White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/115828/

Version: Accepted Version

Article:
Ireland, DI (2017) Great Expectations? The Changing Role of Audiovisual Incongruence in Contemporary Multimedia. Music and the Moving Image, 10 (3 (Fall 2017)). pp. 21-35. ISSN: 2167-8464

https://doi.org/10.5406/musimoviimag.10.3.0021

© 2017, owner. This is an author produced version of a paper published in Music and the Moving Image. Uploaded in accordance with the publisher's self-archiving policy.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Great expectations? The changing role of audiovisual incongruence in contemporary multimedia

Abstract: Film-music combinations that have been labeled as inappropriate, misfitting, or incongruent are often also described as unexpected audiovisual pairings. Various strands of academic research observe a prevalence of such constructions in contemporary multimedia, which arguably implies that such pairings are less surprising or unexpected than they once might have been. This article: identifies three types of audiovisual incongruence from recent multimedia; and discusses these in relation to psychological theories of expectation and ideas from semiotics, which facilitate consideration of any potential disjunction between authorial intent and perceiver reception of a work.

Discussion about the fit and/or perceived appropriateness between the musical and visual components of film dates back to cinema's early years. Examples of such discussion can be found in: Eisenstein's writings on parallelism and counterpoint;[1] and in critics' responses to the practice of "funning", in which performers would select popular songs to use in their live film music because of comic allusions provided by these songs' titles or lyrics, rather than their sonic or musical properties.[2] Discussion of more recent film music that might be labeled as "misfitting" or "inappropriate" often centers on the use of classical music or popular song during scenes of brutal violence, as in the works of *mélomane* directors like Quentin Tarantino.[3]

Conceptualizing all of these constructions as *incongruent* provides a helpful framework for understanding their potential impact on an audience. Drawing on the approach of music psychologist Marilyn Boltz who refers to a 'lack of common stimulus properties between music and film' in mood-incongruent audiovisual pairings,[4] incongruence is defined in the present article as the identification of a lack of shared properties in an audiovisual relationship. This definition is beneficial for various reasons. To some extent the focus on the *lack of shared properties* minimizes more loaded ideas of value that labels which are synonymous with "congruence" and "incongruence" (such as "fit" and "misfit", or "appropriate" and "inappropriate") may connote: instead, this

definition emphasizes the component parts of the audiovisual relationship rather than holistic, subjective judgments about the nature of that relationship (which may of course be influenced by judgments about the perceived congruity between these component parts that comprise the pairing). As such, this approach complements theories that challenge notions of autonomy in the images or audio track: such approaches instead recognize the active contribution of audio-visual difference in the construction of filmic meaning, regardless of whether that difference is complementary or contesting to cite Nicholas Cook's labels.[5] This approach also complements research which recognizes that incongruities can be perceived on various structural, semantic or holistic levels in an audiovisual relationship,[6] given that the present definition does not specify the exact dimensions of the audiovisual relationship on which a lack of shared properties might be identified in order for a perceiver to label the pairing as 'incongruent'. Consequently, this definition highlights that the previously cited Tarantinoesque pairings of violence and popular song or classical music represent just one type of incongruence, albeit a frequently discussed one that arguably focuses on notions of semantic and emotional difference.

Considering such moments using a central concept (incongruence) from research in the psychology of music in multimedia, rather than synonymous terminology from film music studies, is also important for recognizing the impact of audio-visual difference on: perception; subsequent interpretation of meaning; and emotional and aesthetic response. Empirical research suggests that a lack of shared properties in a film-music relationship can result in independent (rather than joint) encoding and memory representation of auditory and visual information,[7] which likely contributes to the memorability and salience of such moments. In this article the term "perceptual space" is used to refer to the opportunity to further consider the audio-visual components that comprise a multimedia text and the nature of their combination, which can be facilitated via this separate encoding of sonic and visual information.

Ideas of expectations (often the violation of expectations) also recur in discussions about film music that has been described as incongruent or misfitting. For example, discussing Hannibal Lecter listening to Bach whilst mutilating his prison guards in *Silence of the Lambs* (Demme, 1991), Stan Link

comments that such audiovisual relationships 'highlight[…] our expectations by thwarting or negating them. The fact that Bach forms an uncomfortable fit makes us very aware of having anticipated something else'.[8] Empirical research also points towards this potential perceptual influence of audiovisual incongruence. Boltz, Schulkind and Kantra found that experiment participants tended to remember filmic events better when a scene was accompanied by mood-congruent music. However, if the music foreshadowed the scene then the participants recalled more when the music had a mood-incongruent relationship with the scene: the researchers attributed this to the impact of the violation of expectations that had been shaped by the music.[9]

However, whilst independent encoding of auditory and visual information, and violated expectations might point to the memorability and salience of certain incongruent film music, other iterations of this trope are not necessarily noticed or remembered, dependent on where in the audiovisual construction attentional resources are focused.[10] Moreover, incongruent audiovisual pairings are not always surprising or unexpected given the trope's frequent use in contemporary multimedia: indeed, the website tvtropes.com has an entire section devoted to the synonymous idea of soundtrack dissonance.[11] To return to the earlier example, Kathryn Kalinak highlights that 'it has become […] commonplace to accompany violence with lighthearted music'.[12] Indeed, as film theorists Willemsen and Kiss note, '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'.[13]

Thus, to explain the effects of incongruent music purely as a result of violated expectations is a problematic simplification: expectations play various roles in the experience and perception of incongruent music in audiovisual media, not least given that such devices are arguably more frequently employed than they once might have been.[14] Elizabeth Margulis' article on expectations and music listening emphasizes that there are various types of expectation.[15] She refers to: expectations that might be more sensory or intellectual in nature; expectations that might be prospective or retrospective, as in when an object provokes expectations about what will follow or respectively may violate

previously-established expectations; and the distinction between a general state of expectancy and specific expectations. Margulis observes that music theorists often use the label "unexpected" variably and appeals for greater specificity, suggesting this may be achieved by detailing the origin, nature, time course, object and consequences of musical expectations to distinguish how these vary.

The same level of rigor is necessary to account for the complex, multifaceted, and evolving relationship between different types of audiovisual incongruence and perceivers' expectations. To explore these relationships further, this article will discuss three types of incongruence from recent multimedia. These examples, and the questions that they raise, provide talking points that illustrate one suggested approach toward theorizing contemporary forms of, and experiences with, audiovisual incongruence. This approach will primarily draw on: David Huron's psychological theory of musical expectations;[16] and the incongruent perspective, a psycho-semiotic approach toward studying film-music incongruence.[17] The incongruent perspective builds on the idea of incongruence as a lack of shared properties in the audiovisual relationship and contends that to more holistically understand such informationally complex constructions, analysis should recognize the perceptual impact of audio-visual difference. Such analysis should also recognize the concerns reflected by semiotic approaches that facilitate textual analysis *and* account for the positions of a text's producers and consumers. Huron's work sits comfortably alongside this approach given the emphasis that it places on perceptual processes, and evaluative and emotional response. Huron's ITPRA model outlines five 'expectation-related emotion response systems', each with 'functionally distinct neuropsychological systems':[18] these are imagination, tension, prediction, reaction, and appraisal responses, and will be explained in greater detail as they relate to the following discussion. This ITPRA model and its underpinning theory emphasizes that expectations relate not only to what might occur, but also predictions about when future occurrences might take place. Huron highlights the evolutionary benefits of such expectations, which help prepare an individual and conserve their attentional and physiological resources to best respond to situations. His work also recognizes the role of culture and prior experience in shaping such expectations,[19] a quality particularly pertinent

for application to aesthetic objects such as film and music. For these reasons, Huron's theory is particularly appropriate to apply alongside the incongruent perspective to consider the examples discussed in this article.

These examples represent three prevalent types of contemporary audiovisual incongruence from a range of multimedia. They represent incongruence as a form of parody, a marker of authorial style, and a means of participation. As such, as a set, these examples provide a range of talking points to consider the various relationships between expectations and audiovisual incongruence: as a group they point towards the ways in which such constructions are not necessarily reliant on shocking the audience but can in fact draw on sophisticated levels of cultural and intertextual understanding. Through discussing these examples, the aims of this article are two-fold: firstly, to demonstrate the benefits of using a psycho-semiotic perspective that can more holistically represent the audience member's potential perceptual, emotional and interpretive responses when analyzing the complex relationship between audiovisual incongruence and perceiver expectations; and secondly, like Margulis, to provide greater specificity by offering some initial tentative steps towards identifying some prevalent forms of audiovisual incongruence in contemporary multimedia and discussing how these might operate in relation to audience expectations.

Example #1: *Modern Family* and incongruence as parody

The first example is taken from the final episode of the first season of ABC's TV series *Modern Family.* The episode, entitled 'Family Portrait', was first broadcast in 2010. In the sequence in question, the character Mitchell is at home with his infant daughter, Lily, and is terrified to discover a pigeon in the living room. Mitchell unsuccessfully attempts to garner help by telephoning his partner Cam, a musician, who is away from the house and about to sing at a wedding, and thus unavailable to come to his aid. Accordingly, Mitchell decides to tackle the bird, who is now positioned in front of the closed door to Lily's room, so that he can reach and comfort his crying child. In the resulting sequence, images of Cam singing Schubert's 'Ave Maria' accompanied by solo cello are intercut with slow-motion footage of Mitchell destroying the contents of the house as he tries to

catch the pigeon. A screaming Mitchell is initially shown chasing the pigeon towards the camera brandishing a sports racket (see Fig. 1). He smashes vases and photograph frames by swiping the racket and throwing fire logs at the bird. He is also shown tearing a cushion, resulting in feathers flying everywhere, and chasing after the bird whilst discharging the contents of a dry powder fire extinguisher, before finally collapsing to the floor. Throughout, diegetic sound in the house is muted and Cam's performance of 'Ave Maria' dominates the soundtrack, which may be considered as an example of spatially displaced diegetic music when presented against the images of Mitchell in the house.[20]

Fig 1. Mitchell and the pigeon (*Modern family* episode 1.24, 2010)

The slow-motion presentation of the destruction of the house and the flying pigeon could be described as fitting or congruent with the steady tempo of 'Ave Maria'. Moreover, the fact that Cam's singing is diegetic (at least in his narrative space) arguably justifies the music's presence in the sequence given its origin in the narrative world of the characters. The cries of Lily and Mitchell before the sequence and Mitchell's muted screams at points within the sequence itself could also be interpreted as thematically linked to Cam's vocal performance.[21] Given our tendencies to search for patterning to facilitate perception, a point that will be returned to in greater detail later in this article, it is important to note such moments of congruity in the sequence when

considering matters of incongruence. However, moments of direct synchronization between Mitchell and Cam's mouth movements, and between the music and the images more generally are not presented in a particularly sustained manner. Thus, despite these localized moments of congruity, the semantic differences presented between the tranquil music and the images of destruction in the house may be of greater salience to the audience. Indeed, research by Scott Lipscomb suggests that in informationally complex stimuli such as this sequence, association judgments surrounding the appropriateness of a film-music pairing can acquire greater perceptual influence than moments of audio-visual synchronicity.[22] Equally, the pairing of such semantically different foregrounded classical music, be that instrumental or vocal works, with stylized scenes of devastation and destruction is, as already noted, not unusual in contemporary film and television. Thus, treating this sequence as a moment of incongruence provides rich opportunity to consider how potential perceiver familiarity with similar multimedia constructions might influence interpretation and response to the audio-visual differences presented here.

Huron's work draws heavily on the idea of schemas, expectation sets - 'encapsulated behavioral or perceptual model[s] that pertain[…] to some situation or context' to aid perception and appraisal of that situation.[23] In relation to music, he identifies concepts such as tonality and genre as schemas that aid perception and understanding of a work, and notes that learned associations can aid the acquisition of new schemas. Given its common usage, audiovisual incongruence of the type featured in this *Modern Family* sequence could be considered as a form of schema. Indeed, there is precedence in film music psychology research for using schemas to understand the influence of music on film perception. Empirical work by Boltz used schemas to explain the influence of music on participants' expectations surrounding the unfolding of an ambiguous scene and their memories of its content. Participants' perception and memory recollections were biased by the music in a mood-congruent fashion: positive music caused them to remember, and misremember, more positive things associated with the scene, and to anticipate more positive outcomes.[24]

Audiovisual incongruence could operate as such a form of schematic framework in sequences such as this example despite the potential perceived

lack of mood congruence between the steady music and the chaotic images, or the ambiguity in the narrative that Boltz's participants experienced. In the context of *Modern Family*, a mockumentary series, an audience's likely familiarity with this type of schematic framework enables it to contribute towards and complement the comedic framing of this sequence. Dependent on where the perceiver's attentional resources are focused, the muting of diegetic sound as Mitchell destroys the house may potentially draw greater attention to the music, and thus the differences that it presents, and perhaps limit the perceived threat of the pigeon, whose coos were clearly audible in the early part of the sequence. This interpretation might thus limit appreciation of the potential peril facing the pigeon and the extent of the carnage that Mitchell causes. Such comedic framing arguably serves to minimize and trivialize the wider presentation of the destruction: indeed, the consequences of Mitchell's actions are also not immediately seen and visually only shown in a later scene through a brief five second shot that pans out to reveal the extent of the damage whilst he talks to his sister on the phone. The only other reference to the destruction takes the form of an argument between Mitchell and Cam, which centers around the latter noting that most of the destroyed items were gifts from Cam's mother that Mitchell disliked.

Miguel Mera's survey of types of comedic film music emphasizes that context is central to the reception of humor: audience expectations can be shaped in part by the presence of comic actors in comic films 'creating [a] correct air of receptivity',[25] which in turn helps to frame their reception of the humor. A similar line of argument can easily be applied to a TV series, whose format permits multiple episodes to establish a narrative formula and tone. In the context of *Modern Family* the comedic framing of Mitchell's situation is thus arguably readily apparent. It can, therefore, easily enable audiences to read this sequence, which involves a likely familiar audiovisual schematic framework, as parody of a common filmic trope or of specific films dependent on their familiarity with such texts. Indeed, reviewing the episode, critics drew comparisons between this sequence and the assassination montage from Francis Ford Coppola's *The Godfather* (1972) and violent scenes from John Woo's films.[26] Regardless of the music, smoke from the fire extinguisher and the flying feathers

may also bring to mind imagery associated with the carnage of battle sequences and war films, adding to the potential intertextual signification.

Incongruence as schematic framework is here not reliant on novelty and surprise then, but perhaps in fact the opposite. Familiarity with this type of audiovisual construction may facilitate recognition of the parody and emphasize humorous elements of the scene: the comedic context of the series helps to frame such responses and likely minimize any confusion at a prominent audiovisual construction of this nature even though *Modern Family* does not really rely on musical humor to the same degree as other series such as *Family Guy* or *The Simpsons*.[27] For knowing viewers, the resultant intertextual allusion can therefore provide extra-layers of meaning and even reward based on their satisfaction at drawing such parallels or at just getting the joke. The audio-visual difference, and resultant separate encoding of filmic and musical information may draw greater attention to the music, the humor of its incongruity with the images, and the additional texts and layers of meaning that it might evoke. Familiarity with such a learnt cultural schema may make expectations about the use of such audio-visual difference, and the type of contexts in which it might appear, operate on a more subconscious level. Thus if the audio-visual difference evokes a sense of surprise, it may do so in a less shocking manner: as Margulis notes, 'you can be surprised without having been consciously expectant'.[28] Theoretically, in this instance the music choice may not be the most expected option for the narrative situation, but may not be completely *un*expected given the mockumentary series context and potential audience familiarity with the learned schema.

However, audiences might still perceive such audiovisual constructions as "misfitting" or "inappropriate" to some extent, which may contribute towards perceptions of them being surprising. Indeed, such attitudes might even constitute part of the learning of the schema, if this schema of incongruent music is considered as being learnt within the dominant context of the schema of classical Hollywood film scoring,[29] from which much contemporary film music practice derives. Against a history, hegemony and tradition in commercial Hollywood cinema that dates back to the studio era, in which inaudible mood-appropriate music that mirrors and reinforces the images and narrative is

deemed typical, music that draws attention to itself and the cinematic construction could well be consequently polarized as deviant or inappropriate. Regardless of any such societally-shaped explanations, the independent encoding facilitated by the perceived incongruity still provides perceptual space to consider a range of reactions be they: surrounding the perceived appropriateness or inappropriateness of the pairing; or reactions of surprise, amusement, or recognition of intertextual allusion.

Huron's ITPRA model distinguishes between reaction and appraisal responses to events. Reaction responses are quick and unconscious but can be based on learned schemas, drawing in part on cultural and social norms. Appraisal responses are more complex, slower and require the engagement of conscious thought. These responses can contrast in response to individual stimuli, particularly in relation to surprise responses, and involve different neuropsychological pathways.[30] If incongruence such as that featured in *Modern Family* is a culturally learned schema, engagement with such moments could also rely on the disjunction between these mechanisms: the reaction mechanism is characterized by an immediate response which may be influenced by expectations of a more congruent, appropriate, or simply alternative accompaniment, whilst appraisal mechanisms might facilitate recognition of the parody or intertextuality with the perceptual space for such conscious appraisal being created by the separate encoding of auditory and visual information. Such space for appraisal could thus negate or complement the presence *or absence* of conscious surprise or shock, and complement the intertextual reading with which perceivers might approach an incongruent relationship when presented in the parodic sense represented here.

Example #2: *The Hateful Eight* and incongruence as a marker of authorial style

Example #1 highlights some of the complexities in the relationships between audiovisual congruence and incongruence, and violated expectations and surprise. However, parody represents just one contemporary use of audiovisual incongruence. Other directors might use similar constructions without the more obvious comedic intent that may be interpreted in examples such as the *Modern*

Family sequence: indeed, for K.J. Donnelly, drawing upon Fredric Jameson, it is such ulterior, and often comedic, intent that distinguishes the parodic use of existing film-music scoring techniques from more intentionally neutral pastiche.[31] The soundtrack for Quentin Tarantino's recent film *The Hateful Eight* (2015) provides one such example, and also raises interesting talking points about contemporary incongruence and perceiver expectations.

In one sequence, David Hess's delicate piano and acoustic guitar song 'Now You're All Alone' is heard whilst the character Joe Gage (Michael Madsen) hunts down a man by following a trail of his blood in the snow and subsequently executes him. In many ways this use of seemingly anempathetic,[32] or incongruent to use the present terminology,[33] popular song is highly congruent with what one might expect of Tarantino's use of music. Lyrics that talk of being "all alone, feeling that nobody wants you" and of "feeling the world closing on you" could readily be interpreted as highly ironic in relation to Gage's victim eventually being found and gunned down. Equally, there are moments when Gage's measured steps through the snow can feel like they are almost in synchronization with beat of the music, which ends abruptly with the gunshot. However, despite these moments of potential congruity, given the content of the scene again it is likely the contrasting character between song and narrative action that is probably most salient for the viewers, and that is also likely most congruent with their expectations of the director.

Yet, this sequence occurs 126 minutes into the 160 minute-long film and is only one of two non-diegetic popular songs featured.[34] Instead, much of the soundtrack is dominated by Ennio Morricone's original music, which represents an atypical approach within Tarantino's *oeuvre* being his first film to feature a newly composed score. Whilst a different approach procedurally for Tarantino, this score is also perhaps not totally unexpected given the director's frequent use of Morricone's work in his films and the pair's attempted collaboration on *Django Unchained* (2012), which resulted in the composer writing a single song for this earlier film. Yet, often in a minor tonality with angular and chromatic melodic content and ominous held accompaniment pitches, Tarantino described in an interview how Morricone's music for *The Hateful Eight* was not what he was expecting.[35] Regardless of director or audience expectations about his

music, Morricone's score stands in marked contrast to Hess's gentle, major-key song and creates intradomain incongruence or difference between elements within the soundtrack itself.[36] Ironically, the song, despite perhaps being the most congruent with a Tarantino-literate audience's expectations of the director's use of music, is thus arguably the incongruent moment within this soundtrack: the soundtrack itself is perhaps unexpected in relation to expectations based on the previous work of both director and composer, and so the moment that is most congruent with these expectations becomes the more unexpected moment in this film.

Recognition of the broader schema of Tarantino's use of incongruent music and the perceptual space that is created by the independent encoding of seemingly mismatched auditory and visual information might influence engagement in different ways. The knowledge of having made an accurate prediction when the incongruent song finally arrives might provide some sense of satisfaction, in accordance with the prediction response in Huron's ITPRA model: this response mechanism accounts for the positive feelings that act like psychological rewards following accurate predictions, and which encourage similar predictive activity given the evolutionary benefits of preparedness for future situations.[37] Equally, knowledge of Tarantino's typical use of music might make this sudden inclusion of a pre-existing incongruent song feel like the fulfillment of delayed expectations, which might heighten satisfaction in the cinematic moment when it finally arrives. Given the importance of musical selection and placement to Tarantino,[38] it is reasonable to speculate that when planning the placement of this song he may have considered such potential consequences of this delayed presentation of a musical trope that has been readily associated with his *oeuvre*.[39]

Conversely, the impact of the semantic differences that the song presents may be perceived as more surprising given the intradomain incongruities and the re-aligned expectations formulated in response to the prevalence of Morricone's music prior to this point. As with the mockumentary format supporting the comedic framing of the *Modern Family* sequence, here too context is everything. In Huron's terms this moment may be considered as one of dynamic surprise, in which the musical work violates expectations that have

been previously established throughout the work itself. However, Huron also highlights that such dynamic surprises rarely occur without some violation of schematic expectations too:[40] this relationship is particularly apparent given that the approach taken towards constructing this soundtrack is so atypical for Tarantino. The resulting intradomain incongruence thus offers some disjunct in the 'tonally, formally, and affectively unified structure' that Lisa Coulthard contends characterizes the musical and sonic presentation of violence in the director's earlier works.[41] However, as this use of song might still feel stylistically and narratively permitted in a Tarantino film, it may yet contribute to a sense of artifice associated with the portrayal of violence, potentially numbing the effects of the shooting.

A third potential consequence of the delay of the song is that its impact becomes diluted to some extent amidst the more prevalent composed musical material. One online reviewer suggests that the scene serves like a more muted, 'quieter and sadder than ever' iteration of the trope, which creates 'the effect […] of a filmmaker dialing back his usual quirks, drawing less attention to the personal stamp that has become so recognizable over the course of his career'.[42] Regardless of which responses, if any, a perceiver may experience, all of these options rely on literacy with the broader schema of Tarantino's use of incongruent music against which the different treatment of this song and soundtrack more broadly can be situated: whilst the song and the audio-visual difference it provides, and the resultant complicated representations of violence, do not feel *un*expected in a Tarantino film, this particular construction still provides scope for consideration of how a knowing *mélomane* director might manipulate elements of their personal style and authorial voice to influence audience expectation and subsequently response.

Example #3: 'While I play unfitting music' memes and incongruence as a means of participation

The final example of audiovisual incongruence for discussion relates to a type of Internet meme called 'while I play unfitting music'. A video featuring looped footage of the character Luigi from the Mario franchise winding a toy box (see Fig. 2) paired with a recording of the Queen song 'Don't Stop Me Now' began this

trend in 2007.[43] A range of similar memes followed that paired supposedly "unfitting" music with a range of videos, largely featuring looped extracts from video games and animation, but also sometimes using a number of songs or more extended audiovisual material, such as the meme that juxtaposes Boromir's death from *The Lord of the Rings: The Fellowship of the Ring* (Jackson, 2002) with the theme from *Ghostbusters* (Reitman, 1984).[44]

Fig. 2. Luigi winds a toy while I play unfitting music

An additional example shows footage of a model steam train with its wheels circling whilst 'Yakety Sax', perhaps better known as the theme tune to the British *Benny Hill* TV series, plays. This example is particularly interesting given its title: 'model train chugs while I play [un]fitting music'.[45] The square brackets around the un of [un]fitting emphasize the fluid, subjective barriers between judgments of what is fitting, appropriate or congruent and what is not. Moreover, they reflect the way in which those creating these memes, or commenting on them, are actively making their own subjective judgments about whether the images and music fit or not. Such opinions can be found in the comments threads to many of the memes' YouTube pages. For example, commenting on the model train meme, the user LinkEx writes 'Not quite that unfitting tbh, Benny Hill theme goes with just about anything'.[46] Responding to the Luigi-Queen pairing user Max Hellen similarly comments 'This is unfitting music?' whilst Dark ShadowFox suggests 'this is pretty fitting ;O', a statement

that they follow with 'but this song fits to everything o:'.[47] The comments threads for many of these memes contain responses arguing for a degree of perceived fit in the audiovisual relationships, despite the memes' titles, and indeed points of congruity within the memes can readily be identified. For instance, the repetitive motion of Luigi winding the box does not stop, just as the Queen lyrics command, and given how engaged he looks in the task, Luigi may well be "having a good time" and "feel ali-i-i-ive". Similarly, the *Ghostbusters* lyrics "If there's something weird and it don't look good" overlapping with two shots showing the wounded Boromir impaled on an arrow, and his friends looking on in horror, could well be interpreted as ironic and thus complementary to some extent. Moreover, the motion of Luigi winding the box and the cycling wheels of the model train do not feel completely asynchronous with their respective soundtracks: indeed, the statement of the second section of 'Yakety Sax' appears to directly line up with the next rotation of the train's wheels. It is likely that the audio and visual movements are close enough that it is not unreasonable for our brains to identify some temporal relationship here. Equally, it is not implausible that our minds want to find such fitting patterns or explanations that justify the audiovisual pairing given the perceptual principles of proximity and fit that we use to encode external stimuli.

This meme trend demonstrates the technology-enabled forms of audiovisual construction that reflect the 'participatory and immersive sensibility' that Richardson & Gorbman identify as a characteristic of contemporary audiovisual aesthetics,[48] practices that are also reflected in the culture of producing mash-up videos for websites such as YouTube.[49] Communications theorist Limor Shifman too recognizes the way in which the Internet can act as a 'facilitator of participatory culture'.[50] Defining the YouTube memetic video as 'a popular clip that *lures extensive creative user engagement* in the form of parody, pastiche, mash-ups or other derivative work', Shifman analyzed a corpus of YouTube memes that had been selected based on a range of measures that could be used to indicate their popularity.[51] Of the six common features she identified amongst these memes, the qualities of simplicity, repetitiveness, and whimsical content, which either references popular culture or reflects the absence of a concrete theme, could be most readily applied to the 'while I play unfitting

music' memes. These memes also demonstrate several salient attributes that Carol Vernallis identifies in YouTube videos:[52] notably, pulse and reiteration, and intertextuality, which in the three memes cited above is achieved via appropriating a pop song, a film title song, and an instrumental track made famous as the theme to a British TV comedy show respectively.

Not only does this meme trend highlight the prevalence of apparent incongruence in various forms of multimedia but it also demonstrates the role of 'prosumers', a term that conflates notions of producer and consumer,[53] in judgments of (in)congruence. Here prosumers develop such texts and debate the nature of audiovisual relationships: in these memes, the (in)congruity, or at least debate surrounding this judgment, *is* the expectation. The website knowyourmeme.com suggests these memes are 'meant to create inconsistency between the mood created by the music vs. what is happening on the screen',[54] suggesting semantic incongruence is the primary purpose. However, the preceding paragraphs have equally presented interpretations and opinions that call such judgments into question. Treating incongruence as a lack of shared properties placates such disagreement to some extent, instead recognizing the perceptual space to make such judgments, which is opened up by the independent encoding of the distinct musical and visual information. To return to Huron, appraisal responses may be drawn upon to respond to the challenge offered to judge the congruity in these memes, perhaps in light of schematic expectations about how incongruence might work in different media contexts. Perhaps the expectation here is that the multidimensionality of judgments of audiovisual fit will fuel such debates.

Conclusions

The examples discussed in this article emphasize the way in which context, subjectivity, and the multidimensional nature of audiovisual constructions can influence judgments of fit, appropriateness, and congruity in sound-image relationships across various contemporary multimedia. They highlight a range of factors that may influence interpretation and response including extratextual knowledge and expectations about the nature of: audiovisual relationships; genre conventions; individual practitioners; or other texts. Different perceivers

will have differing levels of familiarity and conscious awareness of such extratextual knowledge, pointing to the benefits of an analytical framework, such as the present psycho-semiotic approach, which has the potential to account for: the perspectives of producers and any intentionality that they might have when creating an incongruent relationship; the perspective of an audience responding to this text; and the potential incongruity between these when a text is interpreted in a different way to that intended by its creator.

Each of these examples thus points towards issues of reception *and* production. One key factor to consider when conceptualizing contemporary incongruence therefore is production and the functions that the music may be serving. Discussing the comedic use of pre-existing music in film, like the *Modern Family* sequence, Mera states that the music 'must have a strong filmic context to act with or against' otherwise its meaning may become unclear and the humor may be lost.[55] Recognizing this comedic context, to some extent, emphasizes this quotation of a pre-existing composition as reflecting the choice of a director or music supervisor, which in turn might make the allusions and comedic juxtaposition that the music provides more salient. Such intentionality is arguably also central to prosumers' selections of "unfitting" music for their YouTube memes and subsequent tacit invitations for others to challenge or support these judgments online. Equally, familiarity with the broader qualities of Tarantino's use of music provides an earlier production context against which the director's new output may be considered, which in turn can invite speculation about intentionality surrounding different musical approaches in films like *The Hateful Eight.*

All of these examples also highlight the consideration of the perceived congruity or incongruity that can be facilitated by independent encoding of auditory and visual information. The way in which these forms of incongruence permit perceptual space to consider issues surrounding their production corresponds with aspects of Huron's appraisal response. This recognition of the production and functions of incongruence may aid interpretation of the audio-visual difference as a source of parody in the case of the intertextual allusions that the *Modern Family* sequence might provoke, or as commentary on a practitioner's authorial style that may be invited in *The Hateful Eight* or the

'while I play unfitting music memes'. Moreover, recognizing incongruent music devices such as those discussed above as a type of schematic framework in certain narrative or genre contexts (such as the war films parodied in *Modern Family*), or a particular practitioner's output, allows for these points to extend beyond the examples of pre-existing music that recur throughout the examples in this article to broader types of audio-visual difference that do not rely on intertextual allusions in quite the same explicit or specific way.

Whilst audiovisual incongruence might be more prevalent than it once was, to the extent that producers may utilize it for the types of artistic ends discussed throughout this article, this is not to say that the differences that it presents are not unexpected in other ways. As Huron's work emphasizes, there are different types of expectation and therefore different types of surprise too. Shifman suggests incongruence in YouTube memes can present a puzzle for the perceiver to solve, as they consider the nature of the audio-visual difference.[56] However, incongruities do not necessarily require resolution, as reflected by the paradoxical terms 'appropriate incongruence' that Marshall Heiser uses to discuss comedic film music,[57] or 'established incongruence' that Jeroen Vandaele uses to describe the way in which elements of jokes can become established and expected despite their surface incongruities.[58] The distinction that Huron recounts between schematic and veridical expectations helps to emphasize this point: the former label refers to violations of expectation in relation to existing schemas that perceivers might draw upon; the latter term refers to violations of expectation in relation to knowledge of a specific work.[59] It is the distinction between such expectations that can allow known incongruent audiovisual constructions to retain some element of surprise: perceivers may still recognize violations of schematic expectations despite familiarity with a work or genre, dependent on the schematic frameworks they are using to navigate the text.

The ability to use expectations in the ways discussed throughout this article highlights contemporary audiences' potential cinematic literacy, and the ways in which audio-visual difference opens up perceptual space to consider authorial agency, as well as or instead of, the emotional disjunction between sight and sound in such moments. Such incongruities come in a variety of forms and invite a range of responses, including invitations to reconcile the differences

or to appreciate the tensions they present. The incongruities reflect the demands of particular media, and the intentions of various producers. Recognizing when we might expect incongruence or when we might expect a more congruent construction can help to understand this range of responses and intentions linked to audio-visual difference. An analytical framework that considers this plethora of perspectives and responses should account for contextual factors that shape judgments of incongruence in relation to individual perceivers and individual texts. It should also account for the perceptual space that audio-visual difference can open to appraise these texts and should reflect various levels of textual engagement, in regard to production *and* reception, two perspectives which themselves might not necessarily be congruent.

END NOTES:

[1] Sergei Eisenstein, *Film Form: Essays in Film Theory,* ed. and trans. Jay Leyda (San Diego: Harcourt, Inc., 1949/1997).

[2] For further examples of such discussions see James Wierzbicki, *Film Music: A History* (New York: Routledge, 2009), pp.33-35.

[3] For more on *mélomane* directors see Claudia Gorbman, "Auteur Music," in *Beyond the Soundtrack: Representing music in cinema,* ed. Daniel Goldmark, Lawrence Kramer and Richard Leppert (Berkley: University of California Press, 2007), 149-162.

[4] Marilyn G. Boltz, "The cognitive processing of film and musical soundtracks," *Memory and Cognition 32*, no.7 (2004): 1194-1205, p.1202.

[5] Nicholas Cook, *Analysing Musical Multimedia* (Oxford: Oxford University Press, 1998/2000).

[6] Annabel J. Cohen, "Music in performance arts: Film, theater and dance," in *The Oxford Handbook of Music Psychology* (second edition), ed. Susan Hallam, Ian Cross and Michael Thaut (Oxford: Oxford University Press, 2016), 725-744.

[7] For examples of empirical research that specifically touches upon incongruence in these terms, please see Boltz (2004); or Scott D. Lipscomb and Roger A. Kendall, "Perceptual judgment of the relationship between musical and visual components in film," *Psychomusicology 13,* no. 1 (1994): 60-98.

[8] Stan Link, "Sympathy with the devil? Music of the psycho post-*Psycho", Screen 45,* no. 1 (2004): 1-20, p.7.

[9] Marilyn G. Boltz, Matthew Schulkind, and Suzanne Kantra, "Effects of background music on the remembering of filmed events," *Memory and Cognition 19,* no. 6 (1991): 593-606.

[10] David Ireland, "Deconstructing incongruence: A psycho-semiotic approach toward difference in the film-music relationship," *Music and the Moving Image 8,* no. 2 (2015): 48-57.

[11] "Soundtrack Dissonance," at *TvTropes,* http://tvtropes.org/pmwiki/pmwiki.php/Main/SoundtrackDissonance (accessed 18 August 2016).

[12] Kathryn Kalinak, *Film Music: A very short introduction* (Oxford: Oxford University Press, 2010), p.2.

[13] Steven Willemsen and Miklós Kiss, "Unsettling melodies: A cognitive approach to incongruent film music," in *The Cinema of Sensations,* ed. Ágnes Pethő (Newcastle upon Tyne: Cambridge Scholars Publishing, 2015), 103-120, p.170.

[14] An unfolding history of multimedia in which more recent texts add to a catalogue of earlier texts and practices cumulatively creates a greater prevalence of seemingly incongruent audiovisual combinations than previously existed. Moreover, Wierzbicki (2009) suggests an 'endemic' use of music that 'plays "against" on-screen action or emotion', which dates 'at least' to the 1960s and the modernist "new wave" movement in film (p.228). The now 'commonplace' use of light classical or upbeat popular music to accompany scenes of violence that Kalinak (2010) cites (p.2), and the increasing number of *mélomane* directors that Gorbman (2007, p.149) identifies, may also be explained to some extent by the idea of current practitioners emulating or reacting against earlier texts and approaches. Gorbman also notes the rise of compilation scoring practices, and the technological and industrial affordances that accompanied this trend, as some factors that have enabled *mélomane* directors to pursue such use of music in their films.

[15] Elizabeth Margulis, "Surprise and listening ahead: Analytic engagements with musical tendencies," *Music Theory Spectrum 29,* no. 2 (2007): 197-218.

[16] David Huron, *Sweet Anticipation: Music and the psychology of expectation* (Cambridge, MA: The MIT Press, 2006).

[17] Ireland, (2015).

[18] Huron, p.15; p.17.

[19] ibid, p.3.

[20] For more on spatially displaced sound see Jeff Smith, "Bridging the Gap: Reconsidering the border between diegetic and non-diegetic music," *Music and the Moving Image 2,* no. 1 (2009): 1-25.

[21] I am grateful for the comments of the anonymous peer-reviewer who suggested the fruitfulness of exploring these points of congruence in the sequence.

[22] Scott D. Lipscomb, "Cross-modal alignment of accent structures in multimedia," in *The Psychology of Music in Multimedia*, ed. Siu-Lan Tan, Annabel J. Cohen, Scott D. Lipscomb and Roger A. Kendall (Oxford: Oxford University Press), 192-213, p.206.

[23] Huron, p.204.

[24] Marilyn G. Boltz, "Musical soundtracks as schematic influence on the cognitive processing of filmed events," *Music Perception 18,* no. 4 (2001): 427-454.

[25] Miguel Mera, "Is funny music funny? Contexts and case studies of film music humor," *Journal of Popular Music Studies 14* (2002): 91-112, p.97.

[26] James Poniewozik, "Modern Family Watch: Photo Finish," at *Time Online,* http://entertainment.time.com/2010/05/20/modern-family-watch-photo-finish/ (accessed 18 August 2016); Robert Canning, "Modern Family: 'Family Portrait' Review," at *IGN,* http://uk.ign.com/articles/2010/05/20/modern-family-family-portrait-review (accessed 18 August 2016).

[27] See for example, Ronald Rodman, "Auteurship and agency in television music," in *The Oxford Handbook of Film Music Studies,* ed. David Neumeyer (Oxford: Oxford University Press, 2014), 526-555, p.534.

[28] Margulis, p.203.

[29] See for example the theorization of the classical Hollywood score in Chapter 4 of Claudia Gorbman, *Unheard Melodies: Narrative Film Music* (London: BFI Publishing, 1987), 70-98.

[30] Huron, pp.13-14.

[31] K. J. Donnelly, "The Classical Film Score Forever? *Batman*, *Batman Returns* and post-classical film music," in *Contemporary Hollywood Cinema,* ed. Steve Neale and Murray Smith (London: Routledge, 1998), 142-155.

[32] Michel Chion, *Audio-vision: Sound on Screen,* ed. and trans. Claudia Gorbman (New York: Columbia University Press, 1990/1994).

[33] For more on the relationship between anempathy and incongruence see: David Ireland, "'It's a sin […] using Ludwig van like that. He did no harm to anyone, Beethoven just wrote music': The role of the incongruent soundtrack in the representation of the cinematic criminal," in *Constructing Crime: Discourse and cultural representations of*

crime and 'deviance', ed. Christiana Gregoriou (Basingstoke: Palgrave Macmillan, 2012), 97-111, p.108.

[34] Timings here refer to the 2015 Visiona Romantica, Inc. DVD release of the film. At this stage, nearly two hours have passed since the use of the first of these two quoted songs, "Apple Blossom" by The White Stripes, which occurs at just 14 minutes into the film.

[35] Kristopher Tapley, "'The Hateful Eight': how Ennio Morricone wrote his first Western score in 40 years," at *Variety Online,* http://variety.com/2015/film/in-contention/hateful-eight-quentin-tarantino-ennio-morricone-western-score-1201659489/ (accessed 18 August 2016).

[36] David Ireland, "Singin' over rainbows: The incongruent film song and extra-filmic reception," *The Soundtrack 7,* no. 2 (2014): 119-132.

[37] Huron, pp.12-13.

[38] Ken Garner, "'Would You Like to Hear Some Music?' Music in-and-out-of-control in the Films of Quentin Tarantino," in *Film Music: Critical Approaches,* ed. K.J. Donnelly (Edinburgh: Edinburgh University Press, 2001), 188-205.

[39] Whilst I am yet to find any evidence about such a strategy in relation to this pacing of *The Hateful Eight*'s soundtrack, interview material does point to the clear intentional use of this song in this scene: Tarantino describes using Hess's music identically to its use in *The Last House on the Left* (Craven, 1972), a film which also starred Hess. For more information see: Kim Morgan, "The Interview: Quentin Tarantino," *Sight and Sound 26,* no. 2 (2016), 18-27.

[40] Huron, pp.278-279.

[41] Lisa Coulthard, "Torture tunes: Tarantino, popular music, and new Hollywood ultraviolence," *Music and the Moving Image 2,* no. 2 (2009): 1-6, p.5

[42] Timothy Lawrence, 'The Hateful Eight,' at *FilmFisher,* http://filmfisher.com/films/the-hateful-eight (accessed 18 August 2016).

[43] *Luigi winds a toy while I play unfitting music*, at *YouTube,* https://www.youtube.com/watch?v=b-Z5IzXYbCw (accessed 18 August 2016).

[44] *Boromir dies while I play unfitting music*, at *YouTube,* https://www.youtube.com/watch?v=FQmMOS-qumo (accessed 18 August 2016).

[45] *Model train chugs while I play [un]fitting music*, at *YouTube,* https://www.youtube.com/watch?v=yVUlyb_1yS8 (accessed 18 August 2016).

[46] ibid

[47] *Luigi winds a toy while I play unfitting music.*

[48] John Richardson and Claudia Gorbman, "Introduction," in *The Oxford Handbook of New Audiovisual Aesthetics,* ed. John Richardson, Claudia Gorbman and Carol Vernallis (Oxford: Oxford University Press, 2014), 3-35, p.7.

[49] Nicholas Cook, "Beyond music: Mashup, multimedia mentality, and intellectual property," in *The Oxford Handbook of New Audiovisual Aesthetics,* ed. John Richardson, Claudia Gorbman and Carol Vernallis (Oxford: Oxford University Press, 2014), 53-76.

[50] Limor Shifman, "An anatomy of a YouTube meme," *New Media and Society 14,* no. 2 (2011): 187-203, p.190.

[51] ibid

[52] Carol Vernallis, *Unruly Media: YouTube, Music Video and the new digital cinema* (Oxford: Oxford University Press, 2013).

[53] ibid, p.308 n.67.

[54] *… While I Play Unfitting Music*, at *Internet Meme Database: Know Your Meme,* http://knowyourmeme.com/memes/while-i-play-unfitting-music (accessed 18 August 2016).

[55] Mera, p.97.

[56] Shifman, p.196.

[57] Marshall Heiser, "The soundtrack as appropriate incongruity," in *Sounding Funny: Sound and Comedy Cinema*, ed. Mark Evans and Philip Hayward (Sheffield: Equinox Publishing Limited, 2016), 14-28.

[58] Jeroen Vandaele "Humor mechanisms in film comedy: Incongruity and superiority," *Poetics Today 23,* no. 2 (2002): 221-249.

[59] Huron, p.269.