INTERNATIONAL JOURNAL OF FILM AND MEDIA ARTS (2025) Vol. 10, Nº. 2 pp. 176-199

ijfma.ulusofona.pt
doi: 10.24140/ijfma.v10i2.9708

EMBODIED METAPHORS IN FILM SOUND: THE CASE OF AURAL DYNAMICS

TARUN MADUPU
LUSÓFONA UNIVERSITY (PORTUGAL)

MAARTEN COËGNARTS
UNIVERSITY OF ANTWERP/ LUCA SCHOOL OF ARTS (BELGIUM)
THE UNIVERSITY OF THE FREE STATE (SOUTH AFRICA)

Tarun Madupu is a sound designer, composer and filmmaker from India. After graduating from Swarnabhoomi Academy of Music with a focus on audio engineering, he went on to get his Master's degree from Kino Eyes - The European Movie Masters. He is currently preparing a PhD in sound design at Lusófona University. https://orcid.org/0009-0004-8513-3466

Maarten Coëgnarts is Maarten Coëgnarts is assistant professor in film studies at the University of Antwerp, researcher in the arts at LUCA School of Arts and research fellow at the University of the Free State. He is co-author of *Embodied Cognition and Cinema* (Leuven University Press, 2015), author of Film as Embodied Art: Bodily Meaning in the Cinema of Stanley Kubrick (Academic Studies Press, 2019) and co-editor of Projections: The Journal for Movies and Mind. https://orcid.org/0000-0003-2813-4847

Corresponding author
Tarun Madupu
p6970@ulusofona.pt
Universidade Lusófona
Campo Grande 376
1749-024 Lisboa
Portugal

Schedule for publication
Paper Submitted: 30th May 2024
Accepted for Publication: 7th February 2025
Published online: 29th December 2025

Abstract

This article explores the embodied meaning-making capacities of loudness or dynamics in film sound. Drawing on the research program of embodied cognition, we demonstrate how contemporary film sound practitioners use variations in loudness, silence and moments of dynamic contrast between the two, as a tool of metaphorical aural storytelling. We present and illustrate a classification of four strategies for manipulating dynamic range: (1) contrasting dynamic range, (2) relative loudness and silence, (3) spectral dynamics (frequency information and harmonic information) and (4) sound field. It is through these categories that we argue sound practitioners are able to communicate with listener-viewers because the meanings they flesh out in the sound design parallel the inherently embodied processes of human perception and cognition.

Keywords: Aural dynamics, embodied metaphor, loudness, sound design

1. Introduction

Film theory has traditionally modeled meaning in film upon the structuring principles of language (e.g., Carroll, 1980; Metz, 1974). This view can be considered disembodied because it seeks the locus of meaning in arbitrary relationships rather than in the experiential world of the human observer. Recent trends in contemporary cognitive science and neuroscience, however, have presented a fundamental challenge to this classical formal understanding of meaning by going beyond language, claiming instead that all forms of symbolic meaning-making, whether linguistic or artistic, are fundamentally grounded in action and pre-cognitive patterns of sensory-motor experience (e.g., Claxton, 2015; Gibbs, 2005; Lakoff & Johnson, 1999; Shapiro, 2019; Tversky, 2019; Wilson, 2002). One of the key concepts introduced to articulate the relationship between meaning and embodiment is conceptual metaphor (Lakoff & Johnson, 1980). As George Lakoff and Mark Johnson famously argued, metaphors are not simply decorative elements of language; rather, they are fundamental to human cognition and are deeply embodied, drawing on sensory-based experiences. They rely on concepts related to physical form, motion, and spatial orientation to structure and make sense of abstract domains such as time and emotion. Importantly, scholars working at the intersection of music and cognition have noted that this also involves translating auditory properties into metaphorical spatial concepts such as pitch height and musical motion (e.g., Brower, 2000; Cox, 2017, 2018; Johnson & Larson, 2003; Larson, 1997, 2012; McKee, 2007; Saslaw, 1996; Zbikowski, 2008). While scholars have discussed the role of embodied visual metaphors in cinematic storytelling (e.g., Coëgnarts & Kravanja, 2012, 2015; Ortiz, 2011, 2023), the exploration of embodied sound metaphors in film is still relatively modest (e.g., Coëgnarts, 2023; Fahlenbrach, 2008), although recent contributions from the field of artistic and practice-based research have begun to shed light on this area (Chattah, 2015, 2024; Melvin & Bridges, 2021; Ward, 2015). Chattah's use of conceptual metaphors, for instance, illustrates how connections between the source domain of film music and the target domain of narrative can uncover the structure behind cross-modal metaphorical understanding. He shows how musical features like harmony, rhythm, and dynamics can metaphorically represent abstract narrative elements such as conflict, resolution, tension, and pacing. Furthermore, Annabel Cohen's Congruence-Association Model with Working Narrative (henceforth, CAM-WN), developed in the context of film music (Cohen, 2015), offers a cognitive framework for understanding how these sonic metaphors might be processed. According to CAM-WN, audiences construct a "working narrative"—a dynamic, moment-to-moment experience of the film—by continuously integrating input from six channels: text, speech, visuals, music, sound effects, and kinesthetic cues, in conjunction with their long-term memory, prior knowledge, and expectations. Within this model, congruence refers to the structural coherence perceived across sensory modalities, while association involves how viewers draw on personal memories, cultural norms, and existing knowledge to interpret and assign meaning to what they experience.

While Chattah's analysis focuses exclusively on film music and Cohen's model treats sound effects and music as separate informational channels, this paper takes a broader approach by examining the impact of the film's complete soundtrack. This integrated perspective is grounded in two key considerations: First, film score and sound design are often deeply interwoven, working together to create a unified auditory experience for the viewer. Second, from a practical production standpoint, both elements ultimately converge in the hands of the re-recording mixer, who shapes the final soundscape by adjusting the levels and dynamics of all sonic components—deciding what is heard, when, and at what intensity.

Drawing on the first author's practical experience as a sound designer and re-recording mixer, this article aims to contribute

to the debate by providing a metaphorical study of the way contemporary film sound practitioners use aural dynamic range, that is, the application of loudness, silence and moments of dynamic contrast between the two, as a tool of embodied narrative storytelling (Madupu, 2020). In the first part, we provide a theoretical understanding of the concept of dynamics. This investigation will prepare us for the second part in which we explore how film sound practitioners manipulate the variations in dynamic range for various intended effects. Correlations based upon both source and target domains will demonstrate how sound acts bodily and metaphorically, referencing the image or the narrative into the aural medium. To this aim, we will classify the prominent use of dynamics into four categories: (1) contrasting dynamic range, (2) relative loudness and silence, (3) spectral dynamics (frequency information and harmonic information) and (4) sound field. These categories often work in tandem, overlapping and reinforcing one another to produce certain effects. It is through these categories that it will be shown that sound designers are able to communicate with listener-viewers because the meanings they flesh out in the sound design parallel inherently embodied processes of human perception and cognition. It is precisely this embodied parallel that makes it possible for sound designers to communicate with listener-viewers in a very engaging yet unconscious and bodily way.

2. A definition of aural dynamics

Due to the polysemous nature of the word "dynamics", there exists scope for miscommunication with the application of the concept. Terms such as "dynamics", "intensity", "amplitude" and "volume" have often been used interchangeably (Behar, 1984). In the context of the study of acoustic intensity, a definition can be borrowed from the realm of musicology where the dynamics of a piece commonly refer to the variation in loudness between notes or phrases in a piece of music. Brian Moore (2005) provides a more formal definition, describing loudness as "that attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud" (p. 409). Herbert Zettl (2017, p. 336) considers it to be one of five key attributes or elements of sound that one can hear when striking a single piano key or blowing into a trumpet, the others being pitch, timbre, duration, and attack/decay. If loudness defines the apparent strength of a tone as we perceive it, the dynamics are "the variations of perceived strength" (p. 338). In musical notation, it is common to use formal symbols to indicate the loudness: the symbols f (forte) and fff (forte fortissimo) indicate the range between loud and extremely loud whereas the symbols p (piano) and ppp (piano pianissimo) indicate the range between soft and extremely soft. When these variations in perceived intensity are represented over time, they form the sound envelope (Zettl, 2017, p. 339). The envelope encompasses all of the characteristic variations in amplitude over time and consists of four properties: attack, initial decay, sustain, and release (or final decay). Attack (A) refers to the time it takes for a sound to reach a certain loudness level. The initial decay (D) refers to the dip in volume after the attack. The sustain level (S) is reached when the sound maintains its steady volume. Release (R), or final decay, is the time it takes for the sound to fade until it is no longer audible. As the second part will further illustrate, the envelope is typically represented graphically using an amplitude envelope curve, which shows how the amplitude (loudness) of a sound changes over time. While dynamics certainly entails the presence of sound at various intensities, it also implies the lack of sound or sound at an inaudible intensity. In other words, it includes the absence of sound or silence. Through the techniques of sound design and mixing, the various sounds that populate the soundtrack of a film are determined to be at a decided loudness. Some for necessary and functional purposes (such as the dialogue being audible to the audiences) and others for crafting an aesthetic and consequently a meaningful experience to the audiences.

It is important to note that loudness operates across two domains: the physical vibration of sound waves (measured as sound pressure level), and the subjective perception shaped by our biological and neurological processing. Psychoacoustics, the study of sound perception, reveals that loudness is not solely determined by volume, but also by frequency content. For instance, orchestral music may be perceived as louder than speech at the same volume due to its richer spectral content. Research by Munson and Fletcher (1933) showed that human ears are most sensitive to mid-range frequencies, especially those of speech, and less so to very high or low frequencies at lower volumes. This has been formalized through equal loudness contours. While dynamic range in sound is typically understood as variations in perceived loudness (i.e., how quiet or loud a sound is perceived by the listener), we broaden this concept to include not just volume but also the dynamics within the spectral and spatial dimensions of sound. We refer to the dynamics built through the variations in frequency as "spectral dynamics," and the dynamics built through the manipulation of its spatial qualities as "sound field". Spectral information can express a variety of dynamic qualities, such as textural and timbral shifts, harmonic consonance and dissonance, as well as the emphasis or de-emphasis of specific frequency bands. Sound practitioners, through sound editing and design (which involves selecting, creating, and assembling sound effects) and techniques like equalization, manipulate the intensity of particular frequencies within an audio signal to shape its tonal characteristics. Similarly, musical and sound design elements communicate dynamics through harmonic consonance and dissonance, creating cycles of tension and release. Cohen's model of congruence and association offers insight into how such dynamics can serve as fertile ground for metaphor (Cohen, 2015). For instance, cross-modal structural similarities enable sounds with certain spectral qualities to be described as rough or coarse, while others may be perceived as soft or smooth. In contrast, harmonic consonance and dissonance are culturally shaped, influenced by the associations viewers have with these sound characteristics stored in their long-term memory.

This expanded framework is justified not only by the inherent relationship between spectral content and perceived loudness but also from a formal and functional perspective. Just as sound exists on a continuum of loudness, ranging from silence to the loudest possible output of a sound source, frequency and spatial content can also present dynamically. Frequency content can exist on a spectrum ranging from low to high frequencies, sparse to dense spectral information, and consonant to dissonant tonalities. Similarly, the width or depth of the sound field can exist on a range from narrow and focused to wide and immersive. The variations across these dimensions (loudness, spectral, and spatial) are deliberately exploited by sound practitioners to craft metaphors that resonate with embodied human experience. The inclusion of these additional axes of variations captures the full expressive potential of the aural dynamics present in contemporary cinematic sound.

3. The role of aural dynamics in shaping meaning in film

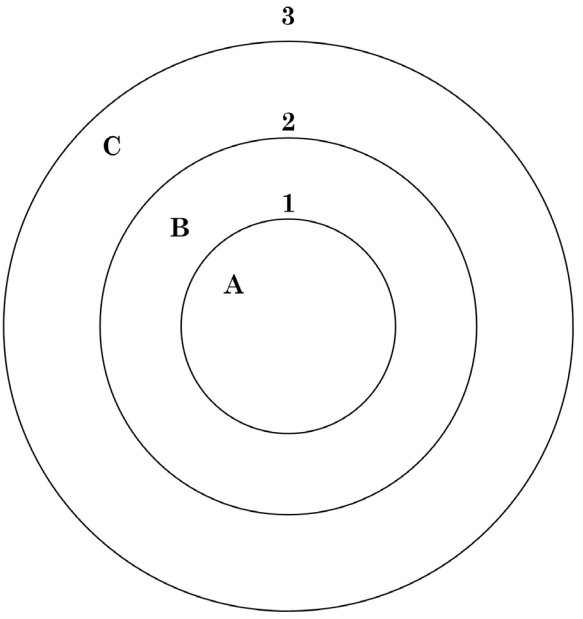
With the concept of aural dynamics now defined, we can explore its significant role in cinematic storytelling. To structure the discussion, we broadly distinguish between four distinctive ways of manipulating sonic dynamics in film, while recognizing they often work together for the purpose of affecting the emotional and cognitive state of the listener-viewer. They are as follows: (1) contrasting dynamic range, (2) relative loudness and silence, (3) spectral intensity (frequency information and harmonic information) and (4) surround width. Before addressing and exemplifying each category, it should be pointed out that these uses take place within what we may call the "audiovisual scene" of a film (Chion, 2019) or the "geography" of the film's "soundscape" (Stilwell, 2007, p. 187).

Film scholars usually conceptualize this sonic geography as consisting of two kinds of aural spaces: the diegetic aural space and the non-diegetic aural space. Following Coëgnarts (2023, p. 67) we may render the boundaries of these zones diagrammatically by means of applying a container logic (see Figure 1).

The diegetic aural space (container 2) is the soundscape of the diegetic world of the film. This kind of aural space is dictated by the point of audition or POA. POA can sometimes refer to situations where the audience shares the subjective aural space of a character's inner mind. Alternatively, sound designers also create aural spaces based on POA in a more objective spatial sense. For example, the spatial position the audience will "hear from" within the diegetic world. Diegetic aural spaces help bring the viewer into the world of the film and empathize with the characters in the film. This space is a representation of the reality of the characters in the film.

Depending on the part of the diegetic world shown on-screen ("framed"), we can distinguish between a "visualized on-screen zone" (space A or container 1) and an "acousmatic off-screen zone" (space B). The latter refers to parts of the story world suggested by the film but not visualized. If the sound source comes from the visual content of the frame (space A) and belongs to the reality depicted there, it is referred to as on-screen or visualized diegetic sound, or source-connected diegetic sound (Zettl, 2017, pp. 312–313). In contrast, if the source originates from within the acousmatic zone (space B), it is termed off-screen or acousmatized diegetic sound, or source-disconnected diegetic sound, meaning the sound is acousmatic relative to what is shown in the shot.

Within the interior of the diegetic aural space, we can also place internal sound, which, although part of the action, corresponds to the physical and/or mental "container" of a character. Following Chion, we can distinguish between



Legend

Space A = The visualised diegetic zone

Space B = The acousmatic diegetic zone

Space C = The acousmatic non-diegetic zone

Container 1 = The framed story world (A)

Container 2 = The story world (A+B)

Container 3 = The audiovisual scene (A+B+C)

Fig. 1 The container logic of film sound (after Coëgnarts, 2023, p. 67)

objective-internal sounds (e.g., breathing, moans, heartbeats, sounds of tinnitus) and subjective-internal or mental internal sounds (e.g., mental voices, memories). Source music, including on-the-air music (i.e., music performed on-screen or coming from a jukebox or radio), also belongs to this aural container.

In contrast, the non-diegetic space (space C) comprises all the non-diegetic elements of the film's soundtrack. This includes sound effects, score, and sound design used to manipulate the audience's aural experience. This manipulation typically does not affect the aural space of the film's characters. On a footnote, it is important to note that not all sounds of the film's soundtrack can be categorized into diegetic and non-diegetic. These sounds that can be considered "trans-diegetic" (Hunter, 2012) fall into neither category. It should also be noted that a large part of the sound designer's role is to blend the subjective aural space of the characters and the objective aural space of the scene. Further, these aural spaces are seldom static, and their categorization is in a constant state of flux. Through the control of dynamic range, sound designers move different elements of the aural space in and out of focus depending on the narrative needs of the film.

Having these conceptual and analytical distinctions in mind, we now turn to a discussion of each of the four distinctive uses of sonic dynamics. To this aim, we will analyze various cases drawn from the author's (©Madupu) own work as a sound designer. They include the documentary *Cypress* (Nosok, 2024) and the short films *Look What You Made Me Do* (Lesmes, 2020), *Play/Repeat* (Lesmes, 2023) and *Pedra Seca* (Mello, 2024).

To render the audio content of the various film examples visually, we make use of the software called Izotope Insight 2

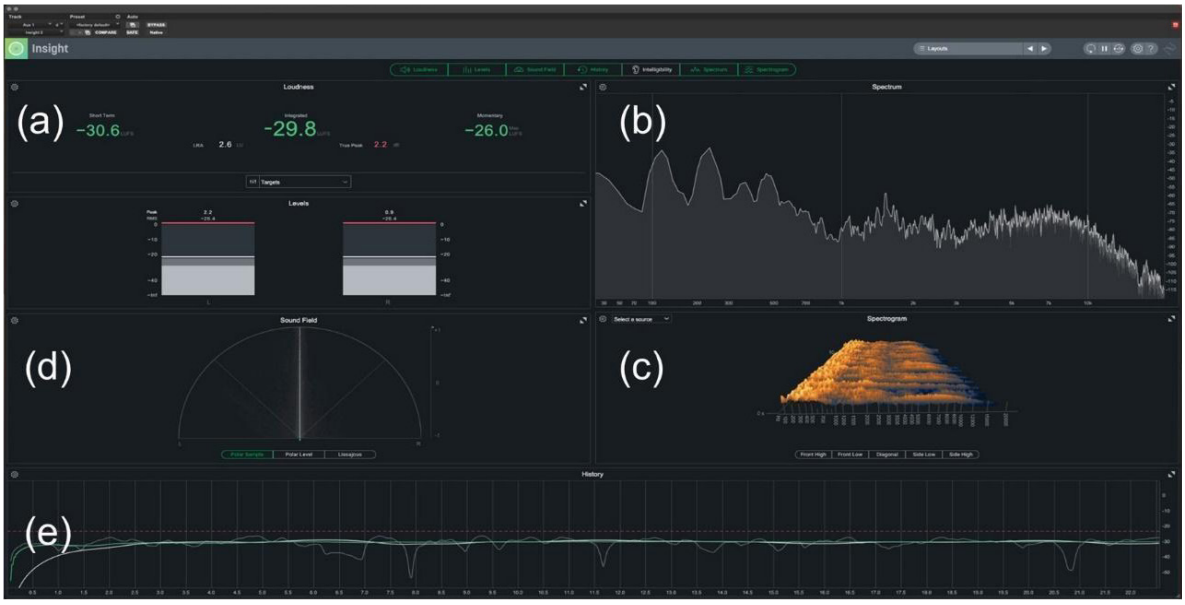


Fig. 2 Screenshot of the user interface of Izotope Insight 2.

(see Figure 2). Insight is an audio analysis, measurement and meter tool commonly used by audio professionals working in post-production, music and broadcast.

Each of the various modules are designed to measure different aspects of audio content. The loudness module, denoted as (a), is used to measure the overall integrated loudness of the programme over a period of time in the LUFS scale. LUFS (Loudness Units Relative to Full Scale) is a standardized unit for measuring perceived loudness of audio, aligning with human hearing sensitivity, providing a more accurate representation of how loud audio actually sounds to listeners. The spectrum analyser (b) and spectrogram (c) measure frequency information. The spectrogram provides a visual representation of the audio over a predetermined period of time and the spectrum analyser offers real-time analysis of frequency content in the audio. The sound field module, marked as (d), allows for the analysis of spatial characteristics stereo and surround and surround sound. Finally, the history module (e) tracks changes in loudness over a period of time. As such, the various modules in Insight are well suited to help visualize changes in auditory dynamics.

3.1 Contrasting dynamic range

Dynamic range refers to the difference between the loudest point and the quietest point of the audio. Sound designers and re-recording mixers use the wide dynamic range that modern sound systems are capable of to enforce narrative aspects of the film. For instance, large emotional scenes may feature music and sound effects mixed loudly and prominently. In contrast, a delicate and tender moment in the film may be mixed with comparable sensitivity to reflect the subtlety and intimacy of the moment.

While the soundtrack of a film is often mixed to be reasonably dynamic over the length of a film, a sudden and abrupt deviation in auditory levels allows sound designers to punctuate particularly significant narrative moments in a film. Take, for example, the sound dynamics that go together with the "Holdo Maneuver" in *Star Wars: Episode VIII: The Last Jedi* (Johnson, 2017). Vice Admiral Holdo chooses to remain aboard a Resistance ship allowing her allies to escape. Left with no chance of escape for herself, she reverses her ship's course to confront the opposing fleet before engaging lightspeed, sacrificing herself as her ship plows through the enemy fleet (Pearson, 2018). As the section of the film plays out, we see the two scenes building to a climax intercut with one another. This conceptual climax is supported by the non-diegetic space through the film´s score. As John Williams' track "Holdo's resolve" builds to a crescendo, the sequence concludes with the culmination of the duel between Rey and Kylo Ren, marked by the shattering of a lightsaber, and Holdo's dramatic crash into the enemy fleet. The crescendo of the brass and string sections allows for the score to reach the loudest point and dominate the soundtrack before suddenly breaking into a brief and sudden moment of complete silence to reveal the destructive results of Holdo's actions. This sudden shift from loudness to near silence jolts the viewer out of the one aural space to another one. After a few seconds of silence, the audience receives a surrogate resolution with the re-emergence of sound with an explosion of the ship over a wide shot. With this the audience are brought back into the diegetic aural space of the characters. The director of the film, Rian Johnson utilizes this sudden shift in dynamic range to highlight the scale of the event, both in terms of the plot and in terms of supporting the image (Pearson, 2018).

The scene marks a crucial narrative moment, revealing the sacrifice of one of the major characters and serves as the culmination of an arc in the film. The grand scope of this moment is mirrored in the sound design through a wide dynamic range. A crescendo in the score transitions abruptly to complete silence, followed by an explosion. This dramatic shift in dynamics metaphorically reflects the narrative's weight

and impact. The source domain of aural dynamics (the crescendo-to-silence pattern) maps onto the target domain of emotional climax and narrative weight. While such punctuative uses of contrasting dynamic range to mark narratively dramatic or 'big' moments are common, dynamic shaping can also take on a more prominent role—especially in scenes where the visuals or plot alone do not provide obvious dramatic cues. In these cases, it is the aural dynamics that lead, creating a space for audiences to engage in metaphorical interpretation through sound.

To provide another example of our own, let us turn to *Cypress* (2020). The film follows Valerii, an elderly self-taught artist who lives in a remote village in the east of Ukraine. His admiration for Greek mythology and art is reflected in his sculptures and paintings. Yet a specific painting remains incomplete due to his lack of firsthand experience of seeing a Cypress tree. Upon a short video of him expressing his lifelong desire to visit Greece going viral, a crowd-funded trip allows him to fulfill his wish. The film explores themes of solitude, national identity and pilgrimage.

Figure 3 depicts the early moments of the film where we are presented with a montage of lush vistas from Greece (denoted as section (a)) before cutting to the title card framed against a backdrop of the pine forest near Valerii's village in Ukraine (denoted as section (b)). Here, the director and the sound designer discussed the need to emphasize not just the visual contrast between the two locations but also to feel the change sonically and emphasize the thematic elements in sound. As such, the sound over the shots of Greece are designed to reflect the idyllic version in Valerii's imagination, lush with sounds of nature - bird calls, crickets and surreal bells balanced on the edge of diegesis. In collaboration with the composer, Margaryta Kulichova, the choral score was treated with a large reverb to add a sense

Fig. 3 Mapping of the integrated loudness of the soundtrack over time, with loudness measured in LUFS on the y-axis and time on the x-axis.

of space and grandeur. Towards the end of the selected passage, the score dies out and the sounds of nature crescendo towards an abrupt drop over the title frame set against a backdrop of a rural Ukrainian pine forest in the cold of winter. The following shots feature a softly mixed cold wind that starkly contrasts the vibrant soundscapes of the preceding shots. This was done with the intention to metaphor ideas of solitude, and isolation and to contrast the harsh reality of Valerii's rural life against the idealized version of Greece that he portrays in his art.

While the aforementioned ADSR envelope is typically used to examine characteristics of individual sounds, we may extrapolate similarities to analyze the overall dynamics of the film segment. When reframed as an envelope, we see a soft attack, a lengthy sustain and relatively fast release. The embodied nomenclature and characteristics of the envelope (softness, length, speed) make it clear that the dynamics were designed to support the narrative intention of sequence. The introduction of the music cue results in a gentle raise of level before the sound design saturates the audiovisual scene, leading to a sense of sustained fullness. The quick release of acoustic intensity provides a contrasting effect in the proceeding moments. Using the terminology of Lakoff and Johnson we may further formalize this through the A is B template whereby the (abstract) narrative beat (A) is further understood metaphorically through the contrast in dynamic range (B). The contrast of idealized desire of our protagonist and the reality of his rural life is manifested in the contrast of loudness.

Yet another instance of the use of contrasting dynamic range can be seen in the film *Look What You Made Me Do* (Lesmes, 2020). The film was sound designed by the first author and mixed by re-recording mixer Israel Bañuelos. The film is a character study of Ingmar, an egocentric and abusive narcissist prone to violent outbursts of temper. Unable to tolerate romantic rejection, he seeks out and meets a new partner, once again perpetuating his cycle of emotional abuse. The film depicts him as an eccentric and explosive individual. Consequently, viewers witness the stark contrast between Ingmar's character as a calm and charming individual one moment before exploding into a fit of rage the next. As such this eccentricity is reflected in the mix and sound design of the film.

Analyzing the loudness dynamics of the first four minutes of the film (see Figure 4), we can see how abrupt shifts in loudness have been used to create moments of dynamic contrast. The film opens with Ingmar's desperate pleas directed at a locked door in the hallway of his girlfriend's apartment complex. The peaks in loudness denoted as (a), (b) and (c) are moments when Ingmar lashes out in rage. While the on-screen actions provide the diegetic motivation for these crescendos, the sound design, through the foley and the mix, amplify the intensity of these moments. Conversely, the ambience of the hallway itself is designed to accentuate its inherent silence, widening the disparity within the dynamic range. The scene ends with a smash cut to the title card which is scored with Free Jazz (d), a subgenre of Jazz characterized with unconventional structures, and atonality, and irregular dynamics. The choice of the music genre aligns with the director's vision, aiming to mirror the essence of the protagonist's character. Once again utilizing a sharp cut to transition out of the title card, we see Ingmar sitting alone in his car. Here we see him adjust his clothing and meticulously pluck his eyebrows before intensely gazing into the rearview mirror. The sound design in this section starts subtly with the soundscape reflecting diegetic elements of the scene such as the metallic patter of the rain against the roof of the car and the distant hum of traffic. However, as the scene progresses the sound design gradually morphs, becoming more surreal and unnatural. While the subtle patter of rain against the car gradually increases in level, the metal clinking sound of a boiling pot is introduced underneath. In addition sounds of stormy waves and thunder are added,

at a low level at first before slowly building to a crescendo (e). An abrupt cut shows Ingmar submerged in a swimming pool. An unnatural silence (f) permeates this moment as Ingmar lies submerged and motionless before finally returning to the diegetic soundscape of the scene when he surfaces for air.

When contextualized as an ADSR curve, the sharp attacks, quick releases and the dynamic mix of crescendos and abrupt dips in loudness embody the essence of Ingmar's psyche, one that is unstable and volatile. Through the sound design and mix, the soundtrack of the film aims to metaphor the character traits of the film's protagonist.

Examining the two examples through the lens of Cohen's CAM-WN model offers deeper insight into the cognitive mechanisms involved. In both instances, the viewer's personal experiential associations (stored in long-term memory) play a crucial role in generating semantic meaning from the sound content. For example, cultural associations with free jazz evoke a sense of eccentricity, while folk choral singing is linked to nature and harmony. Simultaneously, the structural congruencies between the narrative and the sonic dynamic range facilitate metaphorical meaning-making through embodied patterns. In the case of *Cypress*, while both sections feature wide, open landscapes, the differing dynamics engage distinct sensory experiences. The abstract contrast

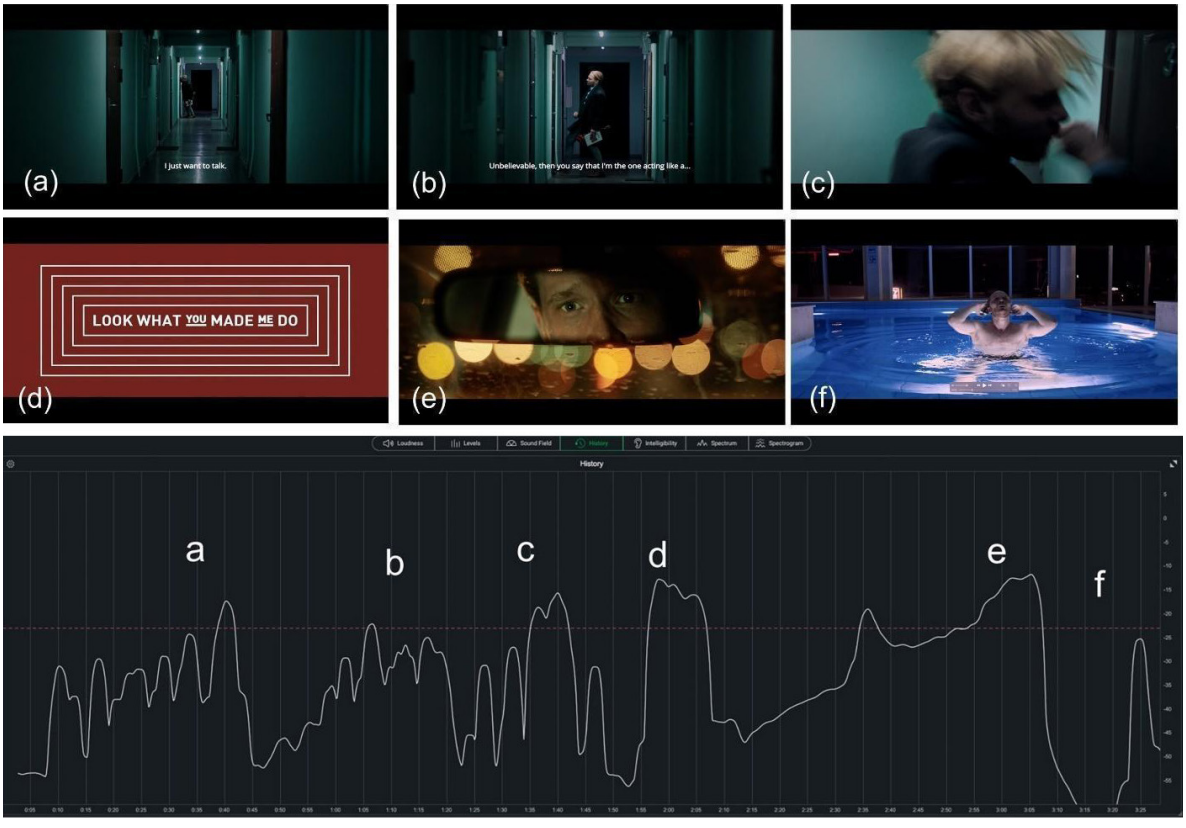


Fig. 4 Emotional intensity as emphasized by peaks in loudness.

between Valerii's idyllic dream and his rural reality is reflected in the sonic dynamics. The sequence over Greece (a), paired with a loud, reverberant choral score and rich natural sounds, evokes an embodied sensation of dreamlike expansiveness. In contrast, the shot over the Ukrainian pine forest (b) immerses the viewer in a more confined, isolated experience, mirroring the starkness of reality. Similarly, in *Look What You Made Me Do*, despite the relatively mundane visual framing of hallway corridors, parked cars, and closed interiors, the sound mix amplifies Ingmar's internal volatility by alternating between moments of intense sonic tension and sudden releases. This dynamic shift allows the audience to physically experience his psychological instability through embodied responses.

3.2 Relative loudness and silence

The amplitude of a particular sound in the context of the soundtrack of the film is one of its defining traits. The general implication of acoustic intensity being that the particular sound demands to be noticed and serves either a functional or an intellectual purpose in the soundtrack of the film. The intensity of a particular sound directs the viewer's focus to its source. While the above example from Star Wars demonstrates the use and effects of loudness in film it is imperative to note that loudness is relative and dependent on the space that the listener is inhibiting at the time. In a dead-silent room, one can perhaps hear the sound of a clock ticking. However, if the same room is occupied by a crowd of people talking, the same clock disappears from our hearing. With the amplitude of every sound that constitutes the soundtrack of a film being determined by the filmmakers, the relative loudness between these various sounds can be used to interesting effect.

In the film *All That Jazz* (Fosse, 1979), for example, we see the protagonist sitting down at a table while reading his new play. At the end of a thunderous laughter, the diegetic space fades to silence with the only sounds played with any audible loudness being those resulting from the main character. We hear the clicking of his wristwatch, amplified. The sound of his cigarette and the creaks of his chair. While these are all sounds that would normally disappear in the mix, the choice to make them audible brings focus to the character while de-emphasizing all the other elements of the scene (Ament, 2014). The amplification of his foley and sound effects in contrast to the de-emphasis on the rest of the soundscape takes the audience out of the objective point of audition of the scene and puts them into the subjective aural space of the characters' inner mind.

Yet another example of such play with relative loudness can be noted in the film *You Were Never Really Here* (Ramsey, 2017), a film about a traumatized mercenary who saves trafficked girls. In this film, the backgrounds (the ambient sounds that form a bed to create atmosphere for a scene) that are found in the home of our protagonist are mixed to be low while the ambiences of the city are mixed to be more prominent. Unlike the previous example in *All That Jazz*, the ambiences here are both diegetic spaces of the protagonist. The amplitude of these soundscapes are designed to metaphor the inner state of mind of the character. The chaos of the city reflects the state of turmoil he finds himself in due to his past traumas. Conversely, the quieter sections are reflective of Joe's emotional state being at ease when he is at home as opposed to the hostile sound of the city where he feels less comfortable. Thus the sound designers use a change in the perceptual state of loudness (from quiet or loud) to dictate shifts in the psychological state of the character, from a state of relaxation to a state of tension. Perhaps the most well-known use of such metaphorization can be found in the scene from *The Godfather* (Coppola, 1972) when Michael Corleone plans to kill Sollozzo and McCluskey in a restaurant. As Juan Chattah (2015, p. 86) comments, "the abnormal intensification of the sound of an unseen train does not correspond to a realistic representation of diegetic sounds; instead, this sonic outburst reflects Michael's increasing psychological tension up

to the moment he shoots." As in *You Were Never Really Here*, the loudness is a metaphor for the psychological state of the central character, a subjective realization of sound, which Chattah (2015, p. 86) calls more formally "the psychological tension is loudness conceptual metaphor," in which soft sounds correspond to a relaxed state and loud sounds correspond to a tense state.

In each of these three instances, the deliberate manipulation of the loudness of specific sonic elements in relation to the diegetic space creates meaning and opens the film to interpretation. Evoking the audience's own embodied knowledge of the harshness of chaotic city traffic or the screeching squeals of trains, the sound design functions as a metaphor for the psychological state of the character.

Such creative blending of relative loudness is omnipresent in cinema sound and to an extent forms a foundational aspect of the mixing process. Several such instances can be noted in the previously mentioned films of *Cypress* and *Look What You Me Do*. For instance, in *Cypress,* an unnaturally slow-ticking clock inhabits Valerri's home, always out of frame and permanently placed in the acousmatic off-screen zone. Its presence in the mix is emphasized during key moments, marking instances that either highlight the mundanity of rural life or build anticipation for an impending adventure. The manipulation of its volume becomes a "container" for metaphor. On one hand, it draws on the audience's cultural and embodied associations, where the sound of ticking evokes the feeling of waiting. On the other hand, through its deliberately exaggerated and almost tactile slowness, it contrasts with our internalized, embodied sense of how time is expected to pass.

Alternatively, the same meaning making effects can be achieved through the use of relative silence. However, when discussing silence in films, it becomes important to distinguish diegetic silence from what Gorbman (1980, p. 193) terms "non-diegetic silence" (see also Coëgnarts, 2023, p. 69). While the former implies silence in the diegetic space of the scene the latter implies silence that affects the space of the audience. Diegetic silence is often used to create tension or mark particular moments in a film. A good illustration of this is *A Quiet Place* (Krasinski, 2018). This sci-fi film is set in a world where large sightless creatures hunt down anything that makes sound, forcing the survivors of this world to carefully mask their sounds**.** Here the silence occurs within the aural space of the characters and the diegetic space where the scene takes place. The characters in the film can hear and react to this kind of silence. In one such scene, the children, Regan and Marcus take refuge in a grain silo. The sound of the hatch breaking attracts one of the creatures to them, as the children wait, the score fades out and is followed by diegetic silence. The sound of the film at this point features the cave-like ambience of the silo and the sounds of metal squeaking. This moment of relative silence is used to create tension. When structured according to Lakoff and Johnson's metaphorical framework, the abstract concept of tension is mirrored in the scale between silence and loudness. Similarly, silence can be used to mark particular moments. In the film *Fight Club* (Fincher, 1999), the unnamed protagonist tries to inform Marla that her life might be at risk. As the scene takes place in a restaurant, the diegetic aural space is populated with regular sounds one would find in a restaurant: light chatter and the sounds of plates, glasses clinking and subtle diegetic music from perhaps a radio. Marla loses her patience, she yells "Shut up" which is followed by a brief moment of diegetic silence as the crowd gets quiet, and the ambient sounds of the restaurant including the music come to a stop. It should be noted that there is a degree of reflexiveness involved here as there is no motivation for the diegetic music to stop, yet, the silence occurs in the aural space of the characters and is used to emphasize the moment.

Non-diegetic silence, on the other hand, is silence that affects the aural space of the audience but not the characters in the film. Kulezic-Wilson (2009) highlights a powerful example of

such silence using the nightclub scene from *Babel* (Iñárritu, 2006). When Chieko, a deaf teenager goes to a club with her friends, there are moments when the audience is immersed in her perspective. The audience are plunged into a "wall of silence" with sound cutting out entirely. As Cohen (2015, p. 14) it points out, this "brings home the idea that Chieko cannot hear (although, whether her world is truly one of acoustic silence we do not know, but metaphorically, her hearing loss isolates her from the sound of voices and music)."

These kinds of examples of silence can also be called "cinematic silence" (Gorbman, 1987). Taking the aural space of the audience as the point of reference, that is what the audience hears, diegetic and non-diegetic silences can work together or independently of one another to create meaning or as metaphor for underlying concepts. However, such instances of absolute silence, where all sound is omitted from the soundtrack of the film. However, this kind of complete omission of sound from the audiovisual scene is less common in the sound design lexicon when compared to its counterparts.

To demonstrate the use of silence and how it can be used as a vehicle of metaphor, we can analyze the short film *Play/Repeat* (2023) from the body of works sound designed by the first author. The experimental movement-based film is an exploration of the complexities of connections and relationships. Through its focus on a couple who find themselves at the tail end of their relationship, the film delves into themes of mundanity, distance and the despair of fleeting affection. On two occasions in the film (denoted as (a) and (b) in Figure 5), as the two characters make physical contact, the soundtrack momentarily dips into absolute silence.

In both instances, the uncharacteristic silence that punctuates the moments when the characters physically connect serves as a metaphor for the condition of their relationship. While the physical touch in action illustrates their attempt to reconnect, the sound design offers contrasting and contrapuntal commentary, inviting the audience to actively engage in deciphering the underlying metaphor. This choice sharply contrasts with the typical expectation of temporal continuity in film sound.

Fig. 5 The metaphorical use of silence in *Play/Repeat* (2023).

Audiences are accustomed to a seamless soundscape, where ambient sounds and subtle foley naturally flow through such moments. By briefly eliminating all sound, the film disrupts this continuity, creating a disorienting sensory void that serves as a counterpoint to the visuals. The silence not only mirrors the emotional emptiness experienced by the characters but also subverts the viewer's embodied expectation that sound should always accompany the image.

As noted by Gorbman (1987, pp. 18-19), this "non-diegetic" silence often appears in dream sequences, portrayals of intense mental activity, or instances of "self-reflexive playfulness" (see also Kulezic-Wilson, 2009, p. 3). To an extent, it prerequisites the audience's awareness that this kind of silence is not part of the film's lexicon and that it is indeed a reflexive attempt to metaphor an idea. The collective embodied knowledge of the audiences and film culture as a whole is leveraged as such atypical silence breaks filmic convention. This unconventional use of silence breaks typical filmic norms, drawing on the audience's collective embodied knowledge and familiarity with cinematic language. As the CAMWN model suggests, such moments engage the viewer's long-term memory and expectations shaped by film cultures and conventions.

To further explore the use of silence as a metaphorical tool, let us briefly consider two cases from contemporary cinema. First, in *Mid90s* (Hill, 2018), which depicts the story of a 13-year-old boy named Stevie and his circle of friends, silence is utilized in a pivotal moment. Near the film's conclusion, Stevie and his friends are *en route* to a party, with the driver of the car inebriated. Suddenly, the piercing sound of a car horn is heard, followed by a sudden cut to black accompanied by absolute silence. The scene then returns to reveal their car overturned, marking a dramatic turn of events.

Secondly, in *The Sopranos* (Chase, 1999-2007), a television series chronicling the life of Tony Soprano, an Italian-American mobster, silence plays a significant role in the series' iconic final moment. Throughout the series, Tony struggles to balance his criminal endeavors with his familial obligations. In the series finale, Tony waits with his family in a diner for his daughter's arrival. Upon the door chime ringing as the diner door opens out of frame, Tony glances up at the door and the scene abruptly cuts to black, leaving viewers enveloped in silence. In both cases, the lack of sound or the absolute silence is a metaphor for death. In the case of the *Mid90s*, cutting back to picture and the re-emergence of sound and picture follows the narrative that Stevie lives on. Whereas in the case of *The Sopranos*, the lack of any such resolution leaves the state of Tony's life unresolved.

3.3 Spectral dynamics (frequency information and harmonic information)

As established earlier, the spectral content of sound, that is its frequency information, plays an important role in defining its perceived loudness. It is through the manipulation and blending of various frequencies, that sound designers are able to weave a soundtrack that supports the narrative of a film. Through sculpting the spectral content through emphasis and de-emphasis of certain frequencies over others, the sound design of the scene can become the container of metaphor. The relationship between spectral density and volume is evident in the earlier segments of *Cypress* and *Look What You Made Me Do*. In *Cypress*, the lush sound design fills the frequency spectrum, creating a rich intensity that suddenly gives way to a sparse soundscape. Similarly, in *Look What You Made Me Do*, the scene with Ingmar in his car layers distinct spectral elements: the metallic clink of a boiling pot fills the high frequencies, while rain and non-diegetic crashing waves spread energy across the spectrum. The drone that builds to a crescendo and the rumble of thunder contribute to the lower frequencies. The title card further amplifies this effect by layering multiple takes of the Free Jazz score, creating a chaotic cacophony that increases the spectral density. These

Fig. 6 From a state of unity (a) to one of solitude (b) (from *Look What You Made Me Do*).

densely layered moments are contrasted with relatively sparse passages, amplifying the contrast. They reinforce the metaphors explored earlier: in *Cypress*, a "dense" spectrum evokes a sense of expansive hope, while "sparsity" suggests isolation. In Ingmar's scene, density reflects anxiety and tension, while sparsity mirrors dissociation. These contrasts tie into universally felt bodily experiences.

Taking another segment of *Look What You Made Me Do,* we see Ingmar on a date with a woman (Figure 6a). Upon noticing the woman's apparent lack of interest in his taste in music, he finds himself offended. The camera slowly tracks in from a two-shot to framing Ingmar alone (Figure 6b). As this happens, the soundscape morphs from realizing the objective elements of the scene to a subjective one. The diegetic music of the restaurant begins to loop, the ambience dips and the woman's voice fades into reverb. As the dolly shot closes in, the sound of dishes and cutlery begins to tremble unnaturally, as if an earthquake is occurring. Through spectral analysis of the scene as depicted in Figure 7, we see that the high frequencies diminish (marked as (a)) and a droning low-frequency hum begins to overpower the soundscape (marked as (b)). The aural dynamics of the scene are composed by shaping the frequency content of the scene and in turn these dynamics channel the emotional state of the character. The attenuation of high frequencies, signaling Ingmar's detachment from the moment, evokes an embodied sensation familiar to audiences— "spacing out" due to disinterest or disengagement. Meanwhile, the emergence of low-frequency rumbles and the trembling of cutlery metaphorically capture the feeling of simmering anger. These sonic metaphors draw on culturally and socially embodied associations between frequency, emotion, and psychological states—associations that audiences have internalized through both lived experience and cinematic convention.

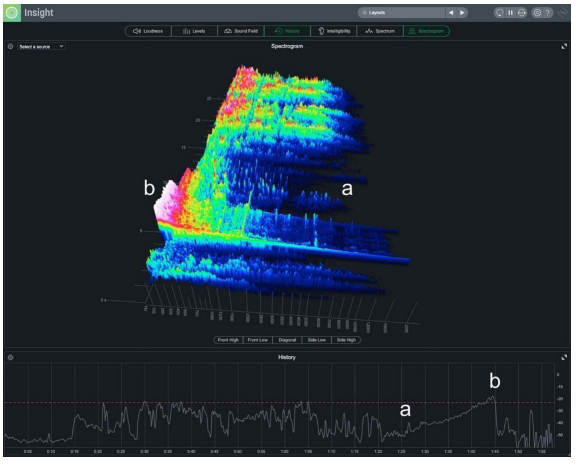


Fig. 7 Spectral analysis of the scene as illustrated in Figure 6. The spectrogram visually depicts the frequency spectrum over time. The color or intensity of the spectrogram represents the amplitude of the frequencies.

The use of such low-frequency sounds is a common occurrence in the vocabulary of film sound. Examples of this are the so-called "braam" effect used in films such as *Inception* and other blockbuster films. In *Irreversible* (Noé, 2002) which features the use of a low frequency (27hz) sound (Noé, 2003). In both cases, these low-frequency sounds have been used to induce a sense of emotional arousal in the audience. The application of sounds with specific frequency ranges emphasized allows sound designers to draw out empathetic reactions from viewers. A well-known example of this would be the dinosaur roars from *Jurassic Park* (Spielberg, 1993), where sound designer Ben Burt stacked and processed various animal roars such as baby elephants and dogs to create the sound of the T-rex (Buchanan, 2015). Studies into this matter have indicated that nonlinearities (the deviations from mid-range frequencies) in sound have the predictable emotional effects of increased arousal (i.e. perceived emotional stimulation) and negative valence (i.e. perceived degree of negativity or sadness) (Blumstein, Bryant, & Kaye, 2012). A notable example spectral dynamics can be found in the film *The Conversation* (Coppola, 1974), a film that revolves around a surveillance expert, Harry Caul (Gene Hackman), and the moral dilemma he faces when his recordings reveal a potential murder plot. Throughout the film, the piano score has varying degrees of distortion applied to mirror the psychological and affective state of its main character. As Harry searches for a possible surveillance device hidden in his home in a furious rage, we hear a sharp ostinato on the piano followed by pounding tritone (G and C Sharp). The tritone being one of the most dissonant intervals reflects Harry's state of mind. Once he gives up searching for the device, the sound returns to the relatively stable main theme in the key of A minor. Thus, film composers may "infuse", as Chattah (2024, p. 15) remarks, "a score with harmonic consonances and dissonances to reflect narrative developments, using music as an off-screen narrator".

In the same way the variations in harmonic consonance and dissonance are used to express dynamics in film scores, sound designers and mixers use changes in timbre and texture to shape perceived dynamics within the soundscape. These variations, often described using cross-modal

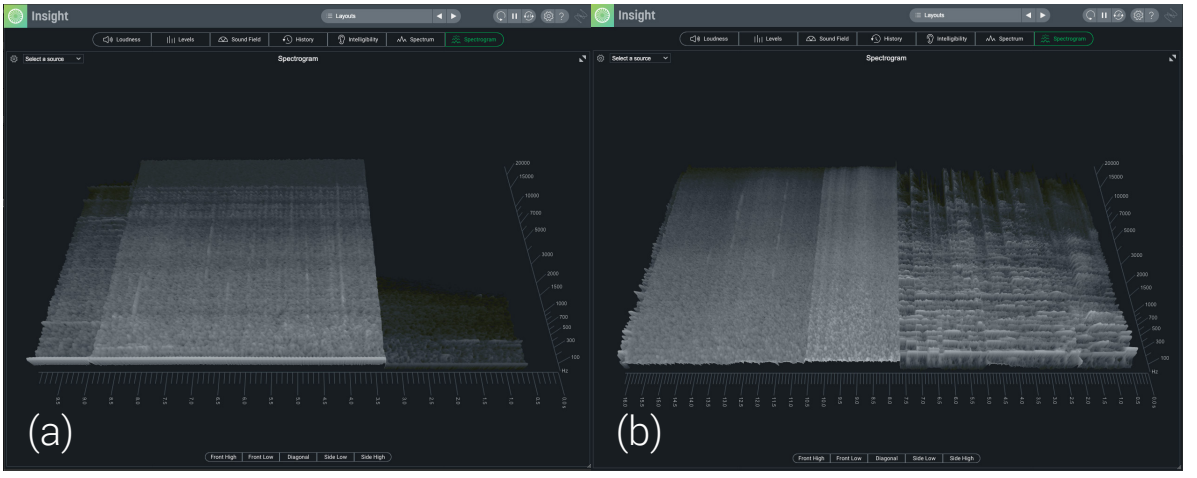


Fig. 8 In the spectrogram view of iZotope Insight, the x-axis represents time, while the y-axis shows the frequency range from low to high. The intensity of each frequency is visualised through the height and brightness of the spectral content, indicating greater amplitude or energy. Here, spectral dynamics are used to create a sense of contrast in *Look What You Made Me Do* (a), and *Play/Repeat* (b).

descriptors, such as smooth, rough, bright, dark map tactile and visual qualities onto the sonic domain. And like perceived loudness, they exist along scales that range in frequency content, spectral density, and articulation. To demonstrate this action, we turn to present the spectrogram analysis of passages from three films.

In each of the passages, spectral dynamics were deliberately manipulated to create contrast and, by extension, convey narrative meaning. In *Look What You Made Me Do* (a), the frequency spectrum is densely saturated prior to a cut, represented in the spectrogram as a dense wash of information followed by an abrupt drop. As previously discussed, multiple layers of sound effects were added to the scene in which Ingmar sits in his car, each emphasizing different parts of the frequency range. When combined, these layers saturate a broad portion of the spectrum. This sonic density functions as a metaphor for the character's emotional state: a sustained, simmering anger, followed by a sudden cut to him submerged in a pool. In *Play/Repeat* (b), a similar technique is used to convey overwhelming emotional weight, reflecting the despair experienced by the couple.

3.4 Sound field

A more contemporary form of dynamic expression comes from surround sound systems that allow sound designers to manipulate the width of the sonic image. Modern cinemas are fitted with an array of speakers strategically placed in the cinema room to create immersion and a sense of space. These speaker arrays allow the mixers to populate both the "visualised on-screen zone" and the "acousmatic off-screen zone" with sounds that reinforce the objective reality of the scene. Further, it is common practice for non-diegetic elements of the soundtrack, such as score and sound design to be pushed into surround space to create greater immersion

Fig. 9 Surround with through (a) all five surround channels versus (b) only the front screen speakers.

for the audience. A significant aspect of the mixing process is precisely situating each sound element within the spatial field. As such, by adjusting the sound stage's width and depth, filmmakers can craft a dynamic soundtrack, not solely reliant on loudness but through manipulation of the stereophonic image.

Such manipulation of space can be heard in the film *Saving Private Ryan* (Spielberg, 1998). Early on in the film, we see the protagonist John H. Miller (Tom Hanks), advancing up the beach with the allied troops. The camera intermittently plunges underwater as the soldiers make their way towards the beach, providing breaks from the chaotic sounds of war. At a later moment in the same scene, Miller is dazed by witnessing the intense fighting and violence. Shell-shocked, we hear the sound through his point of audition, muffled, distant and without detail. In both cases, the mix gets quieter and more subjective, and the penta-phonic sonic image narrows to stereo. As Holman points out "The contrast of all these elements draws us into a different point of view for a while, before we are thrust back to all 6 channels of reality" (Holman, 2008). The surround width narrowing marks the soundscape shifting from the spatial aural space of the character (objective) to the aural space of his inner mind (subjective). It is important to note that these spatial changes work in tandem with dynamic and spectral variations, further contributing to the shifts from wide, detailed sound (target domain: external reality) to muffled, reduced sound (source domain: internal psychological state of disorientation and detachment).

To further illustrate the dynamic use of the sound field to metaphor abstract concepts, we present the analysis of experts from two films. First, the aforementioned passage from the early minutes of *Cypress*. In Figure 9, the six vertical bars in the Insight tool represent the volume of audio passing through the five surround channels (left, right, center, left

Fig. 10 From width to monophonic in Pedra Seca (2024).

surround, right surround) and the additional Low Frequency Effects (LFE) channel that form a 5.1 mix. Through the sequence of shots that show Greece and its abundant natural beauty, the soundtrack utilizes the full range of speakers available. The sounds of waves move through the room, starting from the screen speakers before they wash into the rear surround speakers. The sound of nature, the choral score and the lush reverb utilize the full array of speakers to envelope the audience from all sides. However, upon transition to the pine forest in Ukraine, the soundscape relies solely on the front screen speakers.

The manipulation of the sound field for metaphorical layering is not limited to surround and immersive audio formats like 5.1, 7.1, or Dolby Atmos. While these formats enable for complex and precise spatial positioning, more subtle tools such as panning, reverb, and delays are often used to evoke a sense of space and convey emotional or narrative meaning. These techniques help shape how the audience perceives the setting, characters, and their psychological states. For instance, in *Cypress*, in addition to widening and narrowing the available soundstage across the surround sound channels, the choral score is treated with a lush, long reverb. These reverberations evoke the embodied connotations of specific physical spaces, such as the expansive echo of a forest or the grand reverberation of a church. Again, drawing on the audiences embodied experience of inhabiting environments where such reverberation naturally occurs to accentuate the thematic elements of the narrative.

A similar use of such surround sound dynamics can be highlighted in the film *Pedra Seca* (2024). A film that revolves around the lives of two contrasting women and the ensuing friendship that forms when their paths cross. One is a young traveller journeying through an uninhabited part of the Portuguese countryside while the other is an elderly woman who has decided to live the life of a recluse, detached from any social and emotional connection. While the exterior shots of the films mixed to reflect the width of the scenic landscapes of the Portuguese hillside, the hermetic house of the old lady narrows to monophonic sound utilising the centre speaker exclusively (Figure 10).

Both these films contain thematic elements of isolation and solitude and their respective mixes analogize these ideas by sculpting the sonic panorama. The available sound field is broadened and narrowed as metaphors for the conceptual ideas present in the film.

4. Conclusion

The goal of this article was to leverage the author's practical experience as a sound designer to examine how contemporary sound designers use aural dynamic range—through loudness, silence, and dynamic contrasts—as a tool for embodied narrative storytelling. The first part of the article provided a theoretical framework for understanding dynamics in sound. This set the stage for the second part, where we delved into practical applications, illustrating how sound designers manipulate dynamic variations to achieve specific narrative effects. By classifying the use of dynamics into four categories—contrasting dynamic range, relative loudness and silence, spectral intensity, and surround width—we revealed how sound design operates both bodily and metaphorically. However, it is important to acknowledge that these categories are often not isolated but blend together in the sound design process, with each category influencing the others. This interconnected approach allows sound designers to create a more nuanced and layered soundtrack, enhancing both the emotional and narrative impact of a film. These dynamic variations are not simply deployed in isolation but often in tandem, working together to produce effects that resonate both metaphorically and physically with the audience. Ultimately, the analysis shows that sound designers effectively communicate with audiences by tapping into the

intrinsically embodied nature of human cognition. The meanings conveyed through sound design resonate on a cognitive unconscious level, making the auditory experience an integral part of cinematic storytelling. This embodied connection enables sound designers to engage viewers in a profound and intuitive manner, enriching the overall filmic experience. By integrating theoretical insights with practical applications, this study underscores the importance of considering embodied cognition in the analysis of film sound. It contributes to a deeper understanding of how sound design can evoke meaning and emotion, aligning the audience's sensory-motor experiences with the narrative unfolding on screen. This approach not only broadens the scope of film theory but also enhances the practice of sound design, paving the way for more immersive and impactful cinematic experiences.

References

Ament, V. T. (2014). *The foley grail: The art of performing sound for film, games, and animation* (2nd Edition ed.). Focal Press.

Behar, A. (1984). Intensity and sound pressure level. *The Journal of the Acoustical Society of America*, 76(2), 632. doi:10.1121/1.391117

Blumstein, D. T., Bryant, G. A., & Kaye, P. (2011). The sound of arousal in music is contextdependent. *Biology Letters*, 8(5), 744-747.

Bodner, E., Gilboa, A., & Amir, D. (2007). The unexpected side-effects of dissonance. *Psychology of Music*, 35(2), 286–305.

Brower, C. (2000). A cognitive theory of musical meaning. *Journal of Music* Theory, 44(2), 323-379.

Buchanan, K. (2015, June 9). You'll never guess how the dinosaur sounds in Jurassic Park were made. Retrieved from *Vulture*: https://www.vulture.com/2013/04/how-thedino-sounds-in-jurassic-park-were-made.html

Carroll, J. M. (1980). *Toward a structural psychology of cinema*. Mouton.

Chattah, J. (2024). *Film music: Cognition to interpretation*. Routledge.

Chattah, J. (2015). "Film Music as Embodiment." In M. Coëgnarts, & P. Kravanja (Eds.), *Embodied cognition and cinema* (pp. 81-112). Leuven University Press.

Chion, M. (2019). Audio-vision: Sound on screen. 2nd ed. Colombia University Press.

Claxton, G. (2015). *Intelligence in the flesh: Why your mind needs your body much more than it thinks.* Yale University Press.

Cohen, A. J. (2015). Congruence-association model and experiments in film music: Toward interdisciplinary collaboration. *Music and the Moving Image*, 8(2), 5-24.

Coëgnarts, M. (2017). Cinema and the embodied mind: Metaphor and simulation in understanding meaning in films. *Palgrave Communications*, 3, 1–15.

Coëgnarts, M. (2019). *Film as embodied art: Bodily meaning in the cinema of Stanley Kubrick*. Academic Studies Press.

Coëgnarts, M. (2020). How motion shapes thought in cinema: The embodied film style of Éric Rohmer. *Projections: The Journal for Movies and Mind*, 14(2), 26-47.

Coëgnarts, M. (2023). Shaping films from the inside out: Embodied mental schemas in filmmaking and viewing. *Baltic Screen Media Review*, 11(1), 128-147.

Coëgnarts, M., & Kravanja, P. (2012). Embodied visual meaning: Image schemas in film. *Projections: The Journal for Movies and Mind*, 6(2), 84–101.

Coëgnarts, M., & Kravanja, P. (Eds.). (2015). *Embodied cognition and cinema*. Leuven University Press.

Cox, A. (2017). *Music and embodied cognition: Listening, moving, feeling, and thinking*. Indiana University Press.

Cox, A. (2018). On the subjects and objects of music analysis. In N. Reyland, & R. Thumpston (Eds.), *Music, analysis, and the body: Experiments, explorations, and embodiments* (pp. 15-30). Uitgeverij Peeters.

Fahlenbrach, K. (2008). Emotions in sound: Audiovisual metaphors in the sound design of narrative films. *Projections: The Journal for Movies and Mind*, 2(2), 85-103.

Forceville, C. (2006). Non-verbal and multimodal metaphor in a cognitivist framework: Agendas for research. *Applications of Cognitive Linguistics*, 1, 379.

Forceville, C. J. (2006). Non-verbal and multimodal metaphor in a cognitivist framework: Agendas for research. In Kristiansen, G., Achard, M., Dirven, R., & Ruiz, F. J. (Eds.), *Cognitive Linguistics: Current applications and future perspectives* (pp. 372-402). Mouton de Gruyter.

Gibbs, R. W. (2005). *Embodiment and cognitive science*. Cambridge University Press.

Gorbman, C. 1980. Narrative film music. *Yale French Studies*, 60, 183-203.

Gorbman, C. (1987). *Unheard melodies: Narrative film music*. BFI.

Hunter, A. (2012). When is the now in the here and there? Trans-diegetic music in Hal Ashby's *Coming Home*. *Alphaville Journal of Film and Screen Media*, 3.

Johnson, M. (2007). *The meaning of the body: Aesthetics of human understanding*. The University of Chicago Press.

Johnson, M. (2013). Identity, bodily meaning, and art. In Roald, T., & Lang, J. (Eds.). *Art and identity: Essays on the aesthetic creation of mind* (pp. 15-38). Rodopi.

Johnson, M. (2018). *The aesthetics of meaning and thought: The bodily roots of philosophy, science, morality, and art*. University of Chicago Press.

Johnson, M., & Larson, S. (2003). Something in the way she moves – Metaphors of musical motion. *Metaphor and Symbol*, 18(2), 63-84.

Kulezic-Wilson, D. (2009). The music of film silence. *Music and the moving image*, 2(3), 1-10.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. The University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: the embodied mind and its challenge to western thought*. Basic Books.

Larson, S. (1997). Musical forces and melodic patterns. *Theory and Practice*, 22/23, 55-72.

Madupu, T. (2020). *The meaning and emotional impact of aural dynamics in film*. Kino Eyes The European Movie Master Erasmus Mundus Joint Master Degree.

Mckee, E. (2007). The topic of the sacred hymn in Beethoven's instrumental music. *College Music Symposium*, 47, 23-52.

Melvin, A., & Bridges, B. (2021). Mapping space and time in the soundtrack: Embodied cognition and the soundtrack's spatio-temporal contract. In Filimowicz, M. (Ed). *Doing Research in Sound Design* (pp. 37-55). Focal Press.

Metz, C. (1968/1974). *Film language*. Oxford University Press.

Metz, C. (1971/1974). *Language and cinema*. Mouton.

Metz, C. (1974). *Language and cinema*. Mouton.

Moore, B. C. (2005). Loudness, pitch and timbre. In E.B. Goldstein (Ed.), *Blackwell handbook of sensation and perception* (pp. 408-436). Blackwell publishing.

Munson, W., & Fletcher, H. (1933). Loudness, its definition, measurement and calculation. *The Journal of the Acoustical Society of America,* 5(2), 82-108.

Noé, G. (2003, March 13). There are no bad deeds, just deeds. (j. Tang, Interviewer). Salon.com.

Ortiz, M. J. (2011). Primary metaphors and monomodal visual metaphors. *Journal of Pragmatics*, 43(6), 1568-1580.

Ortiz, M. J. (2023). Embodied cinematography in *Mr. Robot*. *Baltic Screen Media Review*, 11(1), 84-105.

Pearson, B. (2018, January 24). Rian Johnson Talks About The Holdo Maneuver, One of 'The Last Jedi's Most Breathtaking Moments. Retrieved from https://www.slashfilm.com/theholdo-maneuver/

Saslaw, J. (1996). Forces, containers, and paths: The role of body-derived image schemas in the conceptualization of music. *Journal of Music Theory*, 40(2), 217-243.

Shapiro, L. (2019). *Embodied cognition*. Routledge.

Stilwell, R. J. (2007). The fantastical gap between diegetic and non-diegetic. In D. Goldmark, L. Kramer, & R. Leppert (Eds.), *Beyond the soundtrack: Representing music in cinema* (pp. 184-202). University of California Press.

Tversky, B. (2019). *Mind in motion: How action shapes thought*. Basic Books.

Ward, M. (2015). Art in noise: An embodied simulation account of cinematic sound design. In M. Coëgnarts, & P. Kravanja (Eds.), *Embodied cognition and cinema* (pp. 155-186). Leuven University Press.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625-636.

Zbikowski, L. M. (2008). Metaphor and music. In R. W. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 502-524). Cambridge University Press.

Zettl, H. (2017). *Sight sound motion: Applied media aesthetics*. Cengage learning.