper survival instincts, comedy appeals to much higher brain functions. The leading theory around what makes something funny leads back to incongruity, where an incongruous situation violates our expectations based on our normal mental patterns. Such violations are benign in nature, therefore non-threatening and harmless. This could also be a subsequent reason why something scary can also be considered funny, by providing a benign violation to the viewer that, for some, in the context of fictionalized horror, can intrinsically be interpreted as benign, i.e., the realization that it is a farce.
If the benign violation actually elicits laughter, that is another matter. It is preposterous to assume that something is objectively funny simply because it follows the principle of benign violation.

Although in prevalently syntactic films certain reactions are impossible to guarantee, there are important phenomena that aid in the accomplishment of such goals. As we saw with the startling response, where the already altered state of the viewer can enhance the response itself, such an effect can be accomplished before the viewing experience even begins. This is the idea of priming, where the audience's affective state is altered psychologically due to the information about which kind of film it is about to watch. In the case of horror films, for example, it has been proven that the rating of a film and content warnings had a direct impact on the fear and anxiety perceived by the audience.
This can be applied to any form of audiovisual media, in which prior knowledge or experience will influence the viewer to engage in a critical consumer process. The contents of the piece of media are more closely evaluated and scrutinized. And in the case of prevalently syntactic films, the supposed affective elements are more closely analysed in relation to the premise of the viewing. The viewer engages more deeply with these elements. If a film is supposedly terrifying, then the viewer might be keener to pay more attention to figure out how terrifying it really is, and in turn raise fear and anxiety in the anticipation of the pre-established terror (Huston et al., 1997).

3. Horror

The history of horror movies is varied and suggests significant differences across semantic contexts. The pre-history of the genre is widely attributed to German expressionism (late 1910s - early 1930s), with staple films such as *The Cabinet of Dr Caligari* (Wiene, 1920) and *Nosferatu* (Murnau, 1922). Although this movement was still in the silent era of film history, it influenced the genre with its psychological dread, distorted sets, and stark shadows. The usage of negative space is central to most of the later iterations, as the unseen makes use of the viewer's imagination to hide something horrific. This effect was enhanced once sound was introduced, and sound designers were able to use the mismatch between what is heard and what is seen to instill fear.

In the 1930s and 1940s, the most notable output of horror films moved to Hollywood, with the Universal Monster Era. This period established the genre's iconography by carrying literary monsters onto the silver screen. This includes genre classics like *Frankenstein's monster*, *Dracula, The Mummy*, and *Werewolves*. In the later years of the period, filmmakers started favouring the suggestion over the visible portrayal of the monsters.

After the Second World War, horror films increasingly related to commentary on the contemporary political zeitgeist, being reflective of the fears of the time. The communist paranoia of the Cold War in the United States fuelled a wave of alien-invasion films, merging horror with science fiction. Beings from other worlds invading Earth, threatening humanity (Americans) with indifferent de-individualisation, was a central theme in 1950s horror. This includes films like *Them!* (Douglas, 1954), *The Blob* (Yeaworth, 1958), and *Invasion of the Body Snatchers* (Siegel, 1956).

Later in the 60's, horror turned inwards. The threats were no longer abstract alien forces, but fellow citizens. This is the example of *Psycho* (Hitchcock, 1960), arguably the most important horror film of the decade, where the monster was no longer a supernatural being, but a human being, set in a commonplace American location. *Night of the Living Dead* (Romero, 1968) launched the modern zombie genre as a commentary on herd mentality, opposed to constructive collaboration. It is not a coincidence that this film was released at the height of the Vietnam War, where conservative mentality became incongruent with the first televised war.

This brings us into the 70's with heavy hitters like *The Exorcist* (Friedkin, 1973), *The Texas Chain Saw Massacre* (Hooper, 1974), and *Carrie* (De Palma, 1976). This era was built on the themes that emerged from the cultural revolution of 1968. The cultural contrast between "secure" and "paranoid" worldviews and the perceived loss of rationale of the decade allows the *paranoid* to exploit humans' inability to intervene against external forces.

The end of the 70's gave birth to the slasher genre with *Halloween* (Carpenter, 1978). The genre is characterized by recognizable masked killers performing in the framework of multi-sequel franchises. The 80's were also a time of massive innovation in practical visual effects, turning horror towards bodily integrity, perhaps to counterpoint the ease of life of 1980's consumerist America.

At the end of the century, the genre turned self-referential, appealing to the audience's understanding of the medium in films like *Scream* (Craven, 1996).

In the new millennium, the genre explored found footage with *The Blair Witch Project* (Myrick & Sánchez, 1999), which influenced countless films in the style. *Saw* (Wan, 2004) and *Hostel* (Roth, 2005) ushered in a grittier, more visceral subgenre dubbed "torture porn."

The current era of horror can be considered a new golden age, where arthouse craftsmanship deals with current societal topics like grief, trauma, gender, and race. This is thanks to studios like A24 and Blumhouse that brought back the genre into critical respectability (Leeder, 2018).

Horror is built on the premise of negative human emotions, primarily fear. Its history presents a wide array of sources of fear, often appealing to the contemporary political climate and the fears related to the concurrent era. Fears like loss of religion, fascism, atomic annihilation, de-individualisation, group-thinking, and foreign invasions are all linked to the historic specificity of the era in which the films are produced. Some fears are, however, universal. These might include loss of bodily autonomy, like in *The Fly* (Cronenberg, 1986), social

rejection, such as in *Carrie* (De Palma, 1976), isolation and abandonment, like in *The Thing* (Carpenter, 1982), and many more. Themes, of course, can mix and match between concurrent topics and universal fear.

Fear, as a psychological mechanism, functions as a protection from threats, warning the body of any imminent danger. This suggests that all the different types of fears explored in horror films all relate to an ultimate universal human fear: the fear of death.

Fear of death, and by extension fear of non-existence, fear of pain and suffering, fear of the unknown, can be portrayed and understood in a variety of ways. It can extend beyond the simple physical consequences of death to threats to our mind and spirit. Insanity, or eternal damnation, can be understood as metaphorical deaths, whereby our perceived integrity is somehow compromised.

The core mechanism of horror is to use our innate fear of death, often contextualised through the use of fiction, to trigger a fight-or-flight response in the viewer's body. Once this state is resolved, the brain releases dopamine, a neurotransmitter and hormone that provides a surge of pleasure and reward. This happens even though it is a staged fright, which poses no actual threat to our integrity.

3.1. Role of sound design

The fright response is created by using particular visual and auditory stimuli. Sound design and musical score can alert our brain by tapping into our innate survival instinct and recognition of danger (Neurominded, 2023).

With this in mind, we can describe how horror films operate. By abiding by a structure of tension and release, they are capable of delivering the pleasure of "surviving a scary situation" multiple times within the runtime of a given film.

In this mechanism, sound design is one of the strongest driving forces, being able to engage and then disengage the alert state of the viewer.

Although they fulfil a functional role in the dramaturgy of horror, sound design and music have very distinct stylistic features, which are deeply meshed with their functionality.

In the words of William Whittington (2007, p. 131):

> Music and sound effects, particularly ambient effects, in horror films
> present a dark lyricism shaped by the need for a spectacle of intensity or excess, dark emotions, and atmosphere. Orchestrations in minor keys, thunder rumbles and guttural growls can foreshadow a descent into unfamiliar territory or accentuate psychological fracture or weakness within characters. The expressionist results are chilling to filmgoers as the music and sounds of the night rake up and down their spines. The genre contract regarding horror films often stresses visceral impact over narrative causality and unity.

This passage demonstrates that horror cinema deliberately weaponizes sound as an emotional and psychological tool, prioritizing the audience's bodily and emotional response over storytelling logic. The

need for a “spectacle of intensity or excess” informs the soundtrack, emphasizing the importance of functionality relative to aesthetics.

Tension is centre stage here. With rising tension, the viewer becomes aware that something might happen at some point; it is often not clear what and when. This uncertainty is what fuels the tension, and it is also the sound designer’s duty to uphold this. One, if not the most common, scare tactic in horror films is the jump scare.
In the following, we propose a description of some of the most important fear-managing techniques in horror films and the role sound design plays within them.

3.2. Jump Scares

The jump scare is a technique that involves a sudden appearance of a figure, image, or action, often as a payoff to a suspense-filled setup sequence.
The setup establishes the eventual payoff. The viewer is actively aware that something might happen soon; they just don’t know the details yet. They experience the anticipation stage, which subsequently enhances the payoff once it arrives. They are already primed for the jump scare itself. The payoff is where the actual jump scare occurs. With the aid of loud, visceral sound and sudden imagery, the viewer experiences a physical scare as a response to the scene. Jerking, flinching, screaming, a raised heartbeat, and so on (Acosta & Sears, 2026).

Sound design (as one might imagine) plays a massive role in the success of this scare tactic. As a general principle, the sound design requires a pre-established frame in the imagery itself. If the camera work, editing, and acting do not coalesce in agreement over the purpose and direction of a scene, there’s only so much that a sound designer can do to make a jump scare work.

3.2.1. Setup

During the setup, the sound design employs a varied array of soundscapes to create suspense.
Although the space inhabited by the characters always varies, a general practice in sound design can be observed. As a means to raise anticipation, the sound designer may manipulate the soundscape, a classic example being the drone. A low static sound, created by slowing down and/or pitching down a textured sample, or with a synthesizer. A prominent example of this is the “lawnmower” scene from *Sinister* (Derrickson, 2012), characterised by a palpable incongruence between image and sound. The scene establishes that we are watching a soundless video recording. The sound design, however, avoids total silence by offering a very quiet non-diegetic drone, in direct contrast to the frantic and illegible picture on the screen.

Visual cues that obscure the surroundings of the subject are very often used in jump scares. An abundance of shadows and one small, brightly lit subject, which the viewer is forced to focus on, primes them for the payoff.

This idea of grabbing the viewer's attention can also be a common practice for jump scares. This usually happens right before the payoff. The scene builds tension throughout through drones and dissonant music, to then lets go just before the scare. The soundscape is stripped naked, leaving an unnerving silence which is undercut by a quiet noise. Examples of this are "The Red Demon" scene from *Insidious* (Wan, 2010), (Figure 1), or the opening scene of *Nosferat*u (Eggers, 2024) (Figure 2).

In *Insidious,* scene one, the character tells her son about a dream she had in which a "visitor" came after her grandchild. In the dream, we see a demon's hand extending towards the child in his bed. The sequence is filled with horror staples: drones, dissonant music, and creaky doors with a healthy amount of reverb. The demon's hand makes a very distinct crackling sound (somewhat similar to wood creaking).

When the cut returns to the conversation, the drones and the music are gone, and, right before the payoff, we hear the demon's hand's sound once again, somewhat subtly, to then cut to the son with the demon standing right behind him, accompanied by a loud stinger and growl. The clearing of the auditory space leaves much more space for the sound of the demon's hand to live front and center. Although the film establishes the connection to between the object and the sound, it still sounds mysterious. We have to extend our ears; the film invites us to listen carefully as this little sound might be important, but with dire consequences.



Figure 1 "The red Demon" scene from Insidious (Wan, 2010)

Nosferatu's scene is built in a similar way, where after a sonically charged sequence, we find ourselves on a close-up of a woman. Some soft cricket sounds can be heard as she lets out some very subdued moans. Out of nowhere, Count Orlok's hand loudly reaches her neck as the score swells frantically; he growls, and she screams.



Figure 2 Opening scene Nosferatu *(Eggers, 2024)*

The main purpose of this technique is to have the viewers listen closely to what's going on. The unobtrusiveness of these premonitory sounds is deliberate. By shifting the audience's focus to a quieter sound, you prime them for the payoff, therefore catching them off guard.
Another reason for the quietness (or in some cases absolute silence) is to attain a much higher dynamic range between setup and payoff.

Another technique that can be used for setups is the misdirect. Unlike the previous examples where the viewer is drawn in, the misdirect aims at lowering the viewer's guard. An example of this is the "Tall man" scene from *It Follows* (Mitchell, 2014). Here, the sound design plays deliberately with the viewer's expectations. In this scene, much attention is focused on the importance of not opening the door. The sound of the door hinge trying to pry open is very focused, as a drone ramps up in tension. Once the door is open, the drone vanishes, releasing the suspense. Only after that does the tall man walk into frame (Figure 3),



Figure 3 "Tall Man" scene from It Follows (Mitchell, 2014)

accompanied by the swelling synth score.
The misdirect lies in the calmness and false release provided by the soundscape once the door is open. For a brief moment, this makes the audience drop their guard, but immediately after, this relief is violated.

In the previously observed examples, on the other hand, although the soundtrack quietens down for the payoff, the tension is still kept high. This is because the narrative does not signal in any way that the sequence is over. In the "Tall man" scene, however, the conflict that introduces the tension that the soundtrack latches onto is resolved as a misunderstanding and overreaction, only to then subvert this brief release, making the viewers aware it was a misdirect.

In all these analyzed scenes, there's strong usage of drones or music to ramp up tension; while these sounds have a strong semantic connection to the genre, they have the characteristic of being very unstable sounds that don't grant any sort of resolution thus keeping the viewer on edge.

3.2.2. Payoff

As mentioned before, the payoff oftentimes utilizes a sudden image to shock the viewer. It can be a variety of representations; it is usually something that is a direct threat to one of the characters (respectively, the viewer). It could be the appearance of a paranormal entity, such as in the case of *Insidious.* Or it could be an attack on the character, such as in *Nosferatu'*s opening scene. Sometimes it could be something completely mundane or nothing at all, where the threat of a jump scare is enough to scare the audience.

The heavy lifting in the payoff is performed by the sound design with a sudden, very loud layer of sounds that are in direct dynamic contrast to the quieter setup stage. These layers can be a mixture of non-diegetic and diegetic sounds.

An element of these layers is the **impact**, which usually comes off as somewhat non-diegetic or at least exaggerated to some degree. The impact is characterized by a very loud transient that can be achieved in a variety of ways. Often it is left to the orchestration of the score; e.g., with brass or percussion providing a striking and loud impact for the scare. The main objective is to have a full-range instantaneous burst of inharmonic noise with a near-vertical transient. Of course, iterations may vary depending on the style of the movie and composition. An added element that enriches the impact is a low-end transient, which can be either a synthesized or a processed sound. It works similarly to a kick drum, providing the physical reaction to the sudden burst of sound, which is unexpected and therefore startling.

Another common feature in most jump scares is the infamous **stinger**. In contrast to the impact, which is more rhythmical, the stinger is more harmonic in essence. It is the quintessential horror sound, more often associated with the jump scare.

The stinger is a non-diegetic cue that is often dissonant or inharmonic. It can be created by a variety of sources, such as strings or bowed cymbals. It is s more continuous in

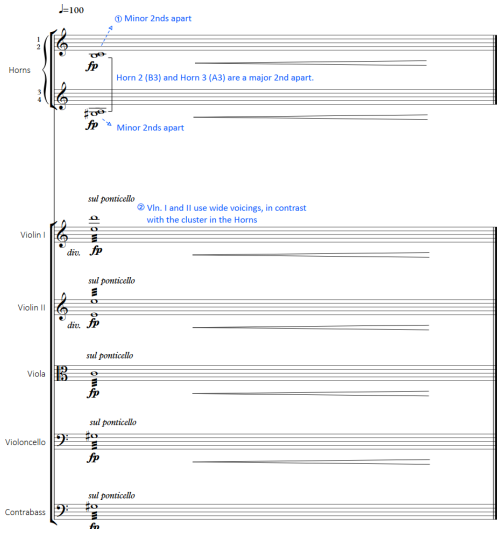


Figure 4 Cluster voicing (FMT, 2023)

length, usually extending over the length of the payoff sequence, although it could have a strong attack, overlapping with, or serving as, the impact sound. Cluster voicings are the foundation of these sounds, using these intervals to create a dissonant texture that is unsettling in accordance with our Western musical sensibilities. Clusters are created by arranging the tones of a composition in very close intervals from each other. A close voicing cluster is exactly that; the tones are played one semitone or whole tone apart. An open voicing covers a wider range by arranging the chord tones in a range wider than an octave (e.g., minor 9th or major 9th intervals) (FMT, 2023).

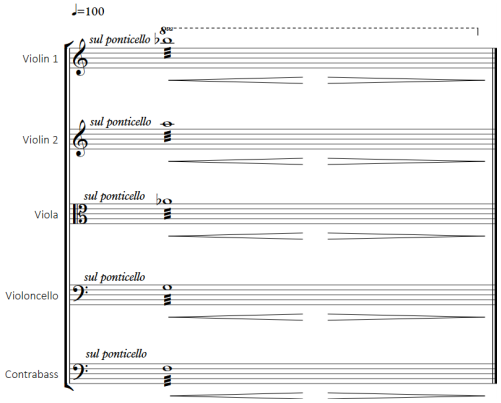


Figure 5 Open cluster voicing (FMT, 2023)

Stingers do not generally resolve the harmonic tension that they create. They instead enforce the tension and harmonic instability until the end, not allowing any form of release for the viewer, thus perpetuating the tension.

Diegetic sounds are also to be expected in a payoff. In the previously mentioned examples, such as *Insidious, Sinister,* and *Nosferatu,* a scream or growl is present in the payoff. Screams can be human or monstrous.

In the example of *Sinister* and *Nosferatu,* the growl helps us contextualise the imminent threat. If the impact and the stinger have an emotional function: scary, unsettling, unexpected, etc., the growls also have a contextual function that enhances the presence of the threatening creature in question, therefore making it even more frightening. Screams and growls have the added perk of being very strong psychological markers of danger. Screams cluster within a restricted portion of the acoustic space (around 30–150 Hz modulation rates) that corresponds to well-known perceptual space, roughness. This roughness activates the amygdala more strongly than other sounds (Arnal et al., 2015).

As mentioned, dynamics play a fundamental role in the successful execution of a jump scare. The startling of the viewer (i.e., the purely physiological reaction of fright) is dictated by the sudden change in dynamics in the scene. Impacts and stingers operate through their very loud onsets, which are in direct contrast to the setup phase. The startle effect occurs because, on a perceptual level, the differentiation between reality and fiction is cancelled. This means that, even though the viewer is aware of the film's fictional nature, the sound design provides concrete acoustic stimuli that can override the fictional nature of the perceived danger, thus shifting the fear into reality. The effectiveness of the jump scare lies in shifting the viewer's worry and empathy for the characters onto themselves (Sbravatti, 2019).

3.3. Foleys

Horror generates terror in us by tapping into our primordial fears. It exploits our innate survival instinct and striving to of stay out of harm's way. Fear, as an emotional response, signals the possibility of physical injury, discomfort, or even death. It would therefore stand to reason that horror filmography taps the nerve of our fear of death, death as the ultimate fear (Nummenmaa, 2021).

Exploring the source of our fear means delving into societal, political, religious, philosophical, and psychological origins.

With this in mind, the way in which these fears are presented and contextualised can be specified for different subgenres. In the case of body horror and slasher, the terror is very evocative of our primordial drive for survival and bodily preservation. Supernatural horror, with its wide array of otherworldly creatures such as ghosts, demons, and the outright indescribable, explores our fear of the unknown. The horror is provoked by entities of unknown origin which threaten our existence in ways we cannot comprehend, and therefore we are betrayed by our imagination.

Other sources of fear for our well-being are social exclusion, loss of reason or identity. These can be a threat to our survival by destabilizing our reality.

While soundscapes and ambience play a contextual role in the sound design, underlining the physical surroundings and psychological world, foleys and sound effects, on the other hand, focus on the strictly bodily dimension.

The foleys can appeal to a more direct physical experience within the viewer. In body horror, the terror lies in the physical threat posed to the characters, and therefore also to the viewer. The viewer is thus engaged in a sort of involuntary mimicry of the emotions and sensations experienced by the characters in the movies (Whittington, 2007).

To achieve such a visceral effect foley artists resort to a wide array of sound sources. Bodily sounds are recreated by handling and manipulating organic sources such as fruits and vegetables, which are represented as experiencing the same level of abuse as the human bodies on the screen.

A typical example is that of celery stalks being snapped in half to recreate the sound of bones breaking. To create a stabbing sound, one might just stab a cabbage with a knife to achieve the crunching sound of a blade entering a human body.

It is interesting to note (thankfully) most people's lack of sonic experience with such bodily excesses. It is far more likely to witness violent acts such as stabbings, gunshots, severed limbs, cracking skulls, squashing, decapitation, etc., as they are represented in works of fiction than to experience them in real life. This means that the foley work does not have to relate to real life, but it needs to clarify and enhance the act represented in the image in order to fulfil its dramaturgical and aesthetic roles. The repeated representation of these seldom-experienced sounds enforces a fictional realism in the collective unconscious. The arguably non-realistic sound becomes a surrogate for the real, a sonic convention that is widely accepted and undisputed (Langkjær, 2010).

Foleys are by nature synchronous with the actions depicted in the film. They are there to support the played action at an acoustic level. It is expected by the viewer that certain actions have their own specific sound to accompany them. This is, of course, because of humans' daily experiences of interacting with their surroundings. Our experience dictates that an action and its consequential sound are the same thing, one being inconceivable without the other.

Stemming from this expectation as a deeply ingrained inevitability, in the case of horror film sound design, a large amount of transgression is commonly used to evoke a feeling of the uncanny. By dislocating the sound from its source, one can achieve suspense. The viewer is offered an audio cue signalling the action of a nearby entity, only without the visual backing to contextualize the movement. With this practice, it is possible to muddy the waters and purposely obfuscate the understanding of the situation. Much is left to the imagination when answering the questions of what is going on where and who is involved. The lack of a visual foundation renders the dislocated sound a blank canvas, where anything might be possible.

Incongruencies in the sounds can be used to great effect to achieve a desired outcome. Genre films have the tendency to overaccentuate foleys for different purposes. This can be achieved by exaggerating a given sound, adding characteristics that might not be implied in the image. This fracture between the visual and aural creates a sense of foreboding and uncanniness. This could be achieved by exaggeration in the loudness, weight, or even roominess.
Horror films very often utilize over-stylized reverbs that do not match the space of a scene. Besides evoking the feeling of disjunction, this effect can also imply the non-reality of the scene. Reverb is a very common tool for dream sequences, memories, and imagined scenarios. In horror, this can be exploited to great effect to undermine the matter-of-fact storytelling and not disclose the *realness* of a sequence.

3.4. Ambiances

Ambient sounds, or *ambiances*, are commonly used to illustrate the environmental sound of a given location or setting. They not only convey information relating to space and time, but also inform on the mood of a given sequence. An important thing to note is that the differentiation between score and sound design in horror is quite ambiguous; therefore, the two aspects will not be treated separately.
Manipulating spaces can be used to impact the mood of a scene. It is important to note that ambiances are not necessarily constrained by the spatial dimension of the room; they can, in theory, break the barrier and merge with more abstract soundscapes, depending on the situation. In the same way as the foleys, soundscapes can be manipulated in order to enhance the mood, frequently by humming and drones.
Infrasound can be used to alert to a possible threat. As these sounds fall outside the human audible spectrum, they are more felt than heard, which also adds a level of “unexplainable” unease in the filmgoers that cannot be formulated (Blog, n.d.).

When it comes to ambience manipulation, dissonances can come into play to bring unease. An important thing to note is that the differentiation between score and sound design in horror is quite ambiguous. Drones, pads, and field recordings can be tuned in such a way as to be arranged in a dissonant or inharmonic way, thus creating unresolved tension.
Ambiances can serve the same purpose as the dislocated foley handled previously. Through soundscape manipulation, something audible, but not yet visible, can be implied. It is very common for sound designers to craft these soundscapes by manipulating organic or human sounds, pitched and slowed down, thus

creating eerie textures that are subconsciously linked to something known, but experienced as an acoustic mystery.

Sometimes, the most unnerving ambiance is no ambiance at all. Silence is a very powerful tool for creating tension. In jump scares, it is utilized to widen the dynamic range, making the startle more intense. On a cognitive basis, it places the viewer in an anticipatory state (Sbravatti, 2019). Silence as an acoustic phenomenon is a stranger to our ears. As humans, we are not used to inhabiting quiet spaces. When crafting ambiances in films of any genre, a quiet room is never silent. Subtle environmental sounds fill the sonic space to prevent total silence. A notable cliché is the ticking of a clock filling the room: it is so quiet that all you can hear is the clock. Sound designers might use this to their advantage to create further tension.

3.5. Quantitative Analysis

Commonalities in horror sound design can also be observed and described quantitatively. The ultimate goals of creating fright and scare are observable in visualisations of the soundtrack, like amplitude envelope and logarithmic amplitude spectrogram.

Redfern's (2020) *Quantitative Analysis of Sound in a Short Horror Film* is an insightful study that conducts such analysis on the horror short film *Behold the Noose* (Brooks, 2014)(Figure 6).

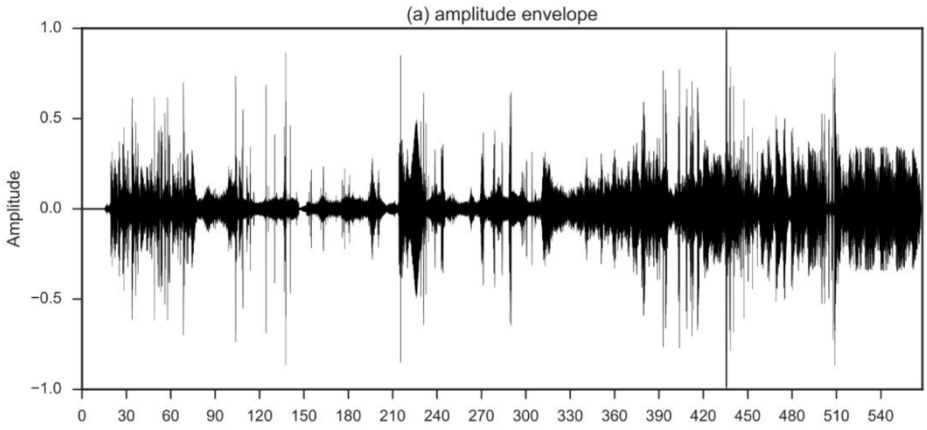


Figure 6 Amplitude envelope in Behold the Noose (Redfern 2020)

By means of the amplitude envelope, the amplitude (normalized) over time can be represented visually, showing clearly the dynamic range of the film in question. The amplitude distribution is widely varied, characterized by sudden peaks throughout the duration of the film. The visualisation also provides insight into the rapidity of the changes in amplitude. Loud parts can suddenly go quiet and vice versa.

Even without knowing the film in question, we could assume that from the time 150 to 210, a buildup to a reveal or a scare is happening. This can be assumed because of the subsequent spikes in amplitude after time 210. This segment is a perfect visualisation of a jump scare sequence. For the best effect, the difference in amplitude should be maximised, as it is the sudden, unexpected increase in amplitude to startle the viewer, and not just the loudness itself.

The amplitude envelope furthermore visualizes the degree of action intensity of sequences throughout the story. That is why, for example, the ending tends to be louder.

Another interesting visualisation is the log amplitude spectrogram (Redfern 2020) (Figure 7), a visual representation of a signal's frequencies over time, where the colour scale represents volume on a logarithmic decibel (dB) scale.

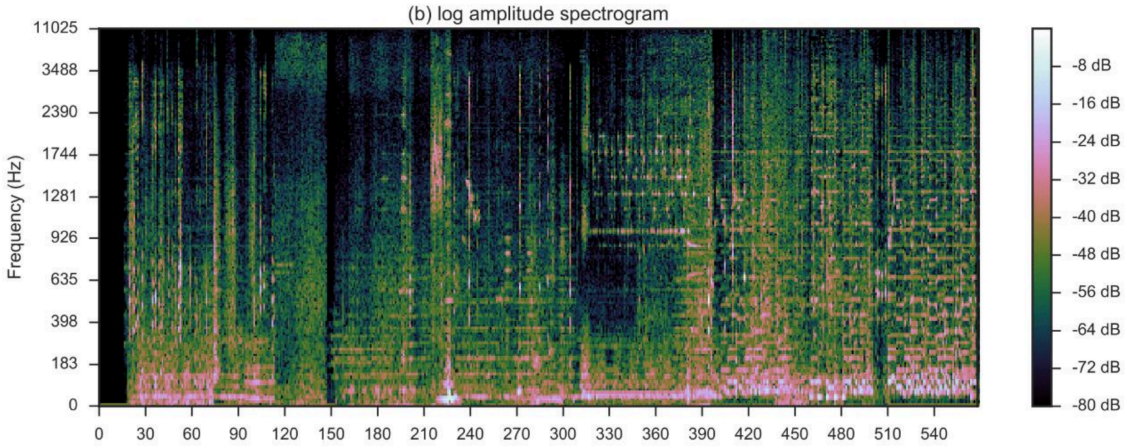


Figure 7 Log amplitude spectrogram (Redfern 2020)

Throughout the runtime of a film, the log amplitude spectrogram reveals a disproportionate amount of energy in the low end. Both in the sub-bass (20Hz - 60 Hz) and in the bass (60 Hz - 250 Hz).
The constant presence of energy in these parts of the audible spectrum, continuous and unflinching, is designed with the intent of generating a sense of instability and unease. Excessively low-end is also used to add an uncanny effect to otherwise realistic actions. This is due to the exaggeration and misalignment of the chosen sound with the action portrayed on the screen, which adds punch and gravitas that the realistic sound would not be able to provide otherwise.
A further aspect that is interesting to notice is the presence of clusters in the mid-range (1 kHz - 2kHz), during the jump scare scene. Not coincidentally, this is one of the frequency ranges where our hearing is most sensitive, thus rendering the sudden burst of sound that much louder to our perception.
An interesting detail can be found in the distribution of ambiances and background sounds which are handled very dynamically. For example, in the buildup to the jump scare, they become the primary element up until the payoff, mainly due to the absence of other notable sounds that would show up in the spectrogram. What is left is a somewhat even distribution throughout the frequency spectrum, with some presence in the low-end.
An unnatural lack of energy can be visualized in certain sequences, like from time 120 to 210, where a noticeable weakness between 1.5 kHz and 3 kHz is measured.
Such dips in the audible spectrum could accomplish two possible goals. The first could be the inverse of the cluster in the mid-range, which is to relieve that specific area of the spectrum, therefore lowering the load on the listener, in turn making the payoff later more impactful. The second could also be to create an unnatural imbalance in the distribution of frequency, thus enforcing the uncanny factor and raising suspense.
Moreover, nearly absolute silence is noticeable at time 150, further illustrating the effective usage of silence in the context of horror films in order to create tension.

3.6. Conclusion

Sound design for horror takes advantage of innate human hearing and psychological acoustic perception to instil fear. Massive changes in loudness can transcend the immersion of the screen, startling the viewer. Dissonances and silence, both phenomena that our ears are not too keen on, create unresolved tension. These sonic predispositions for fear have developed into their own aesthetic language that is easily understood and reproduced. In the case of dissonances, the stinger is the culmination of the aforementioned factors.

Stingers have cemented themselves as the de facto horror sound, which has become a cultural signifier for anything pertaining to the realm of horror, fear, suspense, and so on. As we've explored in this chapter, the success of the stinger is based on its characteristics relating to psychoacoustics and, as a direct consequence, bears cultural significance according to canons dictated by Western harmony.

Sound creation is often related to biological sources, often reaching degrees of unrecognisability, nevertheless retaining its naturalistic texture. The biological sources can also be used to connect us to our bodies and our own fragility, thus enhancing the spectacle of the violence portrayed.

Quantitative research (Redfern, 2020) also suggests that the mechanisms of sound design for horror films can be defined by their psychoacoustic effect, based on amplitude and frequency distribution. Sudden changes in loudness designed to startle the viewer are visible in envelope amplitude graphs, illustrating how they operate as a function of time during the runtime of a given film. Additionally, it visualizes the energy of the soundtrack in relation to the acts and sequences of the film.

The log amplitude spectrogram indicates how the intentional maximisation or reduction of certain frequency ranges can have a direct psychological impact on the viewer. This exemplifies how the aesthetics of sound design for horror films are deeply meshed with their functionality, and most importantly, how they can be separated from the pre-established context of horror aesthetics and yet have the same functionality.

Although plenty of elements like stingers, drones, hyper-realistic and dislocated foleys, and manipulated ambiances present a strong semantic connection to the genre altogether, thus delineating strong context-dependency, they find strong backing and reasoning in the psychoacoustic effects of said elements. It can therefore be theorized that the development of these semantic elements has its roots in the need to generate specific sound-dependent affective states. Said affective states are both influenced by reactions to sudden changes in loudness and specific frequency overloads, and the interpretation of certain sound incongruities.

4. Comedy

Comedy finds its roots in ancient Greek theatre, and shared the same structure as Greek tragedy. Comedy primarily handled the concerns of everyday people, like food, money, sex, and freedom. It would later on take the form of political satire by poking fun at their leaders after Athens' defeat in a war. This would come to be known as the "New Comedy", as opposed to the "Old Comedy". The New Comedy introduced the concept of recurring characters, or archetypes. A fundamental element in comedy are these recognisable

characters, which would endure over the following millennia. Archetypes such as the lecherous old man, wily servants, young lovers, and courtesans with a heart of gold would be the centre of attention during the times of the Roman Empire and the Middle Ages.

These times saw the fall of formal theatre in Europe, and was practised mainly by roaming actors, travelling through the streets of the continent, performing and improvising, leaning heavily on the pre-established characters.

In Italy in the 1500s, the roving actors crystallized into what became *Commedia dell'Arte*. With no known playwrights, every show was built on a simple scenario and then completely improvised. The actors wore distinctive masks and costumes and were also acrobats, dancers, musicians, and quick wits.

The characters birthed by *Commedia dell'Arte* became iconic and recognizable on account of their costume and characterizations. Such characters can still be recognized in modern comedic icons.

Arlecchino, the fool, can be considered to be the ancestor of figures such as Charlie Chaplin and Jim Carrey. Brighella, the aggressive schemer, is recognizable in Ralph Kramden, or also in characters like Barney Stinson from the TV series *How I Met Your Mother* (Bays & Thomas, 2005–2014). Pantalone, the lecherous or miserly old man, can be seen in characters like Archie Bunker, or perhaps even Statler and Waldorf, the heckling old men from *The Muppet Show* (Henson, 1976–1981).

The usage of archetypal figures lends itself perfectly to modern media like films and television. Sitcoms in particular thrive on this, as the format of small episodes does not allow for constant exposition of characters. Instead, the archetypes guide the audience towards an instinctual understanding of a character's identity and function. A perfect example of this would be Kramer from *Seinfeld* (David & Seinfeld, 1989–1998), whose flamboyant and goofy entrances into the scene make him immediately recognizable as the fool (Arlecchino), without any prior information about the character's identity.

Commedia dell'Arte was defined by the physicality of the characters, on which it depended for both comedy and recognisability. The specific physicality of the character (and by extension its characterization) was also what would dictate the plot of a scene. This ties *Commedia dell'Arte* directly to modern slapstick comedy, and its widely recognized connected figures.

As a parallel tradition to *Commedia dell'Arte*, the University Wits emerged in England, where the writers, due to their University education, would focus primarily on wordplay and wit. Shakespeare absorbed both influences, fusing the academic wordplay with the physical clowning of the Italians. The physicality of the *Commedia* troupes in England was most prevalent, as they were unable to speak the language; they mainly resorted to pantomime for their comedy.

Wit has also remained a staple of present-day comedy in any format, as language is the simplest way of creating a setup for subversion (Kaplan, 2013).

Before the advent of synchronized sound for films, all comedies relied on the physical clowning of the actors, their bodies, and the personalities they were to convey. This era was defined by athletes, not witty comedians. The main comic object was the human body performing outstanding tricks. Notable legends of this era are, of course, early Charlie Chaplin and Buster Keaton.

The advent of sound gave a literal voice to comedy, which became far more literary. It was finally able to appeal back to essential tools of literature, which became intrinsic to a film's conception (Mast, 1979). This means that wit was back on the table in a whole new format from the Shakespearean theatre.

The introduction of sound not only brought dialogue; it slowly but surely started introducing sound effects to mark and accentuate the action, and also to take on a life of their own as comedic agents. No longer was comedy only relying on visual elements for the laughs, but it could also fall back on dialogue and sound to deliver jokes.

As mentioned in the Genre Similarities sub-chapter, comedy has many similarities with horror when it comes to ways in which to achieve a desired effect in the audience. Still prevalent is the need for setup and delivery, with the setup as an incongruity and then the delivery (punchline) as a subversion stemming from said incongruity. This does not, however, mean that in comedy a delivery is necessarily funny; sometimes the subversion in itself is the joke. This also goes for horror, where the payoff of a jump scare is not always a given.

Both genres are heavily reliant on perfect timing to achieve their intended effect, often employing the right amount of pause before a payoff, or an overall suspension of the action to allow the audience to sit in a situation to generate anticipation.

The setup and delivery system in horror is quite different in comedy. As opposed to horror, where the tension (set-up) is drawn out as long as possible to then deliver the final payoff, comedy tries to minimize the set-up stage and instead delivers multiple payoffs in a row. This is why expediting the setup with familiar characters, like in sitcoms, and familiar settings or tropes like Mel Brooks' parodies is so effective.

Sound design in comedy, although sometimes more subtle than in horror, plays an integral role in strengthening the relationship with the setup and delivery system. Sound design as a cinematic tool is integral in the accentuation of narrative elements. It can double down on the importance of an object through accentuation of a foley or sound effect, or it can also make jokes of its own through meticulous incongruity, non-fidelity, or off-screen quips. It can create a specific mood that can then be broken at will. And sometimes it can take a step back and leave a purposeful silence to delineate the awkwardness of a situation.

Some sound traditions predate cinema by centuries, but are still present in modern comedies; for example, the so-called slapstick

4.1. Slapstick

The tradition of slapstick, an abbreviation of slap-of-the-stick, is born as a form of comedy based on the pretence of physical violence. From the *Commedia dell'Arte* to British puppet shows, variety theatre, and subsequently animation, cinema, and television, slapstick has been used throughout the centuries as a way of making people laugh through representation of violence, wild antics and disorder.

Its sonic history comes from a physical device, used in the *Commedia dell'Arte* called *Batacchio,* which consisted of two thin pieces of wood joined together to generate a loud slapping sound when struck against something or someone (Figure 9).

Although this specific contraption also came to be called the slapstick, there have been many different iterations of noisemakers, and the term slapstick came to represent the type of comedy based on gratuitous violence.

The aim of the synchronized sound was, and still is, a means by which to enhance the physical violence; the perfect pairing of visual and auditory stimuli is used to enhance comedic effect. The subversion of slapstick sound is achieved by means of exaggerating the physical impact with non-bodily sounds, which is in no way congruent with the expected sound we might expect from abuses to the human body. The percussive material sound mocks the injury, rendering it more brutal yet trivialising it at the same time.

In silent cinema, slapstick was still a prominent style of comedy, leaning entirely on the visual physicality of the "clowns" on screen. In the late 1920s, when film and sound technologies were able to reproduce sound that was synchronised with the image, slapstick got back its primordial acoustic spotlight that had gone missing in the prior decades.

Vaudeville troupes like the Marx Brothers, Laurel and Hardy, and The Three Stooges flourished on the screen, using the idiosyncrasies of slapstick. Synchronized sound gave way to much more varied sound possibilities than a simple wooden stick. Sound effect editor Joe Henrie, who is credited for creating the sound effects for The Three Stooges, did so by means of a wide array of different sources like guitar strings, whip cracks, mechanical ratchets, ripping cloth, and metal bells.

Animated cartoons of the era also implemented the same style of sounds to underline cartoon violence. The associations of these types of sounds are closely related to cartoons, so much so that they are often deemed cartoonish (Lombana Bermudez, 2008).

Although this specific type of sound for slapstick has fallen out of popularity, the term *slapstick* has survived, becoming part of comedy terminology. There are, however, some peculiarities about these sounds that are worth noting and that are relevant to modern sound design practices. The first particularity is the schizophonic mimesis of the sounds. The sound is detached from its source and recontextualized. The new context given to the sound, although incongruous, allows for the sound's mimicry of the visual action. This creates a stark disconnect between what is expected based on experience and what is heard. This incongruity is what makes the violence funny. The recontextualized sound becomes, therefore a metaphor, a decorative substitution of the literal expression of the action portrayed.

The second factor is the intensity and loudness of the sounds; they are in these contexts very loud and hard-hitting. This adds a layer of absurdity to the already incongruous sound, pushing it to the foreground of the scene, blowing it out of proportion. The excessive loudness also implies an excess of force of the violence that is not completely compatible with the actions by the characters. The same thing cannot be said for the receiver of the violence, whose extreme facial expression and physical reaction to the pain are paired with loud, desperate screams. This whole reaction moves to the absurd on account of its sheer clownery, further exaggerating the violence.

These principles relating to this period of slapstick sounds are still relevant in general comedy. What Joe Henrie and other sound editors were doing at the time was a deliberate complete incongruity, matching the action to a sound that is very noticeably not realistic with respect to the action. The incongruity can nevertheless still be exploited in order to achieve a similar comedic effect. The sound can be layered with schizophonic sounds to add a certain oddity to the action, therefore mixing fidelity and non-fidelity.

Non-fidelity can also be achieved by pushing the intensity of a sound, rendering the action acoustically inflated, creating disconnect by means of louder sound, or simply pushing certain frequencies. For example, pushing the low frequencies of the impact can give a hyperbolic weight to the action, which deliberately clashes with the performance.

4.2. Sound Gags

Although slapstick can rely heavily on sound design to enhance violent comedy, the driving force is the performance of the actors and actresses, with sound as the cherry on top.
Sound gags are when the main source of laughs comes from the sound rather than the visual. They inhabit the gray area between diegetic and non-diegetic. Let's take as an example again the "Frau Blücher" gag from *Young Frankenstein*.
Although it's clear by the first iteration of the joke, where the horses' reaction is visible, and in the repetitions of the joke with Dr. Frankenstein's reaction, that the sound of the horses is indeed diegetic, the joke works as an intrusion of non-diegetic sounds.
This joke is parodying established tropes from horror films of the villain's stinger. A villainous, monstrous, or untrustworthy character is introduced with the accompaniment of a screeching sound or a musical stinger to illustrate their evilness. The "Frau Blücher" joke parodies this exact trope by having the horses generate the screeching sound, and then repeat the sound even off-screen whenever a character simply says her name.
This is the reason for this intrusion, where an expectedly non-diegetic sound is placed in the diegesis, subverting the expectation by having the characters also react to this absurd sound, thus making the whole scene more humorous.
This sort of joke only works because of the semantic relationship between filmmaker and audience. The filmmaker assumes common ground with the audience, whereby they understand his employment of this understood trope, to then flip it on its head. This sort of gag relies on semantic subversion, which is very common in parodies, where semantic aspects are a reason for scrutiny from the sound and music department.
Without the semantic context, this joke could perhaps still work, because the incongruity is still there, i.e. horses neighing every time a character's name is mentioned, is not congruent with our day-to-day life experience.
There are, however, sound gags that don't rely on semantics, but are built more on perceptual incongruity.
A prime example of this would be the string of glass door jokes from Jaques Tati's *Playtime* (1967). The recurring character from Tati's filmography, Monsieur Hulot, mistakenly shatters the glass door of a fine dining restaurant. Standing in embarrassment where the door once was, an unsuspecting customer entering the restaurant grabs the door handle still in Hulot's hand and performs a door-opening motion, assuming the door is still intact. Hulot accompanies the motion as if the door were still there, helping the customers enter the restaurant. Although the door is completely shattered, rendering the gesture already absurd and humorous, the sound mix does something unexpected. The volume of the diegetic jazz music coming from inside the restaurant gets louder, as if the imaginary door was dampening the sound (Figure 8).
This joke, although extremely subtle, exemplifies how sound gags can also be done through the use of perceptual incongruity. An added feature again is the disrespect of diegesis. Although the music is established to come from inside the restaurant, its logic is shattered through this absurd joke. In this case, the subversion came from a source that was unexpected because of the trust that we have as viewers in the logic of diegesis.



Figure 8 Glass door scene Playtime *(Tati, 1967)*

This next example encompasses both semantic and perceptual subversion, and was notoriously born from budgeting limitations. It is, of course, the opening scene from *Monty Python and the Holy Grail* (Gilliam & Jones, 1975), where King Arthur's arrival is preannounced by the sound of horse hooves approaching. But once he finally appears, he is not actually sitting on a horse, but he is mimicking the physical movements of a man riding a horse, and his squire behind him is producing the sound of hooves by beating together two halves of a coconut (Figure 8). The subversion is indeed multifaceted. On one hand, it is a perceptual subversion, as the sound establishes the expectation of a valiant knight on his noble steed, but instead it reveals that the horse is non-existent and the squire is conjuring the horse sonically. Only once we see the coconuts is the subversion completed. But on the other hand, the joke is also a semantic subversion. The banging of the coconut is a long-time established foley effect that dates back to the early days of radio (Ouellette, 2025). By showing the squire producing the foley, the joke flips the relationship between sound and source on its head by providing the actual source of the sound as the subversion. The other semantic aspect that is subverted is the expectation that a sound doesn't necessarily have to come from the source, but rather from a source that provides fidelity to the visual counterpart. The joke brings to our attention the relationship we have to said fidelity by misdirecting us with what we think the foley represents, rather than what it really is, and subverts the semantic convention that hoofbeat sounds imply the presence of the horse. This gag also comes with the added perk of working as a metacinematic subversion by explicitly announcing the presence of foley, which would otherwise be "invisible".



Figure 9 King Arthur and his squire in Monty Python and the Holy Grail (Gilliam & Jones, 1975)

To sum up, the sound gag is an effective way to exploit incongruity between what is expected and presented. In the context of comedy cinema, especially parodies, the subversion plays into what is expected from sound conventions of the parodied film or genre. It can purposefully target the semantic aspects of sound design to deliver a joke. In some rare instances, however, it can subvert our perceptual expectations based on the audience's experience with sound in everyday life, therefore detaching itself from the semantic context of a film.

4.3. Silence

We have seen how silence or near silence in horror can influence the tension of a given scene, especially when followed by an acoustic startle. In comedy, silence can fulfil a similar function for an entirely different effect.

Silence in films is usually all but actual complete quietness. It's not simply the complete absence of sound but rather a constructed, intentional sound perspective that signals silence. Although the loudness of such moments is certainly very low, the sound never dissipates completely.

In most instances of quiet scenes, faint sounds still inhabit the space. Ambiances and room sound usually remain present, to give at least a contextual noise floor that informs the audience about the space of the scene and the situation. With the shift into a quiet moment, the realm of what becomes audible changes. Tiny, detailed sounds, coming from either the environment or the characters themselves, become more audible, signalling the quietness of a moment by making such a soft sound audible to the viewer.

As we saw with the example from *Nosferatu*, silence can shift the viewer's focus more closely towards visual cues, in that case on the character's face and facial expressions. This can also be used in comedy to great effect to shift awareness to more visual cues. For example, it can be used after a joke to emphasise a character's facial expression or reaction to the situation. It can give space for the actor's physical performances to shine, unmeddled with by over-stylised audio cues that might negatively impact the scene (Nicole, 2025). Oftentimes, the quietness after a delivery can add moments of awkwardness. The American television series *The Office* (Daniels et al., 2005–2013) uses awkward silences to great effect to further underline the absurdity of a situation.

Silence is also used as a tool for pause and dramatic tension before a punchline. This follows the same structure as horror: setup and delivery, whereby a pause before the delivery can greatly increase the effect of the payoff. This is closely linked to both genres' need for precise timing for the delivery; too early and the impact is not strong enough, as the setup hasn't sufficiently settled in, or too late, and tension is lost due to an unreasonably long waiting period.

Unlike horror, comedy is much more reliant on dialogue and wit. Comedy films featuring large amounts of dialogue will make it a priority to have the line untainted by external factors.

In conclusion, silence is a very powerful tool of great effect in comedy, which can let a situation breathe and let the physical and verbal performances truly shine. Sometimes the best thing to do is nothing at all.

4.4. Dialogue and performance

Due to comedy's reliance on dialogue delivery and performance, it is paramount for a sound designer to keep the dialogue as clear and understandable as possible. This can be achieved by minimizing the possible interference and overlap with other sound elements, as stated in the previous paragraph on silence.
This is, however, not always an imperative. There are some instances where the original voice of the actor or actress can be swapped or manipulated.
Let's take, for instance, *Sorry to Bother You* (Riley, 2018), where a young African-American man gets a job at a telemarketing firm. At first, he struggles with customers until his older coworker suggests he use a "white voice". In that moment, his voice turns to a clearly overdubbed version, performed by other actors. The overdubbed performances are high-pitched and nasal, and employ an exaggerated "white vernacular". The result is a performance that is uncanny as much as it is funny. An incongruity is being created between the characters' mouth movements and how the voice sounds. This incongruity is enhanced by the opposition of the voices before and after.
Furthermore, the joke plays with racial stereotyping and prejudice. The idea that people do not want to hear an African American's natural voice but rather a white, "less threatening" voice. On the other hand, however, the vocal substitution itself works as a benign violation, which can be considered universally funny.
As a finding from Bachorowski et al. (2001) illustrates, participants exposed to movie scenes with the fundamental frequency of the speech shifted higher had a higher likelihood of laughing.
This means that although part of the joke relies on semantic subversion, steeped in cultural and racial views and attitudes, the qualities themselves of the subversion are also part of the joke. Therefore, suggesting that the subversion is not only cultural, but rather an acoustic one as well.

4.5. Conclusion

The development and history of comedy has always been defined by raw comedic performances of the actors and actresses. Very much like horror, its intended affective state is theorized by incongruity theory, where a mismatch is created between what is expected and what is presented.
In comedy cinema, sound has the role of strengthening and creating incongruity of its own. Incongruity can be created by mismatching the scene with our expectations relating to everyday experience, or cinematic language itself.
Sound also aids in the timing and execution of the punchlines that are built on the ground of incongruity by means of pauses, tension, and misdirection. The strong reliance on visual cues from the performers renders the deliberate absence of sound a pivotal tool.
Comedy often relies on semantic subversion to achieve a laugh. This is especially true in the case of parodies, where the genre or film in question being parodied sets up specific expectations. For example, Lord Helmet's introduction in Space Balls, where he's expected to produce the same breathing sound as the parodied Darth Vader, is then subverted when he opens his gigantic helmet and yells, "I can't breathe in this

thing". The audience's understanding of Darth Vader is that his breathing is a sonic signature of the character, but in Space Balls this is put into question; the certainty of the character is violated by revealing that the sound of his breath is caused by the helmet, and he has to open it in the front to be able to breathe. This also plays into our certainty that Darth Vader cannot take off his helmet.

There are, however, many instances of sound generating benign violations independently of the semantic context. The slapstick sounds of Vaudeville subvert the expectation of the sound that would be emitted by physical action, opting instead for a metaphorical substitution. These are sounds that, although completely detached from acoustic reality, still relate to the original gesture. The twisting of the nose in The Three Stooges is accompanied by a cranking sound. This, although unrealistic, still employs a sound that carries the notion of rotation or twisting, which synchronizes with the action seen.

Although slapstick sounds themselves are, deeply semantically tied to the genre of comedy, the mode of its effectiveness can be applied independently of context, as it provides for an acoustic subversion which is incongruous with our expectations. This subversion can be achieved in two ways. The first is the metaphorical substitution of the sound with schizomorphic mimesis; the second is by acoustic exaggeration. These two ways of acoustic subversion are usually intertwined in the context of slapstick.

It's still open to debate whether exaggeration, of either pitch or loudness, is completely independent of the semantics of the genre, as they can be presented in the form of subversion of sound language as a whole.

Changes in pitch and loudness can also be employed in the manipulation of dialogue, resulting in a possible subversion that is entirely based on our experience with human language. It can, however, still be presented or framed in our context of how we expect certain people to speak rather than general human speech.

5. Final Considerations: Horror vs Comedy

Horror and comedy require sound design in order to achieve the desired affective states; they vary, however, in their context-dependency. Horror, since its foundational emotion is fear, can appeal to deeper emotional systems which are more rooted in survival instinct and biology. This means that even the semantic elements in horror, such as the stinger, can be deconstructed to their essential characteristics of pitch and loudness. Through quantitative and qualitative analysis, it was also possible to demonstrate how the semantic elements can be generalized into broader sound mechanisms employed to achieve the desired effect. This means that sound design for horror can be understood as less context-dependent.

Comedy, on the other hand, is a bit more delicate. The emotional engagement in comedy takes the form of laughter, which is reliant on higher neurological functions of reasoning. Evolutionarily speaking its theorized to be a reward system for discovering and double-checking incorrect assumptions. As explained by Hurley et al. (2011) in their book *Inside Jokes: Using Humor to Reverse-Engineer the Mind*

"Our brains are engaged full time in real-time (risky) heuristic search, generating presumptions about what will be experienced next in every domain. This time-pressured, unsupervised generation process has

necessarily lenient standards and introduces content—not all of which can be properly checked for truth—into our mental spaces. If left unexamined, the inevitable errors in these vestibules of consciousness would ultimately continue on to contaminate our world knowledge store. So there has to be a policy of double-checking these candidate beliefs and surmisings, and the discovery and resolution of these at breakneck speed is maintained by a powerful reward system—the feeling of humor; mirth—that must support this activity in competition with all the other things you could be thinking about."

What this means is that for us to be able to experience humor, we must be able to understand both the assumption and the incongruity, rather than just engaging with the stimuli emotionally.
This, when applied to the context of sound design for comedy, illustrates how much more dependent we are on the semantics of the genre.
This means that to generate incongruity with sound design itself, for example, putting horror sounds on a scene that is in no way visually scary, the incongruity necessitates the semantic understanding of "horror sounds" for the incongruity to work.
With this in mind, it's still possible for sound design to work outside of the semantic to achieve humour, and is far more dependent on our presumptions based on experience. With the necessity of shared understanding of said presumption and subversion.

6. Realistic films

6.1. First issue

The term "realistic" when applied to films is fertile ground for contention, disagreements, and misunderstandings.
"Assumptions about realism in the cinema are frequently tied to concepts of indexicality prevailing between the photographic image and its referent." (Prince, 1996)
What this entails is that realism is only dictated by the physical relationship between the object recorded and the resulting recording. It can be argued that an indexical relationship entails a naturalistic approach or result. What the apparatus captures in that specific moment of recording is the final product; no subsequent steps are added to enhance, modify, or correct the final representation. This understanding of realism, although interesting are incompatible with sound practices in cinema, as it rarely is the result of straightforward indexical recording (Altman, 1992, pp. 42-45).
In the field of sound, the rejection of indexicality opens the door for further discussion. Lastra (2000), in his book *Sound Technology and the American Cinema*, provides us with two possible models for sound recording and reproduction. The first model is phonographic, or the so-called "invisible auditor", which focuses on perceptual fidelity.
This model supposes that the ideal recording and reproduction situation would see "placing dummies with microphones for ears in different places around the set and recording multiple tracks. Theatrical playback

would require multiple stereo tracks, with each signal routed to the seat in the audience corresponding to the literal position of its respective dummy at the original performance on the set."

This model comes with its own set of relevant assumptions. One is the assumption that the spectator identifies themselves with the "position" of the camera and microphones. Secondly is that a film is a succession of autonomous, and perceptually specific "observations". Lastly, the model assumes that the auditor is part of the same space of the "original" performance, and that "the space and acoustic quality of the set are identical with those of the represented world".

The effect of the "invisible auditor" model would be absolute perceptual fidelity in relationship to the point of view of the camera; a consequence of this would be the lack of seamlessness from shot to shot. The shot becomes the basic representational unit, rather than the scene (Lastra, 2000, p.159).

The second model is the telephonic or "intelligibility" model. As the name might suggest, this model prioritizes narrative over perceptive fidelity. This sees the objective of reducing interference that could stand in the way of intelligibility, such as excessive reverb, noise, and speech idiosyncrasy. At the same time, maximising the "directness" or "frontality" of recording in favour of the dialogue, regardless of the relative position of the speaker, implies "that he or she is still speaking "to" the auditor, because speech is understood not simply as an abstract sound but as a sound with a specific social and narrative function."(Lastra, 2000, p.139)

According to Lastra, the misunderstanding concerning realism stems from a miscommunication regarding Hollywood's and sound engineers' conceptions of realism. Hollywood's realism often stresses formal unity and narrative plausibility. This inevitably clashes with the perceptual fidelity of the engineer's phonographic model.

He argues in favor of a shift away from reproduction and towards the idea of representation. Representation requires an active act of interpretation by the engineers, since even the "invisible auditor" is an interpretative act.

As a wider consequence of this, expectations of realism applied to sound are treated more as an unattainable ideology than a concrete aesthetic goal. This is because even indexical representations of sound necessitate acts of interpretation which are incompatible with the absolutistic concept of realism. What this ultimately reveals is that sound representation needs to abandon the objective of faithful collection of real events, but embrace the "construction of carefully hierarchized events whose realism is a function of their plausibility and their compatibility with conventions of narrative realism rather than literal, perceptual duplication."

Realism can then be understood as verisimilitude. As recounted by Todorov (Champagne et al., 1978), the meaning of verosimilitude does not simply stop at what is consistent with reality, but rather a relation to what people believe to be reality, in other words, public opinion.

What this means for filmmaking is that realism is treated more like an aesthetic or genre, in which verosimilitude does not stem from conformity to reality but from obeying rules that viewers recognize and believe to be reality.

6.2. Perceptual Realism

Realism is not defined by the film's absolute likeness to reality, as that implies numerous fallacies stemming from interpretation, but it's when the "moving images mime aspects of reality only to such an extent that they activate kinds of perceptual processes that would also be activated if the object considered was not a representational form but rather an unmediated visual experience." (Langkjær, 2010)

Sound design generally has a harder time adhering to this principle because, for the sake of storytelling in fiction, it is intrinsically mediated. The general aim for sound is actually the opposite of realism, as it amplifies sounds to the degree that it becomes an artefact itself, or rather, it is perceived as such.

This amplification of reality, or "hyper-realistic," is necessary for a handful of reasons. An exaggerated foley, which is produced during post-production from a different source, is done so for the sake of legibility. It's necessary when the same sound source as the object on screen is in some way perceptually incongruent or illegible as such.

Hyper-realistic sounds can also fulfill a narrative role, where the indexical sound might represent the sound as "realistic" but lacks the emotional impact that this sound might have in real life. This is more often than not applied to sounds for which the general audience doesn't have enough real-life experience to be able to discern reality. These are cases such as non-realistic objects or entities, like aliens, monsters, spaceships, and so on. Or real-life events of which most people might not have particular experience, like gunshots, physical violence, accidents, and so on. The hyper-realistic, or in some cases hyper-stylised, sound is necessary for conveying the emotional impact of these objects and actions, rather than how they would realistically sound.

Sound design practices generally profit from a high degree of perceptual tolerance concerning what can be counted as perceptually real, unrelated to the reality of the sound. The ultimate goal of perceptual reality is how the representation can be portrayed in a way that it feels like unmitigated reality.

What this entails is that perceptual realism, although "fake" in its essence, is achievable in a variety of different contexts that go beyond grounded filmmaking and can therefore be applied to more fantastical genres, so long as it respects the representation as an imitation of unmitigated reality.

To achieve the illusion of unmitigated reality, sound substitution needs to fulfill two requirements. First is synchronicity, where a sound must match the timing and rhythm of the action, so that the event becomes a single phenomenon experienced by multiple modalities. The second is likeness of the substitute to the original sound. What this entails is not that the replaced sound has to sound exactly like the real object. But rather that it contains likeness to our understanding of the object in question, and that can then be interpreted as such while simultaneously serving as an emotional and narrative tool. Synchronicity also aids with the contextualisation of the similar substitute.

One question remains open about what it means for sound to be non-realistic. Synchronicity and similar substitution are potential clues as to what that might be. Both requisites are in service to the logic of what happens in the scene. It's tightly woven into the logic of the world portrayed, even off-screen. For a sound to therefore be "unrealistic", it must disrespect the established setting. This can range widely across diegesis, but as we've seen with the horror and comedy chapter, this is not always relevant to the realism of a sound, as a diegetic sound can still be unrealistic.

A possible way of determining what is unrealistic, still in the context of imitation of unmitigated reality, is what falls outside this objective in favor of emotional impact. Acoustic startles, horror ambiances, dislocated foleys, slapstick sounds, and dialogue manipulation are all falling outside what one might consider realistic

as they craft a reality that feels deeply mitigated. They are more synchronous with the emotional timing of the story than any action portrayed on screen, or they might lack likeness to the original sound source in favour of the emotion of the action itself.

7. Practical Applications

7.1. Introduction

In the following chapter, I will give an overview of the process of creating sound design for two short films, in the context of the research on contextual dependency or independence of sound design.
The first film is titled *Talion*, directed by Micha Straub, and the second one is titled *Kloster*, directed by Salomé Käsemodel.
The first is a horror film, and the second is a comedy film, but with the particularity of both adhering to a sense of realism. The aim of the research is tightly linked to the process of sound design for these films, as both directors rejected the semantics linked to their respective genres, but still strived for the respective affective states.
I will go over the details of what was the process of creating sounds for these films by adhering, in part, to ideas of context-independent sound design.
An important thing to note is that the final product is still a mixture of both approaches. The priority was still to craft the best possible sound design befitting these films. There are some examples worth analyzing to illustrate how I was able to implement context-independent sound design, instead of purely relying on the semantics of the genre.
Spoilers ahead.

7.2. Talion

Talion is a horror-thriller film set in the forests of Jura, Switzerland. A young couple from the German-speaking region of the country, Aline and Marco, make their way towards the rich forests for a nice summer getaway. Along the way, they pick up a hitchhiker called Noé, an inconspicuous wildlife photographer, in search of animals to photograph. As they venture through forest-surrounded roads, they find themselves off-grid, where even the radio can't get any signal. They do a stop for a quick pee break on the side of the road as Marco spots a deer in the forest. Aline and Noé try to approach the deer as Marco goes to pick up his own camera from the car. After hearing a gunshot coming from deep inside the woods, they freeze, and out of nowhere, Noé is mortally shot in the face. Panic and confusion ensue as a hunter and his son approach, claiming it was an accident. Aline and Marco try to call the police, but there's no signal.
The hunter claims he has a satellite phone in his car, twenty minutes through the forest, but the couple deeply mistrust him. They argue at length as the tension rises, until they hear the satellite phone ringing from the hunter's pocket. Marco makes a run for the hunter's rifle and keeps the hunter and his son at gunpoint. Aline tries to call the police, but it seems not to be working. The hunter exploits a moment of distraction from Marco and makes a run at him, but gets knocked down by a blow from the rifle's stock. The son pins down

Marco in a brawl as Aline makes a run for the car keys, which are still in Noé's trouser pocket. As Aline sprints back, the hunter reaches for his son's rifle and shoots Marco. Aline, deeply startled, runs into the forest, and the hunter hunts her down. She's wounded in her arm but manages to run and hide from the hunter.

We cut to the son, clumsily dragging Marco's body to bury him. He hears a gunshot from deep inside the forest and pauses before going back to dragging the body. As he's clumsily trying to bury both bodies, Aline appears holding the hunter's rifle. She holds the young man at gunpoint, but lets him run away.

The film ends with extended shots of the beautiful nature, as we hear a faint, distant sound of a car taking off.

7.2.1. Approach

The general approach to sound design for Talion focused on diegetic sounds and ambient sounds to heighten the emotional impact of tense and shocking scenes. With that said, it can be argued that there are still some "non-realistic" sounds, such as drones, that are carefully shaped so as not to be front and center, but rather to play into physical perception.

For the sake of orderliness, I will go through the practices listed in the horror chapter and explain how I was or wasn't able to implement them into the film as context-independent, starting with the jump scare.

7.2.2. Noé gets shot in the face and dies

A striking scene for sure is the sudden and graphic death of Noé; it's shocking and seemingly out of nowhere. This is the catalyst of the film. My objective was to have the moment hit as hard as possible with as little anticipation of the event as possible. The objective was therefore not to use the setup as anticipatory tension, but rather as an acoustic setup for the gunshot to create wide dynamic range for an acoustic startle. This was accomplished by carefully shifting the ambient sounds of the forest. When the location is established, the forest feels alive: insects, wind, tree branches, birds, and so on. As Aline makes her way into the forest to reach Noé, some of these ambient sounds disappear gradually. The objective was to make this transition as subtle and believable as possible by using the cuts and their change of perspective as believable shifts in the soundscape.

The gradual attenuation of the soundscapes shifts the attention to the characters and their movements, but also to the gunshot in the distance. The gunshot causes the birds to fly away, and the characters are more aware of their surroundings. The ambiances follow the characters' expressions as they look at each other, trying to discern what is going on.

Out of nowhere, Noé gets shot in the face; the action is very fast. The impact is a full blast of impact sounds, gore, and sneaky sweeteners. This creates a very loud sound in contrast to the setup. The sheer graphic quality of the gunshot to the face, paired with the fast twitching of the character's body, allowed me to fully indulge in a strong acoustic blast that wouldn't come through as overly exaggerated. The reason for this also relates to the perceptual tolerance for phenomena that are uncommon in everyday experience. It focuses on the emotional impact that such a shocking event could represent for the characters, and in turn the viewer.



Figure 10 Noé getting shot in Talion (Straub, 2026)

A nice detail, insisted upon by the director, was the gunshot sound arriving after the bullet, almost like two separate events. This is mostly grounded in physics, as the speed of a bullet should be approximately threefold the speed of sound. With this information, we were able to calculate a delay time that could be entirely realistic.

The whole sequence is built by exploiting and staging diegetic sounds and carefully accentuating synchronous sounds, in a way as not to create excessive incongruity or force intention or emotion by using non-diegetic sound effects.

7.2.3. Where's that hunter, and why is he so heavy?

The foley work started as a simple procedure: spotting and recording the sounds with the simple objective of mimicking the action on screen. This was done in a way that favoured legibility over realism. Many elements were then enhanced during the foley editing and mixing stage. My objective was twofold: one was to give the characters different sound qualities to create opposition and communicate characteristics about the characters, and the second was to dislocate the foley from the sound source to achieve suspense.

For the first objective, I was aided by fashion choices made by the production designer.

The main conflict and contrast is between the couple and the hunter, and as it just might happen, this contrast is already present in footwear. The couple wears similar urban sneakers, which perhaps also showcases the mismatch between identity and environment. Whereas the hunter wears heavy hiking boots, more consonant with the forest. The boy, funnily enough, also wears sneakers, which portrays how ill-adapted he is to the situation as he slips and tumbles while trying to cover the corpses in dirt.

The resulting foley sounds recorded from different shoes resulted in a lighter footstep sound for our sneaker-wearing characters, which also feature some squeaks as they step on branches, dirt, and gravel. This gives them a subtle sense of instability and unpreparedness for what is about to ensue.

The hunter, on the other hand, who is the main threat in the film, has a heavier step on account of his boots, giving a sense of competency and strength.

I then accentuated these already present traits stemming from the choice of shoes themselves by making the urban footsteps slightly thinner by removing some low-end, and the opposite for the hunter, for which I accentuated the low-end to add weight and size.

Some treatment of foleys was also driven by narrative importance; for example, the sound of the keys when they pick up the hitchhiker is a bit louder than it should be on account of its relevance later in the plot.

Some foleys were also dislocated to create suspense, like when, after Noé gets shot, Aline is scanning the woods in search of the culprit. I purposefully placed sounds of footsteps breaking branches, as if someone was walking rapidly and with great force, drenched in a forest-like reverb. This creates suspense as the dislocated foley anticipates an unknown presence that is then resolved once the hunter appears in the frame.

The film omits certain information, such as Aline being able to take down the hunter and steal his rifle. This happens as the son is dragging Marco's body through the forest. And he hears a gunshot in the distance. The gunshot sounds naturally distant, but it is very murky as to what is actually happening with Aline and the hunter.

7.2.4. Why so many crickets?

From the conversation with the director, we came to an understanding that background sounds deserved some special attention. They are, in fact, not only in service of placing the scene in its location's soundscape, but also in service of driving tension and release throughout the film.

The opening scene plays entirely in the car, and I intended to make the film tense from the start, abiding by the principle of priming. The sound of the car and the engine is very deep and rumbly, with extra emphasis on the sub-bass. This was done with the idea of misdirection: that something bad might happen in the car, when it actually happens later. The sensation I was striving for was the sensation of having your head leaning against the window while riding in the car, where the sound is also transmitted via vibrations through the skull. Every bump of the road and hiccup from the engine is felt very closely, and the rumbling is felt in the body. I emphasized this bodily experience of the car engine by sending the sound to an LFE track. This creates a slight incongruity, but it is still entirely believable, as it's deeply tied to the general experience of riding in a car.

As I already explained in the setup process for the jump scare, I manipulated background sounds to guide the emotional journey of the film. This is done by automating volume, mixing different sounds for different moments, and processing certain recordings.

Wind was used to ramp up the moment before certain escalations, adding an audible fullness to the acoustic field in anticipation of certain plot points.

Rather than using excessive dissonant drones to increase tension, I was drawn to the fauna of the forest to act as my own personal dissonance library. Crickets and mosquito sounds produce long, continuous sounds that can be slightly pitch-shifted to fit within the hearing range where we are most sensitive, between 2KHz and 5KHz. By introducing these sounds gradually, I was able to create further tension and give more intensity to the already intense moments of escalation, like Marco reaching for the gun and the hunter charging at Marco.

I experimented with different background sound manipulations to create drone-like sounds that could still be perceived as naturalistic; for example, by slowing down sounds of foliage being moved by the wind to get a

deeper, uncanny characteristic to the recording that still fits the environment. These processed sounds are still layered with more naturalistic sounds to "smuggle" the sensation of the drone while still maintaining immersion in the environment.
Some more "classic" horror drones are still present, especially right after Noé's death. These are, however, heavily focused on the sub-bass region to give the bodily sensation stemming from infrasound stimuli.

7.2.5. Frequency distribution

Many emphases on certain frequencies are based on quantitative research. Particular attention was given to the sub-bass region by often having background or drone tracks routed to the LFE track according to the intensity and anticipation of certain actions and emotions.
Further attention was given to the 2KHz to 5KHz frequency range by attempting to overload this region whenever it seemed fit. This was accomplished by a mixture of sounds already charged in that specific range, pitch shifting, or emphasizing the target frequency with equalization.
Most of the sounds that were used to target these specific frequencies were still congruent with the actions and environment in the film, without resorting to genre-specific sound effects, but rather by altering diegetic sounds to serve the functions that many of the examples seen in the horror chapter exhibit.
A further point for experimentation regarding frequency distribution was the scooping of the mids in the sound field, thereby leaving only the low end and the very high end. The idea behind this is the crafting of an uncanny and unnatural frequency distribution that can cause anxiety and tension.

7.2.6. Outcome

The outcome showcases the possibilities of applying certain sound design practices for horror, in a context-independent. This can be mainly achieved by applying the sound characteristics of semantic elements from horror films to diegetic sources. The manipulation of these sources can, however, cause incongruities that can be perceived as mitigated, despite the best efforts to try and hide such mitigation. This is entirely dependent on the degree of manipulation of the sounds, especially pertaining to generally experiencable phenomena. Whereas non-common experiences benefit from a higher degree of tolerance.

7.3. Kloster

Kloster is a comedy short film set in a convent in Tobel-Tägerschen in the Swiss canton of Thurgau. It's shot in a mockumentary style and plays out as the beginning of a murder mystery. It was conceived as the pilot episode of a potential subsequent TV series.
For the anniversary of the catholic convent of Tobel, a local TV station follows the story of this community on the brink of collapse
They follow Priorin Scholastika, the prioress of the convent; she's hard-working, dedicated, and a bit silly.

As she's interviewed by the TV crew, she tells them about the convent and introduces the rest of the ensemble in a series of cutaways. She introduces us to Schwester Hildegard, the oldest and only sister of the convent. Hildegard is very old and obviously senile, with some gastrointestinal problems. Looking after her is Toni, a young man working at the convent as part of his civil service. He introduces himself to the film crew while the film shows us his responsibilities, such as taking care of the garden, scrubbing the floors, feeding Hildegard, and giving her her laxatives.

During the night, as Scholastika is showing the TV crew the view of the convent at night, they spot a cloaked figure lurking in the dark; they follow the figure as Scholastika suspects it is one of the "Skater Girls" who have already vandalized the convent before.

They follow and corner the figure at the entrance of an underground cellar at the bottom of a staircase. Scholastika unveils the figure, revealing that it is Toni himself sneaking at night. Scholastika suspects him of being a Skater Girl, but he drops his backpack, revealing suspicious pills. He is defensive, declaring he's not a junky, nor a dealer. The door swings open to the sound of dubstep, revealing Lee (They/Them), a hacktivist wanted by the police, who was being sheltered in the convent by Toni. Toni presents them as long-time friends, but Lee corrects him by saying that, to them, Toni is just their ADHD medication dealer ("rita-plug").

The next day, Toni is helping Scholastika by filming a live stream in the courtyard, intended to get more women to join the convent. As she introduces Sister Hildegard live, who was previously rocking back and forth on her chair praying in Latin, Hildegard falls from the chair and dies.

During Hildegard's funeral, it's revealed by Brigitte Braunschweig, a representative from the municipality of Tobel-Tägerschen, that since there's only one nun left in the convent, the municipality has conspired with the priest to turn the convent into a luxury yoga retreat. As the conversation gets heated, Lee barges through the church doors, screaming at Toni that they've spent the whole day on the toilet. Lee reveals the box of pills they took as laxatives intended for Hildegard, meaning that he must have killed Hildegard by giving her Ritalin.

The film ends with the local police officer, Yvonne Gantenbein, who, after talking on the phone with Scholastika, concludes that it must be murder.

7.3.1. Approach

The plot of the film is presented in a very chopped manner, as the editing prioritizes the staging of funny situations over narrative continuity. This works really well in the film's favour as it fits perfectly the mockumentary format, where one situation can be shown while having the voice-over from an interview. The staging of a documentary is the primary reason for attempting to have a realistic approach to sound design, which wouldn't be overly indulgent in the semantics of comedy but rather would be in service of the absurd situations, strengthening it by expanding on grounded pre-established ideas.

Since the focus on comedy leans more towards singular, independent humorous situations, the rundown on the sound design process will be more free-form than Talion's.

7.3.2. Extension, substitutions, exaggeration, and other absurdities

Many of the funny situations established by the film were possible to amplify with simple sound design choices. One of which is the decision to extend certain sounds from shot to shot.
In the beginning of the film, while Scholastika introduces herself, there's some very holy-sounding church organ music playing as a non-diegetic soundtrack. But as the interview continues to other shots of her praying in her room or at her desk, the music continues as diegetic. When she's praying, I placed the music as if it were coming from the next room over, and when she's at her desk, the music can still be heard from the headphones she's wearing. The music from the headphones is also treated in a way that it feels like she's listening to the organ music at an excessive volume.
This simple choice works by bringing a slight absurdity to the convent and Scholastika; the church music is omnipresent, and even when it's not, she still listens to it. The continuation of the music also grants a good rhythm to the scene transitions by tying one shot to the next.
Another example of continuation comes from Scholastika herself. As Toni is feeding Hildegard and talking about himself, Scholastica is there in the room chanting aloud. In the subsequent shot of Toni accompanying Hildegard out of the eating hall, I kept the chanting going by stitching together different takes. The chanting even continues in the courtyard where Hildegard is peacefully smoking, and Scholastika can still be heard singing inside.



Figure 11 Hildegard smoking in Kloster (Käsemodel, 2026)

This extension works as the subversion lies in the persistence of her singing, as even after a non-specific time jump, she's still going at it.
A little gross-out humor never hurt anyone. In the case of Hildegard, I overdubbed all her mouth movements from eating and smoking with my own voice. This was done with the intention of bringing closer focus to the stillness around the convent, and to exaggerate her elderly senile state slightly. Old people are often chronically dehydrated and produce quite gross smacking mouth noises. This was also done to give some idiosyncrasy to an often silent character, which also includes the very necessary fart sound when she's suffering from stomach problems after eating some poisonous berries.

A lot of the humour comes simply from the incongruity of the situations, communicated through either dialogue, character action, or character opposition. A little purposeful addition of foleys and sound effects can really enhance a given situation. An example is the footstep sound of the municipal representative Brigitte Braunschweig, whose chic attire is in direct contrast to the catholic attire of Scholastika, the priest, and even Toni, who's wearing altar boy attire. To strengthen this contrast, I gave Brigitte very sharp, high-heeled shoe-step sounds, which in turn reverberate massively inside the church.

The TV crew also gets its own sonic characteristic. All the movements of the camera, such as shaking and zoom get their own noise. These noises would not be audible in a documentary. My decision to add sounds to these specific motions was to characterize the TV crew as somewhat incompetent, cheap, and clumsy, thus adding more incongruity between the seriousness of the character's intentions and the unseriousness of the production portraying the actions.

Some of the principles seen in slapstick have also been applicable, though not as visible as physical violence. Hildegard's sudden death happens just outside the frame; the information of her falling from the chair is entirely communicated via sound. The sound of her falling from the chair is comprised of different layers of sounds coming from an array of different in-world sources: the body hitting the floor, the chair tipping over, the clothes flapping, the scuffing of the shoes. The combination of sounds gives a quite large and exaggerated impact; although it's not a metaphorical substitution like in classic slapstick fashion, it retains and imitates the characteristics of loudness and strength.

The same thing happens when Lee angrily barges into the church. The door hits the wall with great intensity, halting the action and setting up her screaming for Toni.

In summation, a lot of sound treatment stems from our active understanding of the characters, and the purpose of which is to enhance what we already know about the characters' particularity. Scholastika is devout and determined; Hildegard is old, senile, and gross; Toni is goofy and awkward, Brigitte is pompous, and so on. These added elements of extension, substitution, and exaggeration aid in the communication of the traits of the characters, which widen the incongruity established by the scene itself.

7.3.3. R.I.P. little bird (2026-2026)

All the examples from the previous sub-chapter are closely related to the actions, situations, characters, and objects represented on screen. Unlike these past examples, this next sound gag delivers the punchline entirely through sound, based on visual setup, and in turn recontextualizing the following scene.

The night scene is established by a shot of a bird perched by a window, which then flies away. The following shot is a wide shot of the monastery at night. During this shot, we still hear the bird flying around until we hear it fly straight into a window and fall to the floor. This is entirely unexpected and in direct contrast to the calm and peaceful "holy" atmosphere of the convent. The contrast is only made stronger later by Scholastika's insistence on showing the beauty of the convent to the TV crew.

This sound gag was not intended by the director, but my proposal of it was accepted with open arms.

7.3.4. Let the face do the talking

Certain scenes truly benefited from a step back from sound design, and leaving space for the physical performances. This is especially true for Toni, whose facial expressions of discomfort, awkwardness, and guilt made for truly funny moments. While the other characters are often very active, he's more reactive. These shots are very reminiscent of the "Jim Stare" from The Office. An example of these shots is when Scholastika unveils the illegal sheltering of Lee in the basement and says, "Somebody has to go confess on Sunday, huh?" The camera then pans to Toni's pure reaction of shame and lingers on it for a while in silence. The aim of silence is not only to shift the focus to the facial expression, but also to bring more attention to tiny "inappropriate" sounds for the situation. In this example is a slight camera zoom on Toni's face.
Another example is when Toni realizes that he might have accidentally killed Hildegard, where everyone is silent, and he turns around to look in horror at her casket. As he turns, his shoe emits a faint squeak, which, if it weren't for the stillness, would fall on deaf ears.
These silent moments also allow distinguishing bits from each other and creating rhythm changes that keep the action varied and engaging.

7.3.5. Loud fly on the wall

A unique feature of Kloster that I still haven't touched on is the interaction between the characters and the reporter, who is behind the camera for most of the film. All of the reporter's off-screen dialogue was re-recorded in post. This gave us the unique option of tinkering with the timing of the jokes semi-independently from the dialogue on screen. I use the term semi-independently because the editing still gave timing constraints based on the time of the response. It gave me the chance to further influence the timing of the joke through editing and by anticipating or delaying the delivery.
An objective of the re-recording was also to give a voice to the camera. The reporter gives direction to the camera as to where to shoot at some point. When the TV crew runs with Scholastika, we hear panting breath accompanying the camera movement, paired with the sound of the camera handling.
An added point of humor comes from the interruption of other characters' interactions and actions with questions and intrusions. This is in direct opposition to the tenets of documentary filmmaking, where the filmmaker is supposed to be a fly on the wall, an impartial observer who's not supposed to interfere with the story. This further adds to the insistence of the TV crew, playing on the stereotype of local TV networks.

7.3.6. Outcome

The sound design for this comedy remains, to some degree, dependent on sound design semantics. All of the jokes circling the TV crew are dependent on the audience's understanding of documentary filmmaking tropes and practices, thus relying on the semantics of the genre.
It is still possible to add some elements of context-independent sound design practices, but it is entirely mediated by active understanding of the situation, rather than psychoacoustic reactions to specific sound

stimuli. Perceptual incongruity, by means of exaggeration, is reliant on some form of perceivable mitigation by the sound designer. The strongest clue regarding the possibility of context-independence is the use of silence itself, as it has a direct perceptual shift which can be a source of humour.

8. Conclusion

This study set out to analyze to what degree sound design can be understood as context-dependent and context-independent. How the common practices in sound for horror and comedy are tied to the semantics of the genre and their psychoacoustic effect. These questions are not framed as a dichotomy but rather as coexisting modes of analysis that can influence one another. The questions are asked to gain an understanding of how far these two modes can be separated for practical purposes, namely the application of these practices as context-independent as possible and devoid of genre semantics.

This was done by applying these practices to realistic contexts, whose sound design objective is to present a reality that feels unmitigated, as opposed to the sound design practices in horror and comedy, which, as we've seen by analysing examples, feel heavily mitigated. The theme of realism is a cause for many misunderstandings. Sound design is often detached from any notion of realism in its practice, as it often opts for hyper-realistic representation to favor legibility and narrative.

From both the theory and the practice, we can conclude that horror is more suited, between the two genres, to have sound design practices that can be rendered context-independent.

This can be backed by research regarding fear, which points to deeper psychological mechanisms and reflexes based on evolution and survival, which can bypass cognitive understanding of meaning relating to sounds. This means that fear is more easily induced by specific sound stimuli based on frequency targeting and loudness, regardless of the context.

The analysis of jump scare scenes made it possible to reduce the semantic phenomena into discrete acoustic phenomena. This is partially due to the psychoacoustic reaction to the acoustic startle, which has been proven to elicit a reaction by virtue of factors related to loudness difference and timing, rather than semantic significance of the startle.

The stinger, arguably the single most semantically loaded sound in the genre, recognisable as a marker of shock, turns out to owe its effectiveness to pitch, transient, and voicing rather than to its status as a cultural sign. Therefore, its effect can be accomplished by applying its discrete components as diegetic sounds.

Dislocating sounds for the purpose of creating tension and anticipation is also possible, so long as it follows the logic of diegesis and the visual counterpart.

Based on universal sensitivity and perception to certain frequencies, certain manipulation to the diegetic sound space can be applied to provoke discomfort and instability in the viewer. Such is the case with background sounds, which can be enhanced and manipulated to generate different degrees of sonic intensity for a scene without resorting to semantic sounds such as drones to achieve the same effect. Infrasound emphasis can be introduced as a perceived element, rather than a heard one, to generate a bodily effect.

Crossmodal correspondence can furthermore be the basis for characterisation of character-specific foleys to introduce sonic traits of size and weight with frequency emphasis, generating slight incongruity that can aid in generating the intended affective state.

An important thing to note is that these practices, as they can be perceived as mitigated, can still be read as semantic sounds relating to the genre. The auditor might still perceive the cultural meaning of the sound based on the felt affective state, thus further confirming that although sound design for horror can distance the psychoacoustic effect from the semantic, the two things are inseparable. The mitigation itself might be perceived as the new semantic.

Comedy, on the other hand, is more resistant to the distancing between semantic and psychoacoustic. This can be attributed to the cognitive mechanisms of laughter, which are more heavily dependent on the recognition of incongruities. Laughter is theorized to be a reward mechanism for spotting such incongruities, which requires cognitive understanding of situations and meanings.

As opposed to horror, where the sound manipulations are based on frequency overload and loudness changes, the manipulations in comedy are more reliant on the logic of the film itself. By process of extending, exaggerating, substituting, and adding sound elements, the mitigation becomes much more apparent. The generation of incongruities in comedy is much more heavy-handed precisely because it necessitates the audience's understanding of said incongruity. Though some research exists on the possibility of purely context-independent incongruity, it is not something widely observed in comedy films, where perceptual incongruities are intrinsically mitigated.

Many examples also show that many comedy films heavily rely on the semantics of sound genre to generate laughter, for example in parodies. The understanding of these semantic elements is pivotal to the subsequent subversion of expectation.

The only possible functional context-independent sound design application in comedy seems to be a subtractive approach, where imposing silence on a specific shot can shift the audience's focus to the visual comedy on the screen.

Outlook

This research focused on the attempt of applying context-indeopendent sound elements to realistic film, which, although not being necessarely a genre, has a clear aesthetic identity relating to the feeling of unmitigated reality. Further research may show to what degree these findings could also be applied to other contexts. Perhaps by applying the practices as context-independent to other semantic genres.

Future research may also focus on the semantic analysis of on-screen realism. As we've established, notions of realism are very much at odds with each other. Realistic sound design is built on principles of legibility and verosimilitude. More in-depth analysis could be granted to semantic elements that might evoke verisimilitude and unmitigated perceptions of reality, when at odds with semantic expectations related to genres, i.e., what is realistic in relationship to the real-world experience, and what is realistic in relationship with the audience's expectation of film language.

This idea could very much be pertinent in comedy, as understanding the subversion of overall film sound semantics could further explain the role of sound design in the genre.

Sources

Acosta, E. A. M., & Sears, D. R. W. (2026). (Re)Considering the jump scare in four elements. Frontiers in Psychology, 16. https://doi.org/10.3389/fpsyg.2025.1569394

Altman, R. (1984). A Semantic/Syntactic Approach to Film Genre. *Cinema Journal*, *23*(3), 6. https://doi.org/10.2307/1225093

Altman, R. (1992). Sound Theory, Sound Practice. Psychology Press. (n.d.).

Arnal, L. H., Flinker, A., Kleinschmidt, A., Giraud, A.-L., & Poeppel, D. (2015). Human Screams Occupy a Privileged Niche in the Communication Soundscape. Current Biology, 25(15), 2051–2056. https://doi.org/10.1016/j.cub.2015.06.043

A quote by Robert Bloch. (2026). Goodreads. https://www.goodreads.com/quotes/10364500-comedy-and-horror-are-opposite-sides-of-the-same-coin

Bachorowski, J.-A., & Owren, M. J. (2002). The acoustic features of human laughter. The Journal of the Acoustical Society of America, 111(5), 2480. *https://doi.org/10.1121/1.4778613*

Barbosa, Á., & Dizon, K. (2020). The film sound analysis framework: A conceptual tool to interpret the cinematic experience. *Journal of Science and Technology of the Arts, 12*(2), 81–96. https://doi.org/10.34632/jsta.2020.8528

Blog, W. (n.d.). Infrasound: The Terrifying Tonality | WDAV: Of Note. Retrieved August 15, 2026, from *https://blogs.wdav.org/2019/10/infrasound-the-terrifying-tonality/*

Champagne, R. A., Todorov, T., & Howard, R. (1978). The Poetics of Prose. World Literature Today, 52(1), 188. *https://doi.org/10.2307/40134197*

FMT. (2023, May 27). *Cluster Voicing & Open Cluster Voicing | Film Music Theory*. Film Music Theory. https://filmmusictheory.com/article/cluster-voicing-open-cluster-voicing/?srsltid=AfmBOorXH69v8LKh5pRIEiIV9Mapb7CrQkMMLxpZ2hnEi3M7ahWi2Kda

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Green, A. (2015, October 16). Why Do We Laugh When We're Scared? Mental Floss. *https://www.mentalfloss.com/science/psychology/why-do-we-laugh-when-were-scared*

Hurley, M. M., Dennett, D. C., & Adams, R. B. (2011). Inside Jokes. MIT Press.

Huston, A. C., Wright, J. C., Fitch, M., Wroblewski, R., & Piemyat, S. (1997). Effects of documentary and fictional television formats on children's acquisition of schemata for unfamiliar occupations. *Journal of Applied Developmental Psychology*, *18*(4), 563–586. https://doi.org/10.1016/S0193-3973(97)90029-0

Hutcheon, L. (1991). *A theory of parody*. Routledge.

Hye-Knudsen, M., Kjeldgaard-Christiansen, J., Boutwell, B. B., & Clasen, M. (2024). First They Scream, Then They Laugh: The Cognitive Intersections of Humor and Fear. *Evolutionary Psychology*, *22*(2), 14747049241258355. https://doi.org/10.1177/14747049241258355

Kaplan, S. (2013). The Hidden Tools of Comedy.

Langkjær, B. (2010). Making fictions sound real - On film sound, perceptual realism and genre. MedieKultur: Journal of Media and Communication Research, 26(48), 13. https://doi.org/10.7146/mediekultur.v26i48.2115

Lastra, J. (2000). *Sound Technology and the American Cinema: Perception, Representation, Modernity*. Columbia University Press. https://doi.org/10.7312/last11516

Leeder, M. (2018). Horror film: a critical introduction. Bloomsbury Academic, An Imprint Of Bloomsbury Publishing Inc.

Lombana Bermudez, A. A. (2008). The "New" Sounds of the Slap-of-the-stick.

Manuel, M. (2013). Horror-Comedy: The Chaotic Spectrum and Cinematic Synthesis. Oglethorpe Journal of Undergraduate Research, 2(2). *https://digitalcommons.kennesaw.edu/cgi/viewcontent.cgi?article=1022&context=ojur*

Mast, G. (1979). The comic mind : comedy and the movies. University Of Chicago Press.

neurominded. (2023). The Spooky Science of Horror Movies: Why Do We Enjoy Being Scared? In Inspire The Mind. *https://www.inspirethemind.org/post/the-spooky-science-of-horror-movies-why-do-we-enjoy-being-scared*

Nicole. (2025, December 19). The Quiet Frame: How Silence Shapes Storytelling in Film. BeverlyBoy Productions. *https://beverlyboy.com/film-technology/the-quiet-frame-how-silence-shapes-storytelling-in-film/*

Nummenmaa, L. (2021). Psychology and neurobiology of horror movies. Psychology and Neurobiology of Horror Movies. *https://doi.org/10.31234/osf.io/b8tgs*

Ouellette, J. (2025, April 29). Monty Python and the Holy Grail turns 50. Ars Technica. *https://arstechnica.com/culture/2025/04/monty-python-and-the-holy-grail-turns-50/?utm_source=chatgpt.com*

Prince, S. (1996). True Lies: Perceptual Realism, Digital Images, and Film Theory. Film Quarterly, 49(3), 27–37. https://doi.org/10.2307/1213468

Redfern, N. (2020). Quantitative Analysis of Sound in a Short Horror Film, Humanities Bulletin. ResearchGate, 3(2), 246–257. https://www.researchgate.net/publication/346963997_Quantitative_Analysis_of_Sound_in_a_Short_Horror_Film_Humanities_Bulletin

Sbravatti, V. (2019). Acoustic Startles in Horror Films. *Projections*, *13*(1), 45–66. https://doi.org/10.3167/proj.2019.130104

Spence, C., & Parise, C. V. (2012). The Cognitive Neuroscience of Crossmodal Correspondences. I-Perception, 3(7), 410–412. https://doi.org/10.1068/i0540ic

Stalnaker, R. C. (1999). Assertion. In *Context and content: Essays on intentionality in speech and thought* (pp. 78–95). Oxford University Press. (Original work published 1978)

What is Fidelity in Film Sound? - Team Beverly Boy. (2021). In Beverly Boy Productions. https://beverlyboy.com/filmmaking/what-is-fidelity-in-film-sound/

Whittington, W. (2007). Sound Design and Science Fiction. University of Texas Press.

Filmography

Brooks, J. (Director). (2014). *Behold the noose* [Film]. Thrillvania Productions.

Brooks, M. (Director). (1974a). *Blazing saddles* [Film]. Crossbow Productions; Warner Bros.

Brooks, M. (Director). (1974b). *Young Frankenstein* [Film]. Gruskoff/Venture Films; Crossbow Productions; Jouer Limited.

Brooks, M. (Director). (1987). *Spaceballs* [Film]. Brooksfilms; Metro-Goldwyn-Mayer.

Brooks, M. (Director). (1993). *Robin Hood: Men in tights* [Film]. Brooksfilms; Gaumont.

Carpenter, J. (Director). (1978). *Halloween* [Film]. Compass International Pictures; Falcon International Productions.

Carpenter, J. (Director). (1982). *The thing* [Film]. The Turman-Foster Company; Universal Pictures.

Chaplin, C. (Director). (1925). *The gold rush* [Film]. Charles Chaplin Productions.

Craven, W. (Director). (1996). *Scream* [Film]. Woods Entertainment; Dimension Films.

Cronenberg, D. (Director). (1986). *The fly* [Film]. Brooksfilms; SLM Production Group.

De Palma, B. (Director). (1976). *Carrie* [Film]. Red Bank Films.

Derrickson, S. (Director). (2012). *Sinister* [Film]. Blumhouse Productions; IM Global; Automatik Entertainment.

Douglas, G. (Director). (1954). *Them!* [Film]. Warner Bros.

Eggers, R. (Director). (2024). *Nosferatu* [Film]. Focus Features; Studio 8; Maiden Voyage Pictures.

Friedkin, W. (Director). (1973). *The exorcist* [Film]. Hoya Productions; Warner Bros.

Gilliam, T., & Jones, T. (Directors). (1975). *Monty Python and the Holy Grail* [Film]. Python (Monty) Pictures; Michael White Productions.

Hitchcock, A. (Director). (1954). *Rear window* [Film]. Patron Inc.; Paramount Pictures.

Hitchcock, A. (Director). (1960). *Psycho* [Film]. Shamley Productions.

Hooper, T. (Director). (1974). *The Texas chain saw massacre* [Film]. Vortex.

Käsemodel, S. (Director). (2026). *Kloster* [Film]. ZHdK.

Leone, S. (Director). (1966). *Il buono, il brutto, il cattivo* [The good, the bad and the ugly] [Film]. Produzioni Europee Associate; Arturo González Producciones Cinematográficas; Constantin Film.

Lucas, G. (Director). (1977). *Star Wars* [Film]. Lucasfilm; Twentieth Century-Fox.

Mitchell, D. R. (Director). (2014). *It follows* [Film]. Northern Lights Films; Animal Kingdom; Two Flints.

Murnau, F. W. (Director). (1922). *Nosferatu: Eine Symphonie des Grauens* [Nosferatu: A symphony of horror] [Film]. Prana-Film.

Myrick, D., & Sánchez, E. (Directors). (1999). *The Blair Witch Project* [Film]. Haxan Films.

Riley, B. (Director). (2018). *Sorry to bother you* [Film]. Significant Productions; MACRO; Cinereach.

Romero, G. A. (Director). (1968). *Night of the living dead* [Film]. Image Ten.

Roth, E. (Director). (2005). *Hostel* [Film]. Next Entertainment; Raw Nerve.

Siegel, D. (Director). (1956). *Invasion of the body snatchers* [Film]. Walter Wanger Productions.

Straub, M. (Director). (2026). *Talion* [Film]. ZHdK.

Tati, J. (Director). (1967). *Playtime* [Film]. Specta Films.

Wan, J. (Director). (2004). *Saw* [Film]. Twisted Pictures; Evolution Entertainment.

Wan, J. (Director). (2010). *Insidious* [Film]. Blumhouse Productions; Alliance Films; IM Global.

Wiene, R. (Director). (1920). *Das Cabinet des Dr. Caligari* [The cabinet of Dr. Caligari] [Film]. Decla-Bioscop.

Yeaworth, I. S., Jr. (Director). (1958). *The blob* [Film]. Fairview Productions; Tonylyn Productions.

Television series

Bays, C., & Thomas, C. (Executive Producers). (2005–2014). *How I met your mother* [TV series]. Bays & Thomas Productions; 20th Century Fox Television.

Daniels, G., Gervais, R., & Merchant, S. (Executive Producers). (2005–2013). *The office* [TV series]. Deedle-Dee Productions; Reveille Productions; NBC Universal Television Studio.

David, L., & Seinfeld, J. (Executive Producers). (1989–1998). *Seinfeld* [TV series]. Castle Rock Entertainment; West-Shapiro Productions.

Henson, J. (Executive Producer). (1976–1981). *The Muppet Show* [TV series]. ITC Entertainment; Henson Associates.

Video Links

Brooks, M. (Director). (1974b). *Young Frankenstein,* "Frau Blücher Scene"

http://www.youtube.com/watch?v=vvVmj0s56GY

Derrickson, S. (Director). (2012). *Sinister, "*Lawnmower scene"

https://www.youtube.com/watch?v=ZxXiIOdrfh0

Eggers, R. (Director). (2024). *Nosferatu, "*Opening Scene"

https://www.youtube.com/watch?v=ey5pjdk9Z3g

Gilliam, T., & Jones, T. (Directors). (1975). *Monty Python and the Holy Grail,* "Opening Scene"

https://www.youtube.com/watch?v=6NbCNA0mAtM

Mitchell, D. R. (Director). (2014). *It follows* "Tall Man Scene"

https://www.youtube.com/watch?v=nnrT_zyumAM

Tati, J. (Director). (1967). *Playtime,* "Glass Door"

https://www.youtube.com/shorts/mrSI4tHyIyU

Wan, J. (Director). (2010). *Insidious,* "Red Demon Scene"

https://www.youtube.com/watch?v=CZN6nbd4pB4

Figures

Figure 1 *"The Red Demon" Scene From Insidious. Note.* Screenshot from *Insidious* [Film], by J. Wan (Director), 2010, Blumhouse Productions.

Figure 2 *Opening Scene From Nosferatu. Note.* Screenshot from *Nosferatu* [Film], by R. Eggers (Director), 2024, Focus Features.

Figure 3 *"Tall Man" Scene From It Follows. Note.* Screenshot from *It Follows* [Film], by D. R. Mitchell (Director), 2014, Northern Lights Films.

Figure 4 *Close Cluster Voicing. Note.* From "Cluster Voicing & Open Cluster Voicing," by FMT, 2023, Film Music Theory.

Figure 5 *Open Cluster Voicing. Note.* From "Cluster Voicing & Open Cluster Voicing," by FMT, 2023, Film Music Theory.

Figure 6 *Amplitude Envelope of the Soundtrack of Behold the Noose. Note.* From "Quantitative Analysis of Sound in a Short Horror Film," by N. Redfern, 2020, *Humanities Bulletin, 3*(2), p. 249. CC BY.

Figure 7 *Log Amplitude Spectrogram of the Soundtrack of Behold the Noose. Note.* From "Quantitative Analysis of Sound in a Short Horror Film," by N. Redfern, 2020, *Humanities Bulletin, 3*(2), p. 249. CC BY.

Figure 8 *Glass Door Scene From Playtime. Note.* Screenshot from *Playtime* [Film], by J. Tati (Director), 1967, Specta Films.

Figure 9 *King Arthur and His Squire in Monty Python and the Holy Grail. Note.* Screenshot from *Monty Python and the Holy Grail* [Film], by T. Gilliam & T. Jones (Directors), 1975, Python (Monty) Pictures.

Figure 10 *Noé Getting Shot in Talion. Note.* Screenshot from *Talion* [Film], by M. Straub (Director), 2026, ZHdK.

Figure 11 *Hildegard Smoking in Kloster. Note.* Screenshot from *Kloster* [Film], by S. Käsemodel (Director), 2026, ZHdK.