

# Unsettling Melodies: a Cognitive Approach to Incongruent Film Music

Steven Willemsen, Miklós Kiss

The University of Groningen (the Netherlands)  
E-mails: m.kiss@rug.nl, s.p.m.willemsen@gmail.com

“One cannot listen and look at the same time. If there is a dispute, sight, as the most developed, the most specialized, and the most generally popular sense, always wins. Music which attracts attention or the imitation of noises is simply disturbing.”

Jean Epstein (*Magnification*, 1977, 15)

**Abstract.** *Incongruent film music* is a soundtrack, either diegetic or non-diegetic, which expresses qualities that stand in contrast to the emotions evoked by the events seen. The present article aims at covering two interconnected areas; the first is comprised of a critical recapitulation of available theoretical accounts of incongruent film music, whilst the second part of the paper offers an alternative, embodied-cognitive explanation of the audio-visual conflict which arises from this particular type of incongruence. Rather than regarding it as a phenomenon that works through disrupting conventions, we stress a perceptual-cognitive reason behind incongruence’s emotional strangeness.

**Keywords:** film music, affective congruence, polarization, incongruence, cognitive-ecological theory.

The remarkable torture scene from Quentin Tarantino’s 1992 black comedy crime film *Reservoir Dogs* is a cinematic moment that has worked its way into the collective memory of moviegoers over the last twenty years. It depicts Michael Madsen’s gangster character, “Mr Blonde,” torturing a police officer that is being held hostage, seemingly for no other reason than for his own sadistic pleasure. When the scene culminates violently in Madsen’s cutting of the helpless cop’s right ear with a razorblade, this shocking event is not the only contributor to the scene’s idiosyncratic, memorable qualities. The clear and foregrounded presence of a “bubble gum” pop song, which happens to be on the radio, conspires with the visuals as an evenly disturbing choice. From the moment when Mr Blonde turns on the radio,

the song, Stealer's Wheel's *Stuck in the Middle with You*, immediately calls for attention. As the cop loses his ear and is doused in inflammable oil, the unbearably light song remains constantly present, adding an emotive quality that feels highly inappropriate to the horrific events it accompanies. It is precisely this phenomenon that is of our interest here. We will refer to it as "*incongruent film music*," a musical track in narrative film, either diegetic or non-diegetic, which expresses qualities that stand in sharp contrast to the emotions evoked by the events seen.

Conflicting soundtracks generate an emotional impact that is of a substantially different nature than the affect of traditional, congruent film scores. Yet, despite their often memorable status, these salient moments have not often been described as a specific affective device. The goal of this paper, first of all, is to contest the assumption that the striking effect of such music arises merely or mainly from its novelty and estrangement from film music's conventions of congruence. After all, if it were only for such music's unexpectedness or self-consciousness to create an audio-visual shock, the practice would already be out-dated given its presence in a wide array of (more or less "mainstream") films. Secondly, although the instances of musical incongruity may differ in some respects, we believe that there are overlapping resemblances in the emotional impact that these scenes attempt to invoke. Yet, rather than providing any typology, tracing a history of cases or investigating cultural origins and implications, this paper focuses on the perceptual-emotional conflict that incongruent film music forces us to endure. More precisely our aim is to analyse and explain how audio-visual incongruence works on our multi-modal cognitive processing, interpretation, and emotional evaluation. We propose that the particular affect produced in the emotional collision of music and visuals results from a distinct cognitive interplay. This interplay, as we argue, results from our (evolutionary) propensity to perceive and process cross-modally, combining sensory data from different senses pre-consciously. Looking at cognitive and ecological underpinnings, our approach will explain the practice and effect of incongruent film music that precedes any culturally influenced interpretation. This cognitive approach also enables us to define incongruent film music as a historically recurring strategy of audio-visual affect in its own right. Lastly, such a perspective helps to re-evaluate conventional assumptions on congruence that are implied in many earlier theoretical accounts on film music.

In different forms, incongruent film music can be found throughout the history of narrative film soundtracks,<sup>1</sup> in fact, some instances have become highly

---

1 By focusing on classical narrative film, we mean to exclude not only documentaries, but also, more importantly, other traditions of non-classical, artistic filmmaking.

prominent in narrative film history. One can think, for example, of scenes in the films of Stanley Kubrick (e.g. Vera Lynn's *We'll meet Again* in the final scene of *Dr. Strangelove* [1964], Malcolm McDowell's sadistic rendition of *Singing in the Rain* in *A Clockwork Orange* [1971]), or the violent climax of the baptism scene in Francis Ford Coppola's *The Godfather* [1972], when the church organ accompanies the assassinations). The practice can moreover be found recurrently in classical Hollywood movie scenes as a means of ironic comment. Musical-emotional incongruences also recur in more explorative narrative cinema, like in the films of David Lynch (such as the song *In Heaven*, performed by "the lady in the radiator" [Laurel Near/Peter Ivers] in *Eraserhead* [1977] or the playing of Roy Orbison's *In Dreams* in *Blue Velvet* [1986], which outrages Dennis Hopper's Frank Booth character), or the shockingly unexpected, and seemingly unmotivated, appearance of Naked City's *Bonehead* in the exposition of Michael Haneke's *Funny Games* (1997 and also 2007), to name just a few.

Yet, as we will prove shortly, traditional film music theory uncritically and persistently builds on assumptions of how, in order to be emotionally effective, classical narrative cinema's film music should remain congruent and subordinated to the images and the story. This is in accordance with the view that film techniques, such as editing, supposedly should remain "invisible" and subordinated to narrative action, and in this way, film music's role has been constrained to be that of an unnoticed emotive manipulator. According to this somewhat uncritical view, film music should remain at all times "unobtrusive" – to be "heard" only on a subconscious level. Conflicting music then, as these studies assume, could point directly to a film's artificiality and its manipulative intentions – a simple explanation which foregrounds non-diegetic music's fundamentally unnatural role in any story world.

One of the theoretical works that most enduringly represents this view is Claudia Gorbman's eloquently entitled *Unheard Melodies* (1987). Gorbman's widely acclaimed book is firmly rooted in the 1970s and '80s paradigm of psychoanalytical film theory and its reliance on conceptions of film viewing as a "suspension of disbelief." Jeff Smith, in his 1996 article *Unheard Melodies? A Critique of Psychoanalytic Theories of Film Music*, aptly shows how the idea of unobtrusiveness in film scoring has fundamentally shaped psychoanalytic

---

Experimental film, art cinema, and other avant-gardes have long since experimented with a broad range of incongruent elements, including sound and image. The reason for exclusion is that such films lack the mimetically evoked immersion and emotions that classical narrative films exploit, making them different objects of study that are (for now) outside of the scope of this paper.

film theory's accounts of film music. For Gorbman, film music has a crucial role in strengthening viewers' immersion and attaching them emotionally to the visually presented fiction. Crucially, she argues for a central position of the "unheardness" in film music's affective capabilities. According to Gorbman, the double functions of film music are to *semiotically* prevent viewers from finding unclear or ambiguous signification, as well as to *psychologically* enhance their immersion by means of film music's capacities to bridge visual discontinuities and propel the flow of events and narration (Gorbman 1987, 58–59).

The view proposed by Gorbman entails that it is only by means of congruence that the music in narrative film can enhance viewer immersion. After all, it is congruence that helps music to remain unobtrusive and supplemental to the narrative. As Gorbman notes, "were the subject to be aware (fully conscious) of its [that is the music's] presence as a part of the film's discourse, the game would be all over" (Gorbman 1987, 64). From this psychoanalytical point of view, congruency is not only regarded as the most common mode of film scoring (which it undoubtedly is), but is also deemed to be a necessary precondition in order for film music to be emotionally effective. However true this may be for the majority of film soundtracks, this explanatory framework is not able to unfold the emotional affect of incongruence.

Even though our viewing experiences prove that incongruent film music does not necessarily break immersion as radically as psychoanalytic accounts presume, its salience may put the viewer in a troubling emotional state or bring semiotic confusion. Whilst cheerful sounding tunes do not complement explicit violence either psychologically or semiotically, we can relate to our own confrontations with scenes in *Reservoir Dogs* or *The Godfather*, where neither displeasure nor disturbing awareness of obtrusive techniques blocked our emotional experience and narrative involvement. Although we were consciously *aware* of the collision taking place, this awareness did not divert but rather enhanced these scenes' emotional effects. In conclusion, we strongly feel that the somewhat logical, but still problematic stance of psychoanalytic theories on musical-emotional congruence provides some untenable assumptions in these cases.

Modes of incongruence in film music have been observed by many film theorists over the years. Scholars and critics have mostly noted a practice that is labelled by David Bordwell and Kristin Thompson as "*ironic contrasting*" (Bordwell & Thompson 2008, 302). Claudia Gorbman similarly describes how directors have often used (mostly diegetic) appearances of (mostly popular) songs to achieve a contrasting irony towards the accompanied narrative events (Gorbman 1987, 23–

24). Drawing on what Michel Chion called such music's "anempathetic" relation to the fiction (Chion 1994, 8–9), Gorbman notices a sense of irony arising from the music's unawareness or indifference to the unfolding dramatic developments. In this way the incongruent music comes to function as a kind of authorial ironic comment that expounds the narrative action.

Defining incongruent film music as ironic contrast allows for many examples – see the previously mentioned cases of films by Stanley Kubrick for instance. Yet, there are several arguments against reducing incongruent film music's manifold functions to being merely ironic. First of all, in his article *Popular Songs and Comic Allusions in Contemporary Cinema* (2001) Jeff Smith has rightly noted that in these cases irony functions more often linguistically than musically. To achieve an ironic effect, popular songs in film (either diegetic or non-diegetic) often create a play of meaning through bisociative implications of a song's title or lyrics (Smith 2001, 428). In such instances, the sense of irony is not primarily triggered by incongruent musical emotions, but mostly as a result of noticed lexical ambiguities – like indefinite pronouns in the songs' title or a lyrical shift of meaning, for example from figurative to literal (e.g. the re-contextualized meaning of being "stuck in the middle with you" during *Reservoir Dogs*' torture scene). Our second argument presents a perhaps more obvious point, according to which, not all incongruent film music is felt to be ironic. Although many examples provide clear emotional collisions between their musical and visual tracks, they do so without any trace of irony (see the light classical music that accompanies the assassinations in *The Godfather* or the explicit torture scene in Chan-wook Park's 2003 film *Oldboy*, which is countered with a light Vivaldi piece from *The Four Seasons*). Moreover, even when an ironical component seems to be present, "irony" may sometimes be hardly sufficient to fully cover the complex emotional impact of the audio-visual collision. In *Reservoir Dogs*, for example, the gruesomeness of the torture of cutting off someone's ear is simply too abject to hold a merely ironic relation to the music by use of some comic effect. Actually many viewers have experienced this scene as not being amusing at all and, to some extent, have even considered it revolting.<sup>2</sup> It seems to us that in these cases feelings of irony serve more as a "naturalization" of an emotional conflict. We will consider the sensation of irony to be one possible hermeneutic result from the felt conflict of incongruence, where the likelihood

2 Multiple accounts exist of viewers walking out of the first *Reservoir Dogs* screenings in 1992 because of this specific scene – see John Hartle's illustrative article in the *Seattle Times* (Hartle 1992).

of such a response depends on the viewer, the context of the narrative situation, and the nature of the music.

Aiming for a more thorough investigation related to functions of incongruent film music, we propose a shift in the explanatory perspective. The first key question is “how can we be, at the same time, *aware* of musical incongruence, as well as being emotionally *affected* and manipulated by it?” Although audiovisual incongruence can surely create a sense of confusion or disturbance, we argue that it does so *within* our immersive experience of the narrative world, and not by wholly disrupting it in some Brechtian fashion. As noted, much of the problematic points in psychoanalytic approaches to film music emerged from the discrepancy between the music’s ability to give clear cues and bits of narrative information on the one hand, and notions of its supposed “inaudibility” to avoid a rupture in the immersion on the other (Smith 1996). We agree with Smith that an appropriate theory should not reduce the viewer to a passive receptacle who treats all the presented emotions as its own, but acknowledge a dynamic relation between musical information and the viewer. In his article *Movie Music As Moving Music* (1999), Jeff Smith provided approaches to film music from a cognitive perspective, attempting to theorize the relation between film music and emotion while overcoming some of the problems of psychoanalytical methods (Smith 1999, 148). Drawing from the work of Smith, who builds on Peter Kivy, Joseph D. Anderson, and Annabel J. Cohen, our proposed framework brings together the findings of key researches in the field of film, music, and cognition. This will enable us to study film music on the fundamental levels of perception, cognition, and emotion, thereby bypassing the untenable theory that discards incongruence in psychoanalytic approaches.

According to Smith it is too often overlooked that all emotions are composed of both affective and cognitive components (Smith 1999, 155). Following Peter Kivy’s division between music as *expression* and music being *expressive of* (Kivy 1989, 12–26), i.e., between primary feeling and recognizing emotions in music, the distinction is between emotions that are *evoked* by a scene and its accompanying music (e.g. creating empathy, offering the spectator emotional *arousal*), and emotional components that are *communicated* by film music (e.g. indicating a character’s emotional state, thus influencing the viewer’s *judgement*). Affect and judgement “each may take precedence over the other at different moments to produce a range of possible responses” (Smith 1999, 156). Because of its reliance on psychological immersion, psychoanalytic theory tends to overlook these dynamics of communication, presuming that viewers claim all associated

emotions as their own. We can note that in the case of radical audio-visual incongruence, music never seems to directly *arouse* the emotions expressed through its own musical qualities. What is more, cheerful music accompanying visual brutalities does not make us *feel* more cheerful (on the contrary).

Nevertheless, we cannot ignore the questions raised by psychoanalytic theory regarding audio-visual incongruence and immersion. As Gorbman presumes, one would think that in highly conflicting situations viewers would instantly decide to discard the accompanying music as incongruent, conflicting, or simply mismatched in order to consciously marginalize its emotional impact. Yet, we can note that film viewers still tend to accept even clearly conflicting music as referential to the images. How else for instance could *Oldboy*'s non-diegetic Vivaldi piece affect us at all? At this point it is crucial to ask what compels us to accept such music as referential to the images, without harming our immersion in the fiction and our (partial) belief in its realism.

Convention has trained viewers to know and feel that musical accompaniments tell us something useful about how to interpret and evaluate narrative situations. Our reliance on film music's "usefulness" is partly habitual. Yet we must also note that this dependence too has a strong cognitive-ecological origin. In his 1996 book *The Reality of Illusion: An Ecological Approach to Cognitive Film Theory* Joseph D. Anderson reminds us of this aspect: "we are programmed by evolution to check and cross-check our perceptions multi-modally. [...] The simple addition of musical accompaniment provides a second modality against which to check our impressions and provides confirmation on at least two levels" (Anderson 1998, 87). But why should we do this when the musical addition is clearly "false" and has seemingly no motivated role in the diegesis? Building on Lawrence Marks's (1978) multimodal theory, Anderson overcomes this question by arguing for an evolutionarily evolved cognitive process that binds sound and visual together, thus creates a percept of a single event (Anderson 1998, 86). On a basic cognitive level the sounds and images are perceived by film viewers as coming from *one* source – a single instant generating our perceptions. This is not strange if one considers the idea from an evolutionary perspective. The real world's natural environment never offers false sounds or artificially added music, and thus, ecologically speaking, anything we *hear* must somehow be part of the same situation we can also *see*. Detecting synchrony is a crucial part of our experience and perception of everyday reality.<sup>3</sup> In any case of film sound, synchrony masks

3 Synchrony alone is sufficient for the human processing system in linking the auditory and visual elements as one event. Multiple researches show a fixed tendency or even

the “true” non-diegetic source of the sound. When looking at a film, viewers are ecologically, even evolutionarily, biased in seeking for synchrony as a means to connect sound and image.

The presumption that all information originates from one coherent source can easily lead film viewers to search for patterns and correlations for bridging the modalities. After all, to unify our perceptions into something coherent, we must distil and interpret an unambiguous meaning from the multi-modal flows of information. With film music we favour congruence, as an inclination towards seamless connection being our basic perceptual disposition. We would argue that this bias of a single generating instant is what makes us readily ascribe the music, in terms of emotional or narrative information, to whatever is happening onscreen. For this reason, music and narrative images tend to conflate on a pre-conscious, or “unheard,” level.<sup>4</sup> Music cognitivist Annabel J. Cohen has similarly suggested that in the cognitive selection and determination of narratively “useful” information, the semantic meaning of the musical track is detached from its acoustical source (Cohen 2001, 262). The music’s semantic components are deemed useful as a source of relevant information and are processed separately in terms of their assistance to understanding what is seen. The acoustical dimensions, on the other hand, generally remain consciously unattended by cognition (although they do steer our visual perception in certain directions, for example, through temporal correlations). We suspect that immersion-breaking audio-visual collisions may be prone to emerge primarily when the incongruence concerns the music’s fundamental acoustic or temporal elements. For example this may be at the heart of our estrangement when large deviations in rhythm or clearly flawed synchrony occur, as happens in Jean-Luc Godard’s asynchronous or abruptly pausing soundtracks.<sup>5</sup> Verifying this claim would nevertheless require deeper, preferably empirical, research.

Everything considered, we assume that the binding of sound and image achieves an interpretational influence before it would be consciously evaluated as disruptively mismatched. Nevertheless, as Anderson too notes, following the

---

necessity of recognizing and attributing synchrony that is already present in very low-level, bottom-up, processes of multi-modal perception (Spelke 1979, Hummel & Biederman 1990, Revonsuo 1999).

4 Presumably, it can only be on the basis of such traits and “flaws” in human cognition that film music has been able to gain its status as an almost unquestioned realistic effect in the cinema, despite its lack of mimetic abilities that we would objectively consider “realistic.”

5 See for example the recurrently interrupted extra-diegetic music in *Pierrot Le Fou* (1965) or the temporal disjunctions of sound and image in *La Chinoise* (1967), *Weekend* (1967), and *Passion* (1980).

initial comparing of rhythm, meter and temporal congruencies with the visual, “there is comparison of the tone and emotion of the sounds and/or music with the event unfolding on the screen, perhaps the very emotion of the music confirming or denying the validity of the viewer’s response to what is seen” (Anderson 1998, 87). For incongruent film music, this emotional response becomes the most significant, bringing about the question: how does incongruent music influence and impact viewers’ visual perception and emotional evaluation?

There is no need for theoretical explanations to realize that it is always the music that exerts a great deal of influence over the emotional qualities of an image, whereas the visual content is hardly able to alter the emotional affect in the music. Music’s emotive meaning is thus constant; that of the visual track is relatively unstable and changeable. This insight is at the core of the two cognitive processes, *affective congruence* and *polarization*, that Jeff Smith (1999) distinguishes as the two overarching affective functions of film music. In the process of *affective congruence* the music emotionally matches the visual and narrative events, thereby heightening the viewer’s emotional experience. When the emotional and semantic qualities of the visual and narrative events are largely similar to those expressed by the music, they reinforce each other, possibly leading to a stronger arousal of emotions in the viewer. During *polarization*, on the other hand, the emotive qualities of music and the visual are not unambiguously matched – i.e. they do not express exactly the same emotional meaning. This is not to be confused with the incongruence we are concerned with in this article. As for polarization one should think of a slight discrepancy, or perhaps an emotionally “neutral” image accompanied by music that is expressive of some emotion. This audio-visual constellation leads to an interaction in which “the affective meaning of music moves the content of the image towards the character of that music” (Smith 1999, 160). Annabel J. Cohen has conducted various experiments on this process, for example, by testing the effects of different accompanying music on the evaluation of short animated films presenting moving geometric figures (Cohen & Marshall 1988, Cohen 1993). In these cases the emotive qualities that viewers ascribed to the film were moved towards the qualities expressed by the music. In this sense film music functions as a semiotic signifier, guiding viewers’ judgements. Cohen explains how the music guides us through the interpretation of an image, directing viewers’ attention (since temporal congruencies tend to steer viewers’ attention), consequently altering the perceived meaning. As Jeff Smith concludes, auditory elements systematically shape the denotative and affective meanings of the visual (Smith 1999, 161).

Incongruent cases seem to disregard this general rule. Even though incongruent music may feel disorienting, unsettling, or ironical, when it stands in clear conflict with the visual, music does not move the visually perceived towards its own contrasting emotional qualities. It is likely that whenever the audio-visual incongruence becomes radical, the natural processes of congruency and its corresponding modes of affect are suspended. Empirical findings back up this assumption. For example Annabel J. Cohen noted that when music and visuals are in clear conflict, the musical affect is marginalized and the visual information tends to take precedence over the audio (Cohen & Marshall 1988). Although this only describes a tendency and not a fixed mechanism, it is assumable that when primary, cognitively impenetrable, tests of audio-visual correlation fail, viewers generally tend to rely on the visual information over the audio.<sup>6</sup> This could be an underlying reason of why violent images are specifically effective in creating and maintaining audio-visual incongruence: violent images are viscerally strong and emotionally unambiguous hence capable of resisting the interpretational polarization of the accompanying music. The lack of these effects demonstrates that audio-visual incongruence is different from Smith's category of polarization.

But what functions remain for the conflicting music then? Anderson notes that "If musical and visual information are in conflict in any of these instances [i.e. in *synchrony*, *rhythm*, or *emotion*], the conflict will force the viewer to go back and re-evaluate earlier reactions, to reinterpret the patterns and the significance of the filmic events" (Anderson 1998, 87). Yet, re-evaluation does not seem to be at the core of incongruent film music. While watching *Oldboy*, we do not re-evaluate or reinterpret the whole film, nor the particular scene of torture because of the incongruent music. Anderson's statement seems to be directed more at the narrative implications, rather than at any emotional effect. Marilyn G. Boltz (2004) offers another suggestion, following empirical tests on film music and cognition in terms of our interpretation, emotional affect and memory. The results of Boltz's experiments with viewers watching short clips paired to diverse musical soundtracks show that mood-congruent pairs of film and music are jointly encoded, leading to an integrated memory code. This encoding underlines our argument on how film and music tend to be coupled strongly in cognition.

---

6 In cases of narrative films, on the one hand, the primacy of visual information over the audio could include, even be reinforced by, viewers' *narrative* interest. On the other, *ecological* point of view, visual information's dominance could be easily justified from an evolutionary perspective, as it could refer to survival skills, which dodge those very rare situations, when nature's voices aren't reliable to the dangers of the – seen – environment.

“A joint encoding,” according to Boltz, “may be more likely to occur in cases of mood congruency in which musical affect can direct viewers’ attending toward those aspects of a film with a similar connotative meaning and thereby integrate music and film into one coherent framework” (Boltz 2004, 1196). As for cases of incongruence, Boltz’s findings suggest that viewers tend to perceive conflicting film music as separated from the visuals. After all, “the emotive meaning of [incongruent] music conflicts with that of the film, so that it is not always clear where attending should be directed or how the conflict of information can be resolved within one interpretative framework. Given that music and film seem relatively dissociated from each other in this situation, each may be encoded independently of the other” (ibid). However, we are bound by justice to say that this finding refers only to a tendency measured in short clips viewed under rather unnatural circumstances. Even Boltz herself notes that in an immersive, full-length narrative fiction film, the outcomes may be different: “On an experiential level, the use of ironic contrast often seems to result in a vivid memory of the film information. For example, ironically contrasted scenes from *A Clockwork Orange* are very memorable, perhaps even more so than they would be if they had originally been accompanied by mood-congruent music” (Boltz 2004, 1203).

Concluding the above, it is not some separate encoding or a disjunction in memory representation that makes incongruent film music an emotionally powerful audio-visual strategy. Nevertheless, there is a key notion in Boltz’s statement that, when film music remains highly incongruent, it is not clear for viewers *where* to direct their attention or *how* to solve the emotional problems within a single interpretative framework. In relation to Boltz’s conclusion, Anderson’s remark, originally concerning silent films, may be applicable to incongruence too: “Consider the possibility that the absence of the opportunity to confirm our perceptions cross-modally might account for our discomfort in viewing a silent film without accompaniment. If we are programmed by evolution to check and cross-check our perceptions multi-modally, the inability to do so might well make us fundamentally, vaguely uneasy” (Anderson 1998, 87).

Crucially, the cognitive accounts of film music – whether it is Smith’s affective congruence and polarization, Cohen’s and Boltz’s experiments or Anderson’s ecological considerations on multi-modality – all point to the same processes of viewers seeking to interpret in favour of congruence. Essentially, these processes are nothing but economically driven attempts to bridge the gap of multimodal information in order to create a single, clear, cognitively consonant meaning. With incongruent film music – although the possibility to cross-check our perceptions

and interpretation is essentially given – these very processes fail on us as the audio-visual discrepancy blocks our shortcutting meaning-making routine. Although the music presents itself as a source of information (relying on our cognitive and habitual disposition), incongruent film music obstructs the process of cross-modal testing, since the affective qualities of information offered are in unanticipated conflict. We would argue that in the clearest cases of incongruent film music – that is when no affective polarization takes place or the conflict can not be inferentially resolved in terms of simple ironic allusions – the felt result follows from the evasion of the processes of correlative interpretation. The role of incongruent music is therefore one of a deliberate disturber of unambiguous meaning making and a clear-cut judgement. Sonic perception's failure to confirm or even disconfirm visual perceptions creates some vague uneasiness, an unsettling feeling, similar to the one as described by Anderson concerning silent cinema's soundless viewing experience. Due to audio-visual incongruence, which makes finding correlation between information from multimodal sources difficult if not impossible, a thwarted perception of a scene results in emotional uneasiness.

On the basis of these findings and sources, our paper concludes in a three-step explanation on the elemental functioning of incongruent film music. Note that this is not a definitive description of all instances, but rather principles that generally seem to surround cases of musical incongruence.

1.) *The music is presented as an aspect of the visual.* We argue that it is on the basis of both acquired convention and our cognitive disposition that one readily and pre-consciously ascribes emotional qualities in music to images. Following this viewer bias, filmmakers further justify incongruent music's place by means of stylistic devices. The most common is a method, through which incongruent music gets some diegetic validation. A good example is a diegetic song, only "accidentally" played on the radio, appearing at an inappropriate moment, as seen in *Reservoir Dogs*. Incongruence's conflict may become less jarring by calculated editing that creates temporal congruencies and correlations between music and visual. This can be achieved through rhythmic scene pacing (as can be seen in the case of *Oldboy*) or by smart crosscutting (as happens in *The Godfather*, where the church organ music originates from the baptism scene, yet is increased in volume after each cut to the assassinations). Such stylistic tricks create a stronger and more convincing perceptual and evaluative connection between the emotionally colliding audio and video tracks.

2.) *Our cognition attempts to bridge the multi-modal gaps.* It is when we try to bring together two contradictory flows of input, the same way we automatically

do in cases of congruence (cf. Anderson's single event perception), the emotional conflict becomes tangible. It becomes salient because the emotional gap between visual events and contrasting music holds emotional implications that are too strong to be seamlessly naturalized.

3.) *The conflict is upheld or even climaxed.* Incongruent film music disturbs unceasingly our emotional response to the given scene. We tend to feel somewhat "uneasy" because it seems as if our senses are letting us down. Confronted with emotionally contradictory flows of information, one cannot apply a single interpretative framework to the unsettling experience. This is the moment where the sense of irony or comedy may arise, since such feelings result precisely from a sustained or sudden incongruence. Incongruence's uneasy experience may necessitate further viewer response, such as different inferential strategies and interpretations that serve to naturalize the conflict, for instance ascribing it to an (implied) author persona or treating the conflict as symbolical.

It seems that when incongruent music is not guiding us towards a singular emotional response (through an inferential shift producing "a judgement" – cf. Smith), in order to resolve what exactly is happening, the situation requires the viewer to focus more on the visual aspects of the actual narrative event. In this case, following an ecologically determined trajectory (cf. Cohen & Marshall 1988), the viewer is exerting an inferential effort to derive meaning from the visual track *only*. Thus, ultimately and somewhat paradoxically, incongruent music may emphasize and strengthen the power of the image on its own.

## Conclusion

This preliminary explanation is only a first attempt to connect all the different modes of incongruence and the cognitive-emotional responses they evoke. We can conclude that although psychoanalytic theorists were right in claiming that incongruent film music creates "uncertainty," they fell short in explaining the exact nature as well as the emotional consequences of such a complex audio-visual stimulus. A mere conflict between inconsistent semiotic interpretations cannot fully explain why incongruent film music recurrently invokes feelings of uneasiness. Rather than regarding it as a phenomenon that works through disrupting conventions, we stressed a perceptual-cognitive reason that ensures incongruence's emotional strangeness. The consequence of this conclusion is that the effect of incongruent music cannot be fully habitualized through overuse of the practice. It even seems that the previously described unsettling impact is

at the heart of all radically incongruent cases, even if they are considered comic or ironic. Other possible emotional and intellectual responses, we would argue, arise from this initial uneasy sensation.<sup>7</sup>

## References

- Anderson, Joseph D. 1998 [1996]. *The Reality of Illusion: An Ecological Approach to Cognitive Film Theory*. Carbondale: Southern Illinois University Press.
- Boltz, Marilyn G. 2004. The Cognitive Processing of Film and Musical Soundtracks. *Memory & Cognition* vol. 32 no. 7: 1194–1205.
- Bordwell, David & Kristin Thompson. 2008. *Film Art: an Introduction*. Boston: McGraw-Hill.
- Chion, Michel 1994 [1990]. *Audio-Vision: Sound on Screens*. New York: Columbia University Press.
- Cohen, Annabel J. and Sandra K. Marshall 1988. Effects of Musical Soundtrack on Attitudes toward Animated Geometric Figures. *Music Perception* vol. 6 no. 1: 95–113.
- Cohen, Annabel J. 1993. Associationism and Musical Soundtrack Phenomena. *Contemporary Music Review* vol. 9: 163–178.
- Cohen, Annabel J. 2000. Film Music: Perspectives from the Cognitive Psychology. In *Music and Cinema*, eds. J. Buhler, C. Flinn & D. Neumeyer, 360–377. Hanover: Wesleyan University Press.
- Cohen, Annabel J. 2001. Music as a Source of Emotion in Film. In *Music and Emotion: Theory and Research*, eds. P. N. Juslin and J.A. Sloboda, 249–274. Oxford: Oxford University Press.
- Flinn, Caryl. 1992. *Strains of Utopia: Gender, Nostalgia and Hollywood Film Music*. Princeton: Princeton University Press.
- Gorbman, Claudia. 1987. *Unheard Melodies: Narrative Film Music*. London: British Film Institute.

7 Further research on incongruence could apply this hypothesis to more specific scenes, testing how it may give rise to different felt and understood types of meaning. Other research may be directed to the question of whether immersion and film music's naturalness are as unstable as psychoanalytical theory implies. It seems that more severe disruptions are required to really break the audience's absorption in the fiction. Furthermore, it could prove interesting to see whether the differences between perceiving diegetic and non-diegetic music, related to their incongruent use, are as different as assumed. Although the difference between the two is very clear on an inferential level, they may just function differently while underscoring incongruently, since we essentially deal with non-diegetic film music, either perfectly congruent or highly conflicting, as an informative aspect of the diegesis.

- Hartle, John. 1992. 'Dogs' Gets Walkouts and Raves. *The Seattle Times*. October 29. <http://community.seattletimes.nwsource.com/archive/?date=19921029&slug=1521437>. Last accessed at 20. 09. 2013.
- Hummel John and Irving Biederman. 1990. Dynamic Binding: A Basis for the Representation of Shape by Neural Networks. In *The Twelfth Annual Conference of the Cognitive Science Society*, ed. M. P. Palmarini, 614–621. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kivy, Peter. 1989. *Sound Sentiment: An Essay on the Musical Emotions*. Philadelphia: Temple University Press.
- Liebman, Stuart and Jean Epstein. 1977. Magnification and Other Writings. *October*. vol. 3 (Spring): 9–25.
- Marks, Lawrence E. 1978. *The Unity of the Senses: Interrelations Among the Modalities*. New York: Academic Press.
- Powrie, Phil & Robynn Stillwell 2006. *Changing Tunes: The Use of Pre-existing Music in Film*. Aldershot: Ashgate.
- Revonsuo, Antti. 1999. Binding and the Phenomenal Unity of Consciousness. *Consciousness and Cognition* 8 (2): 173–185.
- Smith, Jeff. 1996. Unheard Melodies? A Critique of Psychoanalytic Theories of Film Music. In *Post-Theory: Reconstructing Film Studies*, eds. D. Bordwell & N. Carroll, 230–247. Madison: University of Wisconsin Press.
- Smith, Jeff. 1999. Movie Music as Moving Music: Emotion, Cognition and the Film Score. In *Passionate Views: film, cognition and emotion*, eds. C. Plantinga & G. M. Smith, 146–167. Baltimore: The Johns Hopkins University Press.
- Smith, Jeff. 2001. Popular Songs and Comic Allusions in Contemporary Cinema. In *Soundtrack Available: Essays on Film and Popular Music*, eds. P. Robertson Wojcik & A. Knight, 407–432. Durham: Duke University Press.
- Spelke, Elizabeth S. 1979. Exploring Audible and Visible Events in Infancy. In *Perception and Its Development*, ed. A. Pick, Hillsdale, 221–235. NJ: Lawrence Erlbaum Associates.