Breaking the 4th Wall of Sound: The Paradox of Screen Music
Professor Milton Mermikides
16 January 2025

I: Firesides, Curtains and Screens

Human beings are storytelling creatures. Other animals report to each other information about the world, and sometimes extensively, such as the long songs of whales telling of their identity and location, and bees whose elegant 'waggle dance' language communicates the direction and distance of resources. However, it seems like a *story heritage*, and in particular the use of *fiction* (stories that are *knowingly untrue*) are unique to us humans. This insatiable need to recount and embrace fictions, might be born of evolutionary pressure: perhaps a way to charm (or deceive) others, or as a mental training ground for real events. Its use may be directly advantageous, a form of social bonding, or it may be a 'useless' byproduct of our ever vigilant and tetchy minds, incessantly probing the '*what ifs*' and '*remember whens*' in order to gain treasures or avert disasters: Unable to turn off, our story-telling minds fabricate worlds for no reason at all. Whatever its origin, we spend much of our waking lives drifting into daydreams, reliving, or imagining future scenarios where we are the victor or victim, avenging or avenged, spectators, heroes, or villains of our own stories. And we continue the process even when we sleep, weaving in our dreams yet more fantastical tales. If this natural state of continual internal storytelling is not enough, we have constructed through eras and cultures, communal arenas for storytelling. The parent-child, and group act of storytelling have served us for millennia as a means to create shared imaginary worlds. These stories not only explain the world but provide comfort, distraction, community, morality tales, entertainment, and a shared history whether real or imagined. The story-teller – in order to lure the listener deeper into the tale – has often used a (literal as well figurative) veil at the threshold between the real and imagined worlds: The flickering flames at a campfire hinting at the forms within, the curtain that opens to reveal the fantasy land, and the silver screen of the cinema whose flat surface is made to dissolve into a deeply immersive neverland. The now portable (small but life-consuming) digital devices, and the brave new virtual reality worlds are but sophisticated evolutions along this ancient theme: Veils between our 'real' and the 'imagined' worlds.

This positions the storyteller as the "veil-lifter," someone who reveals deeper truths beyond the visible. Sometimes, the storyteller's role is neutral - reporting on the happenings of this imagined world without breaking the illusion. However, the canny self-aware storyteller (or fictional character) - who stands at the threshold of these two worlds and able to speak of and to the characters in the story-world, and to us (the audience) directly - is an ancient idea. The Greek chorus, the Shakespearean narrator, and the countless examples of direct address from Dionysus to Daffy Duck to Deadpool act as intermediaries between the audience and the imagined world: Simultaneously weaving – and revealing the stitching of – the story-world. In this lecture I take on that role as we examine one particular arena of storytelling, **the cinema** (and other of its screen-based descendants) and the powerful role that sound and music play in constructing and subverting the storytelling process.

II: Let There Be Light, Sound and Music

The introduction of 'moving images' – a sequence of still images 'coming to life' – represented a major evolution in our story-telling ability. The point of origin – and precise definition – of this craft is nebular but early varied examples include the Ancient Chinese *chao hua chich kuan* 'fantasy pipe' (180 AD) and its 19th Century revival in the *zoetrope* ('life-turn') carousel of images; Linnett's 1868 patent of the flip book; La Prince's – the 'father of cinematography' – 1888 2-3 second film footage of a Roundhay Garden Scene; and the 1895 *cinématograph* exhibition in Paris by the appropriately named Lumière Brothers. In every case it was found that when presented with a series of still images (whether illustrations or pieces of film) at a sufficiently high frame rate[1], the spectator creates the illusion of continuity. We 'fill in' a smooth motion between the fixed images to create an illusory new world. This sense of motion is so intuitive and 'real' it is embedded in the term: 'moving images', and many of the earlier labels derive from 'life' 'movement' 'animation' and being 'written' (cinematography = movement writing). Not only do we associate movement with a series of images but we also ascribe agency – a purpose and intent – to even very abstract images when presented in succession: Heidel and Simmer's 1944 psychological study presents an animation of geometric shapes which was interpreted almost universally by their subjects as a meaningful narrative, demonstrating at the same time our innate story-telling flair and the 'realness' of moving images.

Despite the extraordinary innovation, the lack of sound in the moving image was a limitation and challenge for storytelling. The literal and figurative flatness of images might have been made more apparent by their silence, and even though a narrative is presented visually, exactly what it says and how it should be emotionally received is unreliable without sound or music[2]. One can track in the history of cinema attempts to bring the moving image into narrative life with the use of *intertitles* acting as both narrative signposts and a stand in for dialogue. The larger-than-life on-screen theatrics of eye-lined melodramatic gestures to Buster Keaton's astonishingly impressive – and dangerous – stunts, compensated somewhat for the awkward silences and small screens of early cinema[3]. Although it was initially impossible to synchronise audio with video, music – performed locally at cinema performances – were used to enhance immersion. This served the multi-purpose of drowning out the distracting and spell-breaking noise of the projector, evoking a time and place for the narrative, and to help guide the viewer to what they were 'supposed to feel' with emotionally unambiguous music. This music was usually improvised live by the musicians at the venue (often using popular songs of the day or a repertoire of Western canon pieces); sent out as a score as part of the product; or taken 'off-the-shelf' with specially prepared editions of 'photoplay' music (including useful moods: 'Love scene' and 'Chase') – a practice which in essence remains the same in contemporary digital library music. 1908 saw the first bona fide film score for Western cinema, composed specifically by Saint-Saëns for *The Assassination of the Duke of Guise* (*L'Assassinat du duc de Guise*) to be performed by an orchestra live – quite spectacularly no doubt – at specific venues. Erik Satie's 1924 score for René Clair's *Entr'acte* (part of Picabia's ballet *Relâche*) is among the first examples of a music score where the interaction of music with each shot is tightly considered. A typical silent movie model is shown in Figure 1: the audio and visual was separated at the screen, and the story world and its telling was separated 'in' the screen with on-screen action and intertitles/ credits respectively.

---

[1] The frame rate required for the illusion of movement is surprisingly low, 24 frames per second has been used quite happily for a century now. Compare this to the 44,110 (and more) audio snapshots needed every second (for each ear) to reproduce continuous sound.

[2] The power of music to accompany motion pictures was discovered early, even the 1895 Lumière Brothers exhibition included live music. That said the lack of synchronised sound certainly enforced and inspired the staggeringly rapid evolution of the cinematic narrative visual language (Eisenstein's 1925 epic silent movie *Battleship Potemkin* is a seminal work of visual story-telling).

[3] This increasing drive to lure the audience 'through the screen' continues to this day, aspect ratios of screens and loudspeaker arrays have expanded ever wider, enveloping the audience.
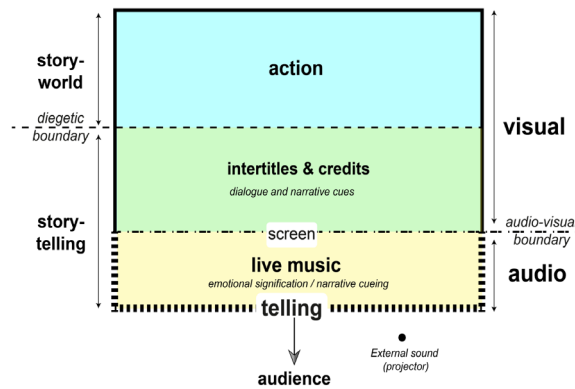
Figure 1: A typical model for the separation of audio and visuals (at the screen); and the story-world ('diegetic' or 'diakosmic') from its telling ('non-diatonic' or 'paradosic') (see IV)

With the introduction of technologies to synchronize sound to film (such as *Vitaphone* which mechanically linked the projector and audio playback, and the more sophisticated and reliable *Movietone* which printed an optical sound wave directly on the film), filmmakers gained an unprecedented ability to marry sound and image and exploit the powerful symbiotic effect such a sensory bonding[4] had on the audience. Cinema took a short while to come to terms with this superpower, its introduction in 1926, was the premiere of the silent film *Don Juan*, which used the technology to retrofit a symphonic musical score and sound effects (dialogue was a step too far). The first 'talkie' *The Jazz Singer* (1927) used dialogue sparingly (synchronisation was a challenge and many intertitles were retained) but the use of music was largely on-screen performance, an 'overture' (dutifully labelled with an intertitle) and a touch of underscoring. The use of a sound is cautious: conservative and single layered – we hear a performance in a bar (but no background noise), spoken dialogue is not underscored. Now with all audio living 'behind the screen'. the relationship between audio, vision; and the story from its telling became a little more fluid. Dialogue could now either be heard 'in the story world' with synchronised audio or represented more abstractly as intertitles. On-screen events could now be both seen and heard (even though these were limited to dialogue and music, rather than foley). Notice that music now takes on a flexible status: it could 'live in' the story (as on-screen performance), or as film scoring external to the story world, but used to guide and enhance our experience. Figure 2 illustrates these relationships in the 'first talkie' – *The Jazz Singer.*
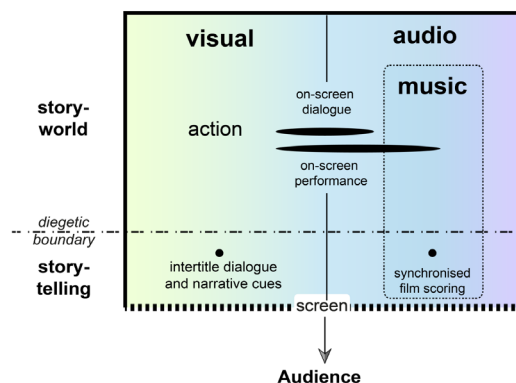

Figure 2: A representation of the relationship between audio, visuals, the story-world and storytelling in *The Jazz Singer* (1927)

---

[4] This audio-visual bonding, film theorist Chion terms *synchresis*. Consider the hybrid effect of dissonant violin and on-screen knife - both stabbing - in Hitchcock's *Psycho*.

III: The Principles of Classic Screen Music

Hollywood was remarkably quick to find its footing with these new technologies. With *King Kong* (1933) featuring complete dialogue and an elegant score by Max Steiner, the industry had quickly gathered its audio-visual and film scoring cinematic language, one that remains remarkably intact to this day. Claudia Gorbman's seminal 1987 work, *Unheard Melodies*, outlined key principles that define how such 'Classical Hollywood' film music (in the model of Max Steiner) supports the cinematic experience:

1. **Invisibility** (or *ephemerality*): The technical apparatus of screen sound need not be visible: we don't see how the sound is gathered. We might extend this to say that sound and music in film is somehow *ephemeral*. We as the audience have a privileged access to sound – a 'floating ear' that can hear clearly all environmental, dialogue, even inner thoughts, and is untroubled by the acoustic realities.

2. **Inaudibility** (or *subordination* and *subliminality*): Music should subordinate itself to dialogue, visuals, and the narrative. This subordination manifests in terms of volume and timbral balance (ducking and avoiding the dialogue's frequency range), entrances of the film score are usually undetected 'felt rather than heard', and screen music tends to be 'stretchy' (with the use of ostinato, flexible tempo etc) so that its duration and shape is malleable so as to be governed by the narrative, rather than the reverse.

3. **Signifier of Emotion**: Music expresses emotion to guide the audience's interpretation of the scene. Emotional colouring tends to be unambiguous and 'readable' – threat, serenity, an untrustworthy character, elation, disorientation, romantic love etc. – how we are 'supposed to feel' about a character or situation is clearly signified.

4. **Narrative Cueing**: Music provides formal structure ("here is a new section of our story") and a sense of time/place/culture (a future city with gritty techno, the vastness of space with harmonically distant epic chords, the deep South with slide guitar, a refined social gathering with a string quartet). A wordless narration through the story's acts and scenes.

5. **Continuity**: To smooth the *joltiness* of edits, camera changes and movement, music connects shots and scenes, creating a seamless flow across otherwise awkward silences, visual edits, imposing a temporality and 'narrative lubricant' to shots whose ordering in time is ambiguous, distracting, awkward or otherwise confusing.

6. **Unity**: Music contributes to the overall coherence of the film with the use of repeated musical material (leitmotifs and instrumental markers for characters, places, objects or the film itself)

To these rules (principles really) Gorbman adds the 7th that any of these can be broken if in service to any other. Let's call this principle:

> 7. **Flexibility.** All other principles may be twisted or subverted to enhance the narrative experience.

And despite the remarkably robustness of the first 6 principles of audio-visual cinema's century, this flexibility happens regularly – the *Star Wars, Indiana Jones, Mission Impossible* and *James Bond* themes are clearly consciously heard when they appear in their respective scenes (breaking 2), adding to a sense of communal spectacle and unity across the franchise; the use of *counterpoint* - chillingly serene classical music during a villain's murderous acts, or the tender ballad *What a Wonderful World* during the wartime scenes in *Good Morning Vietnam* (1987) signify the 'wrong' emotion with a conspicuous irony or lack of empathy (breaking rules 2 and 3). Occasionally we 'hear' through the ears of a character (Beethoven's impairment in *Immortal Beloved*, the white noise of a character zoning out in a conversation, or underwater submersion[5]) revoking our auditory privileges (breaking rule 1) in order to enhance the subjective

---

[5] See the opening sequence of *Saving Private Ryan* (1998) where the subjective ear whether physical (underwater or after explosions) or psychological (traumatic detachment) is used to powerfully immersive effect.

immersion. The *Better Call Saul* (2015) theme tune cuts off abruptly (almost like a technical glitch) just one the cusp of its final tonic (breaking 2 and 5) indicating perhaps that this is a story like no other, and the non-linear nature of the narrative. The opening car chase of *Baby Driver* (2017) includes music playing in the protagonist's headphones (of which we are privy to thanks to rule 1). The music should (according to rule 2) be subservient to the onscreen events, but it is a pre-existing track[6] – fixed – and holds dominance over the ensuing action (breaking rule 2), as if every frantic lurch, skid and collision are dictated by the music which makes no plausible (or even acoustic) sense, but says something of Baby's skill, control over the circumstances and general swagger.

IV: Diegetic Boundaries and Walls of Sound

Let's return to our concept of the veil between the world of the story and that of the audience. The characters live behind the screen (in the 'story world') and we the audience can see into their world. Although we can see them, they – in general – can't see us through the veil[7] (or 'fourth wall') as if it is the 'one-way' mirror of an interrogation room. Interlocutory narrators, and some special characters are imbued with the superpower – the self-aware lucidity – to *know they are in the story world* and can 'see through' the veil, break the 4th wall and even address us directly[8]. However, the general principle remains that there is a dividing line – a one-way screen – between the story world and the audience. How does sound and music interact with the veil? Well in the case of the silent movies the distinction is simple, music is added in a higher-level 'presentation layer' for our consumption. The story world is silent, but our audience experience is enriched and guided by music.

The introduction of synchronised sound however complicates matters. Now, sound (and music) can emerge not only from the 'story-world' but from a 'presentational layer'. Characters may speak or even sing 'in the story' while music on 'our side of the veil' is played. We now hear two layers of sound – while the story world only has one.[9] In order to navigate this complexity, Gorbman formulated the terms *diegetic* sound/music for sonic materials emerging from the story world and its characters; and its opposite the *non–diegetic* (sonic material external to the characters' awareness and the story-world environment).[10] Simply put: diegetic (or *diakosmic*) sound 'belongs' to the story, and non-diegetic (or *paradosic*) sound is part of its *telling.*

As opposed to a silent movie ('externally' scored with instrumental music), synchronised audio in a film now has multiple narrative functions and sonic 'spaces'. An example: a scene starts with a gentle guitar chord setting up mood, environment, and broad cultural reference as we see a forest at sunset, the sun reflecting off the surface of a lake. The music is (we presume) non-diegetic and serves Rules 3 and 4, locating us in geographical, cultural, and emotional space. With the music we hear bird song and rustling leaves. We don't see the source of these but naturally assume it is diegetic, unseen elements that live in the story world. A sung voice joins the guitar (we recognise the melody from elsewhere in the film - rule 6) as the camera pans to reveal a girl leaning against the tree strumming and singing – we reassess the music as diegetic, and perhaps notice that it is now more acoustically natural and 'plausible'. A narrated voice joins "I

---

[6] *Bell Bottoms* by the Jon Spencer Blues Explosion

[7] What do we imagine the story characters see on the other side of the veil? A surrounding environment perhaps, or in the case of a more familiar staged performance, an in-story audience?

[8] Examples of 4th wall breaks and 'direct address' are numerous and diverse in cinema (see Brown 2012), and vary in their nature and 'intensity'. From a glance (e.g., Eddie Murphy's look to camera as if to say to us "can you believe this guy" in *Trading Places*), a character's moments of lucid awareness (Abed in *Community* using the terms 'episode' and 'season') or more extreme technical breaks such as Jerry Lewis calling attention to the cameras and walking 'off set' in *The Patsy* (1960). Films can also be *about* 4th Wall breaks, such as the 'awakened' film characters in *The Purple Rose of Cairo.*

[9] The description is of course completely illusory, all the sound lives 'in the film' and propagates from the same speakers. There are some conventions in some speaker arrays (e.g., dialogue emerges from the centre speaker in 5.1 while the film score is spread across the sound field), but in general we – from context – 'demix' the sound elements and understand where they 'live' in – or above - the story.

[10] Despite the brilliance and influence of Gorbman's work, I find the term *diegetic* (with its Greek roots in 'narration') a little misleading, and its counterpoint *non-diegetic* (which can include narration!) rather negative. I have suggested *diakosmic* (from the (small) cosmos of the story) and *paradosic* (from the Greek to deliver) as more compelling terms, while holding no hope of changing film music studies terminology any time soon.

spent countless hours every summer, singing and playing by that lake". This is clearly non-diegetic (*paradosic*) dialogue, unheard in the scene and for the audience's privilege (or are we listening in to its diegetic telling?). The source is clear, we see who the voice belongs to, but the speech is happening at a different time and place, and not vibrating the virtual air molecules of our story world. As the narration continues, we feel (but don't notice) a low drone that recasts the major melody in a minor tonality. The drone morphs (Rule 5) into the aircraft engine noise as we dissolve into a war scene. This relatively simple segment – employing now standard cinematic devices – involves elaborate and complex (even if unconscious) moments of identification, evaluation, navigation, and subliminal manipulation of the audience. From one continuous waveform a dynamic and overlapping field of soundworlds are conjured, some part of the scene (whether or not visualised), other layers above, and others somehow floating between these boundaries.

V: Fantastical Gaps

Sound is an extraordinarily malleable and mercurial medium, it can represent any 'real' or imagined element; these elements can be layered, remaining distinct and identifiable, or congeal into some other 'whole'. It can creep in without being noticed or startle you, splinter into independent moving elements, and – crucially for screen music – can seamlessly transform from one element to another. Conversely an identical sonic event can have multiple narrative meanings. One particular opportunity for such transformations afforded to screen music is known as the *fantastical gap*[11]. This exploits sound's ability to move between, or coexist in, the diegetic (story) and non-diegetic (storytelling) worlds. Imagine a scene of a pianist practising, we see her fingers move and so the music is on-screen and part of the story. The scene cuts to a new place and time but the music continues with the subtle introduction of a string ensemble (felt by all but only noticed by the attentive musicians in the audience). The music has transformed from the story to the storytelling in one continuous motion. Conversely, a scene might start with music which we assume to be film scoring but is later revealed to be happening in the story – a move (or 'correction') from the storytelling to the story world. In such cases the boundary between the story-world (diegetic/diakosmic) and its presentation (the non-diegetic/paradosic) is not clearly delineated, but a liminal space, a porous and magical membrane. The nature, extent and complexity of such transformations are too numerous to list, but a handful of representative examples are presented below hinting at the vast diversity of cinematic practice. These show the liminality and continuous nature of the boundaries between the diegetic and non-diegetic as well as the blurred lines between other sound spaces be they the 'real and imaginary, objective, and subjective, high and low fidelity, visualised and acousmatic, music and foley.

- In *Elf* (2003) the protagonist (Buddy the Elf) is beguiled by a solo female vocal (his co-worker Jovie) emerging from the showers, he moves closer and starts harmonising undetected (even though their voices are perfectly balanced). A non-diegetic orchestra sneaks in to accompany the duet, subliminally expressing and foreshadowing their eventual union. When Jovie suddenly notices the harmoniser, her shock is represented by the hard cut of the orchestral accompaniment as we are brought back down to Earth for the comedic conclusion. This is an example of subtle 'hybrid-scoring' where in-story and 'presentational' music interact for narrative effect.
- In *The Amazing Spiderman 2* (2014), Peter Parker, in an anxious slump, puts on headphones to listen to a track to console himself, we hear the lyrics through the tinny headphones through the ambient space, "*to make you well…to make you well*". As the lyrics sink in, the music swells and increases in its *fidelity* occupying both the emotional function and sonic characteristics of non-diegetic film scoring. And as if to complete the fantastical transition, Peter – having found a new purpose – removes his headphones (the original source) and the music just continues to grow enveloping the scene's sonic and emotional space.
- In *The Shawshank Redemption* (1994), Andy Dufresne, having briefly commandeered the prison's PA system, places a record on the turntable and plays Mozart's *Canzonetta sull'aria* from *The*

---

[11] Proposed by Gorbman (1987) and further developed by the likes of Stillwell (2007) and Smith (2009)

Marriage of Figaro. At first, the aria begins as a diegetic sound, mono and constrained in frequency of the loudspeaker, this low fidelity is reinforced when he initially plays it through the tannoy system, reaching the ears of prisoners and guards alike. The camera cuts to inmates, frozen mid-step, their faces lifted toward the sky as though touched by a force beyond their grim reality. As the aria unfolds, its sonic texture subtly expands, transcending the diegetic source. The prison yard and the world within Shawshank are flooded with the rich, ethereal harmony of Mozart's duet, breaking the boundaries of the tannoy's mechanical fidelity. In that moment, the music becomes not just a sound but a voice of liberation, enveloping the scene with a swelling resonance lifting them beyond the prison walls, and of diegetic space itself.

- *Birdman* (2014) features a beguiling score, composed – in fact often improvised – by Antonio Sanchez on drum kit and layered percussion. Throughout the film, it is unclear whether this music lives in the score or in the streets around the action. So much so that in one continuous scene, the drums underscore a conversation between the two protagonists (Riggan and Mike), when the characters continue into the streets the score builds and we see the in-story source (in fact a stand-in for the composer) delivering the score 'live'. Later, Riggan in a deep psychosis on a building rooftop – summoning an earlier role as a superhero – says "cue the music" – which instigates the film score.

- As a segue to *Pulp Fiction*'s (1994) dramatic cold open, Dick Dale's iconic surf-rock track *Misirlou* plays over the opening credits operating as a form of theme music, at an arbitrary moment during the credits radio static cuts into the track and emerges into *Jungle Boogie* by Kool and the Gang. The black background dissolves into Vincent and Jules chatting in the car as the music drops in volume and fidelity indicating its source in the car radio. It's easy to miss this – it just has a general gritty and compelling feel – but unpicking the process (which the listener might do subconsciously) suggests that it's one of the two protagonists switched the radio station 'off–screen'. From what 'station'? One broadcasting in the restaurant perhaps? Even this attempt at justification is impossible, it turns out that the restaurant scene occurs *later* than the car conversation, but one musical sequence connects multiple locations in geographical, temporal, and diegetic spaces. The non-linear and disarming nature of the film's narrative is reflected in these strange transformations.

- In *Reservoir Dogs* (1992), *Stuck in the Middle with You* by Stealers Wheel is initially heard in mono through the radio in the background, when Mr. Blonde first turns it on during the tense torture scene. As the violence escalates, the music subtly increases in fidelity, volume, and stereo spread, moving from the diegetic to non-diegetic in terms of sonic characteristic and immersion, pulling the audience as accomplices to the brutality unfolding on screen. The lyrics "stuck in the middle with you" capture the ironic contrast, at **once** belonging to the familiar song and referencing the entrapment of the victim. The music has moved from the in-scene radio to *anempathetic* scoring – coldly indifferent to (if not psychopathically with) the brutal scene.

- In *Inception* (2010), "Non, Je Ne Regrette Rien" by Édith Piaf is used as a key device to blur the diegetic boundary, creating an unsettling and disorienting effect as the characters move through dream layers. Initially, the song plays as diegetic sound from a music box, subtly shifting in volume and distortion to indicate the dream world's collapse. As the layers of the dream deepen, the track's sound decreases in fidelity, becoming slower (and pitch), more warped, and enveloped by the film's score, while remaining logically 'in the scene' its role now is also to cue both the audience and the characters to the blurring of dream and reality.

- The *Bill and Ted Adventure* (1989) series have a repeated device, where in excitement they play air guitar gestures ("Excellent"). We hear distorted metal guitar 'twiddles'. What is this sound? It is visualized in the scene, but not actually sounded. We might call this 'implied' sound, or meta-scoring – the music they are hearing in their heads to which we are also privy, despite its non-acoustic nature. We are used to similar audio-visual cinematic devices, where the implied or potential sound inherent in an object or character - such as Janet Leigh's inner monologue in

Psycho (1960) looking at telephone wires and hearing the conversations, or the (historically reversed) radio broadcasts as we drift backwards from planet Earth in *Wall-E* (2008).

- In the opening of *Blazing Saddles* (1974), the use of big band jazz (specifically by the Count Basie Orchestra) is startling, as it immediately breaks from the traditional Western score that feels completely out of place in the dusty, rugged world of the frontier. The camera then pans to reveal the heightened absurdity of the scene: the music is not a film score but played in the story-world by the Count Basie Orchestra themselves who interact with the character. Suited up, with full band staging, in the middle of the desert, for no discernible reason whatsoever. Such a distracting, unconventional and absurd use of screen music, reminds us of the whole artifice of the cinematic process. The use of absurdist or deliberately inappropriate scoring and sound design might be termed *extradiegetic*[12] – the furthest boundary in the screen music landscape.

- In *The Truman Show* (1998), Truman Burbank's entire life has been secretly broadcast as a 24/7 reality show (again *The Truman Show*), with everyone around him—including his family and friends—being actors, while Truman remains unaware that his world is a massive, meticulously controlled television set, with directors, extras, and ad placements. As part of the show's plot points (and his 'real' life experience) Truman reunites his long-lost father. We witness the scene being live-broadcast for the (in-story) audience's entertainment. To heighten the emotion (in story) film scoring is cued and performed live, alongside the director's careful camera instructions. The layers are particularly porous here, the film scoring (in the 'film proper' for our benefit) actually *precede*s and entangles with the in-story film scoring; we feel the emotional resonance of Truman's authentic experience, as well as the in-story's audience reactions. To further entangle the layers, the actor playing the music for the scene is in fact the (higher-level) composer himself (Burkhard Dallwitz)[13]. Though we, as viewers, are aware of the orchestration behind the scenes, we still experience the same emotional significance as Truman, underlining the film's commentary on authenticity, control, and the power of media manipulation.

Several of these boundaries and devices are illustrated in Figure 3, which reveals a little of the fantastical complexity Screen music and sound affords in storytelling.
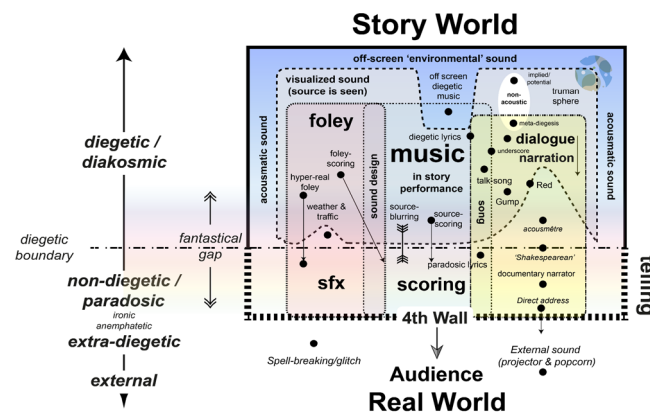

Figure 3: An impression of some of the many narrative devices of screen sound and music, blurred boundaries between the story-world (diegetic), and its telling (non-diegetic/paradosic); dialogue and forms of narration; sound effects, foley and composition and the fantastical gap between diegetic music and film scoring. This whole world can also be recreated within a film itself, indicated as a *Truman Sphere*.

---

[12] 'Extradiegetic' is also often used simply as a synonym for 'non-diegetic' - anything not in the story world (film scoring, a 'higher-level' narrator etc.). However here the distinction is that these elements actually challenge, break or otherwise disrupt the story-telling process itself. *Apodiegetic, dysdiegetic* or *antidiegetic* ('away from', 'impairing' or 'against' the story) suggest similar mechanisms would we need to invent a new term. The infamous end of *Monty Python and the Holy Grail* (1975) may fall into this category - where the dramatic final battle scene is stopped by the police in 'our' world, and we are left abandoned with chirpy organ 'hold' music.

[13] In another scene, the film's other composer, Philip Glass, can also be seen playing piano (and thus 'live-scoring') as Truman sleeps.

VI: Tearing Down the Walls

As complex and various as these screen music devices have become, and as ubiquitous as screens are, technology continues to expand, providing new opportunities for storytelling through music. Video games, whose stories are often non-linear, require these devices to adapt dynamically to the choices we make, indirectly allowing us to guide the score. In *Red Dead Redemption 2* (2018), diegetic music becomes immersive, with spatial realism as characters walk around the players. In *Grand Theft Auto V* (2013), we create our personal soundtracks by choosing in-game car radio stations. Some games take an even more direct approach: *Monument Valley* (2014) uses individual notes in a harmonic context to represent the sequential completion of puzzle sections, *Machinarium* (2009) requires players to replay on an in-game instrument a melody heard earlier to solve its final level, and *Returnal* (2021) organizes levels by musical keys, blending architectural and harmonic structures seamlessly. In *Guitar Hero* (2005)*, Rock Band* (2007) and countless other games, the successful completion of the score *is* the game.

Virtual reality has dissolved the walls yet further; we no longer need to be lured into the screen but can step right though it. We enter the story-world itself living among its characters. Our need to create and tell stories remains so strong it seems that we now make stories about storytelling itself, blurring the boundaries between creators, audiences, and the narrative worlds we inhabit.

© Professor Milton Mermikides, 2025

Glossary of Key Terms

- **Acousmêtre**: A term introduced by Michel Chion, describing a voice whose source is unseen. Unlike a detached narrator as in the documentary genre, an acousmêtre is part of the narrative 'story world'.

- **Diegetic/Nondiegetic Sound/Music**: Claudia Gorbman's framework where diegetic sound/music is heard within the story world by characters, while nondiegetic sound/music exists solely for the audience. I have suggested the terms **diakosmic** (i.e. from the *kosmos* – the 'story world') and **paradosic** (from the Greek to deliver) to represent similar distinctions between the story and its telling.
- **Extradiegetic:** A synonym for non-diegetic, or here distinguished as non-diegetic elements, such as music, sound, or dialogue, that momentarily break the narrative's internal logic or boundaries, creating a disruptive or self-aware interaction between the story and the audience. *Apodiegetic* (away from the story), *dysdiegetic* and *antidiegetic* – evoke similar 'spel-breaking' mechanisms.
- **Fantastical Gap**: A concept from Gorbman and refined by Stilwell, Smith and others, referring to moments where the line between diegetic and nondiegetic becomes ambiguous.
- **Gorbman's Rules**: Principles governing the use of film music, including invisibility, inaudibility, emotional signification, narrative cueing, continuity, and unity.
- **Intertitles**: a word or group of words (such as dialogue in a silent movie or information about a setting) that appear on-screen (that is non-diegetically/paradosically)
- **Leitmotif**: A recurring musical theme associated with a particular character, idea, or emotion, often used to enhance narrative cohesion.
- **Meta-Diegesis**: A narrative layer where sound or music represents a character's internal state, blending diegetic and nondiegetic elements.
- **Synchresis**: Michel Chion's term for the spontaneous mental fusion of sound and image when they occur together, regardless of their actual relationship.

References and Further Reading

Bibliography

Brown, T. (2012). *Breaking the Fourth Wall: Direct Address in the Cinema*. Edinburgh University Press.

Chion, M. (1994). *Audio-Vision: Sound on Screen*. Columbia University Press.

Didion, J. (1979). *The White Album*. Simon & Schuster.

Gorbman, C. (1987). *Unheard Melodies: Narrative Film Music*. Indiana University Press.

Gottschall, J. (2012). *The Storytelling Animal: How Stories Make Us Human*. Houghton Mifflin Harcourt.

Kalinak, K. (2010). *Film Music: A Very Short Introduction*. Oxford University Press.

Smith, J. (2009). Bridging the gap: Reconsidering the border between diegetic and nondiegetic music. *Music and the Moving Image, 2*(1), 1–23.

Stillwell, R. (2007). The fantastical gap: Representing sound and music in cinema. In R. Stilwell & P. Powrie (Eds.), *Composing for the Screen in Germany and the USA* (pp. 184–202). Indiana University Press.

Filmography

The Amazing Spider-Man 2 (2014) – Columbia Pictures / Sony Pictures

Baby Driver (2017) – TriStar Pictures / Sony Pictures

Better Call Saul (TV Series, started 2015) – Sony Pictures Television / AMC

Bill & Ted's Excellent Adventure (1989) – Orion Pictures

Birdman (2014) – Fox Searchlight Pictures

Blazing Saddles (1974) – Warner Bros. Pictures

Don Juan (1926) – Warner Bros. Pictures

Elf (2003) – New Line Cinema

Entr'acte (1924) – Société Générale des Films

Good Morning Vietnam (1987) – Touchstone Pictures

Immortal Beloved (1994) – Columbia Pictures

Inception (2010) – Warner Bros. Pictures

The Jazz Singer (1927) – Warner Bros. Pictures

L'Assassinat du duc de Guise (1908) – Pathé Frère

Monty Python and the Holy Grail (1975) – EMI

The Patsy (1964) – Paramount Pictures

Psycho (1960) – Paramount Pictures

The Purple Rose of Cairo (1985) – Orion Pictures

The Truman Show (1998) – Paramount Pictures

King Kong (1933) – RKO Pictures

Pulp Fiction (1994) – Miramax Films

Reservoir Dogs (1992) – Miramax Films

Saving Private Ryan (1998) – DreamWorks Pictures / Paramount Pictures

The Shawshank Redemption (1994) – Castle Rock Entertainment / Columbia Pictures

Trading Places (1983) - Paramount Pictures

Wall-E (2008) – Walt Disney Pictures / Pixar Animation Studios

Ludography

Guitar Hero (2005) – Harmonix / RedOctane

Grand Theft Auto V (2013) – Rockstar Games

Machinarium (2009) – Amanita Design

Monument Valley (2014) – ustwo games

Red Dead Redemption 2 (2018) – Rockstar Games

Returnal (2021) – Housemarque / Sony Interactive Entertainment

Rock Band (2007) – Harmonix / MTV Games