The Music-Image:

Closing the Methodological Gap between Musicology and Film Studies

Will Jeffery

A thesis submitted to fulfil requirements for the degree of Doctor of Philosophy

Faculty of Arts and Social Sciences

The University of Sydney

2023

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Abstract

What is film music and how should it be analysed? Film Music Studies has a problematic history that has separated the analysis of music from the image due to the opposing methodologies of Musicology and Film Studies. This thesis discovers a new way of thinking about film music called the music-image. The music-image is the semiotic ambiguity that interlocks 'music' and 'image' through time, analysing a whole image with no fixed meaning or communication, bridging the methodologies of Musicology and Film Studies to do so. The music-image is a philosophical and theoretical model that encourages new film music analysis, rather than an applied analytical model. Theorists who have attempted this ‘bridging’ have leaned towards creating unified theories of musical score and visual form, addressing mainly the mechanics of how music and images *work*. However, such models rarely consider the spectator’s experience as part of the puzzle. There needs to be a new *way of thinking* about film music to interlock music and image in one analysis to consider a whole audio-visual *experience*. I term this ontological interlocking of music and image, quite literally, the music-image. The music-image adopts Gilles Deleuze’s concept of the ‘spiritual automaton’ into the music-image automaton which interlocks the semiotic systems of music and visual image in a whole audio-visual system of meaning. This philosophical and theoretical model avoids any top-down ‘Grand Theory’ as the music-image needs the experience of the spectator and cinema to exist, as any film music needs cinema to exist. The methodologies of Musicology and Film Studies are bridged to explore the music-image as one ‘whole’ image.

Acknowledgements

This thesis would not have been completed without the support of those around me.

Firstly, I would like to extend my sincere thanks to my supervisor, Assoc. Prof. Bruce Isaacs. I have appreciated his endless enthusiasm for my thesis topic and film in general, his encouragement especially in the more challenging areas of the thesis, and the detailed advice he provided along the way. It has been an enjoyable and rewarding experience.

Special thanks to my colleagues at The University of Sydney for your shared love of film, friendship and support.

To my close friends, and to everyone I had conversations with about my thesis over the years, I have loved talking to you about film and it means a great deal how much you care and have supported me – thank you.

To my family, living close by and abroad, thank you for your love and support during the writing of this thesis.

Finally, to my parents Barbara and Martin, this thesis would not have been possible without the love of film you shared with me from an early age. A massive thank you to my mother, Barbara, who proofread my thesis. To my father, Martin, thank you for your infectious love of film music which ultimately inspired me to write this thesis. Thank you for making me CD compilations of classic movie soundtracks all those years ago which began my love of film music. For these reasons, I dedicate this thesis to my parents.

Table of Contents

Page

Abstract **2**

Acknowledgements **3**

Table of Contents **4**

List of Figures **6**

Introduction **12**

Chapter 1: Towards a New Way of Thinking about Film Music **22**

- The problem with Film Music Studies in academic literature 23
- Closing the (methodological) gap between Music(ology) and Film (Studies) 32
- Grand Theories and ‘communications’ models 34
- Analysing film music through film experience 39
- Eisenstein’s Vertical Montage and towards a new way of avoiding separatism 45

Chapter 2: What is the Music-Image? **50**

- The difference and repetition of film music 51
- The music-image automaton 54
- Music and image as one image 62
- The music-image 68

Chapter 3: The Nondiegetic Music-Image **73**

- What is underscore and how is it sourced 74
- Audio-visual harmony and dissonance 77
- Audio-visual time: duration, rhythm, and tempo 87
- Audio-visual tone and texture 97
- Audio-visual space 106
- Conclusion 113

Chapter 4: A Music-Image Analysis of *Jaws* (Spielberg, 1975) **114**

- The *Jaws* title sequence 114
- The death of Chrissy 120
- The death of Alex Kintner 124
- Face to face with the shark 128
- The death of the shark 135
- Conclusion 140

Chapter 5: ‘When’ is the Music-Image? 142

Is nondiegetic film music ever really ‘unheard’? 142

‘When’ is the music-image in *The Hours* (Daldry, 2002)? 153

When music-images repeat 161

When is the first music-image? 166

The Overture 168

The Title Sequence 171

Conclusion 190

Chapter 6: Where is the Music-Image? 192

What is diegetic film music? 193

Redefining diegetic film music 194

Principles of film narrative 197

Visibly sourced film music 198

The invisible source 204

Is diegetic film music actually diegetic? 213

The film is a stage 220

Sourceless musical sources 232

Idealism as a diegetic bridge 236

La La Land (Chazelle, 2016) 240

Conclusion 247

Chapter 7: The Breakdown of the Diegesis 248

The breakdown of the music-image 249

Towards a non-binary diegesis 260

The transdiegetic music-image 267

When nondiegetic film music enters the diegesis 284

Conclusion 289

Conclusion: Beyond the Music-Image 291

Bibliography 296

Filmography 312

List of Figures

Figure 1 Gabriel playing his oboe in *The Mission* 42

Figure 2 A wide shot of a landscape in *The Mission* 42

Figure 3 Natives closing in on Gabriel in *The Mission* 42

Figure 4 Natives approach Gabriel in *The Mission* 43

Figure 5 Bart riding on horseback in *Blazing Saddles* 44

Figure 6 Bart approaches an orchestra in *Blazing Saddles* 44

Figure 7 Sergei Eisenstein's 'Vertical Montage' 47

Figure 8 The spectator's experience of film inside the model of an automaton 57

Figure 9 George in a close-up shot in *The Time Machine* 60

Figure 10 George in a mid-shot *The Time Machine* 60

Figure 11 Judy Garland singing 'Somewhere over the Rainbow' in *The Wizard of Oz* 68

Figure 12 Jesse in blue light on the runway in *The Neon Demon* 81

Figure 13 Jesse in blue light on the runway in *The Neon Demon* 81

Figure 14 Jesse in red light, becoming the 'neon demon' in *The Neon Demon* 82

Figure 15 Jesse in red light, becoming the 'neon demon' in *The Neon Demon* 82

Figure 16 Wide shot in *The Social Network* 85

Figure 17 Wide shot in *The Social Network* 85

Figure 18 A shot fragment of reaching Hamlet at the top of the staircase in *Hamlet* 92

Figure 19 A shot fragment of reaching Hamlet at the top of the staircase in *Hamlet* 92

Figure 20 A shot fragment of reaching Hamlet at the top of the staircase in *Hamlet* 92

Figure 21 A shot fragment of reaching Hamlet at the top of the staircase in *Hamlet* 92

Figure 22 A shot fragment of reaching Hamlet at the top of the staircase in *Hamlet* 92

Figure 23 A shot fragment of reaching Hamlet at the top of the staircase in *Hamlet* 92

Figure 24 Hashem notices his son in *The Color of Paradise* 100

Figure 25 The beginning of a crane shot in *The Color of Paradise* 100

Figure 26 Mohammad's hand lights up gloriously in *The Color of Paradise* 101

Figure 27 The connoted, poetic opening image of *Saving Private Ryan* 103

Figure 28 The denotated, prosaic opening image of *JFK* 103

Figure 29 An empty hallway of the Nostromo spaceship at the beginning of *Alien* 108

Figure 30 An empty hallway of the Nostromo spaceship at the beginning of *Alien* 108

Figure 31 Justine notices blood on her shoulder in *Raw* 111

Figure 32 Justine looks up to see blood being poured on her and her classmates in *Raw* ... 111
Figure 33 Justine and her classmates are splattered with blood from above in *Raw* 111
Figure 34 An opening title card in *Jaws* 117
Figure 35 The opening underwater shot in *Jaws* 117
Figure 36 The opening underwater point of view shot in *Jaws* 118
Figure 37 The E-F ostinato in *Jaws* 118
Figure 38 The *Jaws* title card 120
Figure 39 A dominant 7th chord progression played by a tuba in *Jaws* 120
Figure 40 A point of view of Chrissy from the depths of the ocean in *Jaws* 122
Figure 41 Beginning of the ‘Trombone shot’ in *Jaws* 127
Figure 42 End of the ‘Trombone shot’ in *Jaws* 127
Figure 43 Brody and Quint notice the shark coming towards the Orca in *Jaws* 131
Figure 44 The shark coming towards the Orca in *Jaws* 131
Figure 45 The shark’s monstrous size revealed for the first time in *Jaws* 132
Figure 46 The shark approaches the boat with the ostinato in *Jaws* 133
Figure 47 The shark is threatened by the boat in a reverse angle pirate melody in *Jaws* 135
Figure 48 Low angle shot of Brody climbing the mast in *Jaws* 136
Figure 49 High angle shot of Brody climbing the mast in *Jaws* 136
Figure 50 Brody’s Theme from *Jaws* 136
Figure 51 The low E returns with Brody in a wide shot in *Jaws* 139
Figure 52 Close-up of the shark in a montage in *Jaws* 139
Figure 53 Mid-shot of Brody shooting the shark in *Jaws* 139
Figure 54 Close-up of Brody’s joy in *Jaws* 139
Figure 55 The destruction of the shark in *Jaws* 139
Figure 56 The descending shark carcass in *Jaws* 140
Figure 57 Over the shoulder shot of Isak in *Wild Strawberries* 146
Figure 58 Isak and the arm in *Wild Strawberries* 146
Figure 59 The low F note in *The Thing* 149
Figure 60 Childs drinking beer in *The Thing* 150
Figure 61 MacReady watching Childs drink beer in *The Thing* 150
Figure 62 The aftermath of the explosion in *The Thing* 150
Figure 63 The shapeshifting chord progression in *The Thing* 151
Figure 64 A close-up shot of Virginia declaring the life of the poet in *The Hours* 156
Figure 65 A close-up shot of Leonard in *The Hours* 156

Figure 66 Music that leads nowhere melodically in *The Hours* 156
Figure 67 Virginia and Leonard walking away from the camera in *The Hours* 157
Figure 68 The cut to Laura driving her car in *The Hours* 157
Figure 69 Laura and Richard face to face in *The Hours* 159
Figure 70 Laura and Richard's fragile relationship musically in *The Hours* 159
Figure 71 *Ben-Hur*'s Overture 171
Figure 72 A sample of *Enter the Void*'s strobe-lit title sequence 172
Figure 73 A sample of *Enter the Void*'s strobe-lit title sequence 172
Figure 74 A fragment of *Enter the Void*'s fiery title card 173
Figure 75 A fragment of *Enter the Void*'s fiery title card 173
Figure 76 A fragment of *Enter the Void*'s fiery title card 173
Figure 77 *Vertigo*'s title card 175
Figure 78 Example of the spiralling imagery in *Vertigo*'s title sequence 175
Figure 79 Example of the spiralling imagery in *Vertigo*'s title sequence 176
Figure 80 The opening image design of *North By Northwest*'s title card 177
Figure 81 *North by Northwest*'s title card image now the side of a New York building 177
Figure 82 Alfred Hitchcock misses a bus as *North by Northwest*'s title sequence ends 178
Figure 83 An example image from the title sequence of *The 7th Voyage of Sinbad* 179
Figure 84 An example image from the title sequence of *The 7th Voyage of Sinbad* 179
Figure 85 An example image from the title sequence of *The 7th Voyage of Sinbad* 179
Figure 86 An example image from the title sequence of *Jason and the Argonauts* 180
Figure 87 An example image from the title sequence of *Jason and the Argonauts* 180
Figure 88 An example image from the title sequence of *Jason and the Argonauts* 180
Figure 89 *Catch Me If You Can*'s title card 182
Figure 90 The opening segment of *Catch Me If You Can*'s title sequence 182
Figure 91 The title card for *Star Wars* beginning larger than the film frame 184
Figure 92 A sample from the opening title scroll for *Star Wars* 184
Figure 93 A sample from the opening title scroll for *Star Wars* 185
Figure 94 The opening title scroll for *Star Wars* is now the image of space 185
Figure 95 The opening title scroll for *Star Wars* become part of the narrative landscape .. 185
Figure 96 An out of focus shot of a bedroom in the title sequence for *Lolita* 187
Figure 97 The title card for *Lolita* 187
Figure 98 The title card for *Halloween* (Carpenter, 1978) 189
Figure 99 the title card for *Halloween* (Gordon Green, 2018) 189

Figure 100 The jack-o'-lantern coming to life in the title card for *Halloween* (2018) 189
Figure 101 The jack-o'-lantern coming to life in the title card for *Halloween* (2018) 189
Figure 102 The jack-o'-lantern coming to life in the title card for *Halloween* (2018) 189
Figure 103 Harpo Marx playing the loom in *Go West* 200
Figure 104 Harpo Marx playing the harp in *Horse Feathers* 200
Figure 105 Harpo Marx playing the harp in *A Night in Casablanca* 200
Figure 106 Harpo Marx playing the harp in *A Night at the Opera* 201
Figure 107 An audience watching Harpo play the harp in *A Night at the Opera* 201
Figure 108 A reverse angle of Harpo Marx playing the harp in *A Night at the Opera* 201
Figure 109 Al Jolson performs in a shot lasting 63 seconds in *The Jazz Singer* 203
Figure 110 The camera cuts away from Al Jolson performing in *The Jazz Singer* 203
Figure 111 The camera cuts away from Al Jolson performing in *The Jazz Singer* 203
Figure 112 The camera cuts back to Al Jolson performing in *The Jazz Singer* 203
Figure 113 Samson lets go of his right ear in *Samson and Delilah* 205
Figure 114 Samson unblocks his left ear in *Samson and Delilah* 205
Figure 115 Samson turns on the amplifier in *Samson and Delilah* 207
Figure 116 Samson dancing in a wide shot in *Samson and Delilah* 207
Figure 117 Delilah watches Samson dance in *Samson and Delilah* 208
Figure 118 A mid-shot of Samson dancing in *Samson and Delilah* 208
Figure 119 The songwriter striking a chord on his piano in *Rear Window* 211
Figure 120 A neighbouring apartment in *Rear Window* 211
Figure 121 A neighbouring apartment in *Rear Window* 212
Figure 122 Ally is told she is late for work in *A Star is Born* 218
Figure 123 Ally decides to leave her job in *A Star is Born* 218
Figure 124 Ally and Ramon travel by car to the music gig in *A Star is Born* 218
Figure 125 Ally and Ramon travel by plane to the music gig in *A Star is Born* 219
Figure 126 Ally and Ramon arrive at the venue for the music gig in *A Star is Born* 219
Figure 127 Ally and Ramon arrive onstage in *A Star is Born* 219
Figure 128 An onstage performance from *Yankee Doodle Dandy* 222
Figure 129 An onstage performance from *Yankee Doodle Dandy* 222
Figure 130 An onstage performance from *Yankee Doodle Dandy* 222
Figure 131 An onstage performance from *Yankee Doodle Dandy* 223
Figure 132 An extreme wide shot of the stage in *Rhapsody in Blue* 226
Figure 133 The onstage camera shows the theatre audience in *Rhapsody in Blue* 227

Figure 134 George in close-up as he performs onstage in *Rhapsody in Blue* 227
Figure 135 The piano accompanist is acknowledged offscreen in *You'll Never Get Rich* .. 229
Figure 136 Robert and Sheila dance in *You'll Never Get Rich* 229
Figure 137 An extreme high angle shot from the rear of the stage in *The Red Shoes* 231
Figure 138 The audience as a turbulent seaside in *The Red Shoes* 231
Figure 139 A mid-shot of Bob, Cleo, and Maurice at the piano in *Cleo from 5 to 7* 234
Figure 140 The camera swings around the piano in *Cleo from 5 to 7* 234
Figure 141 The camera moves into a close-up of Cleo's face in *Cleo from 5 to 7* 235
Figure 142 A close-up shot of Cleo in *Cleo from 5 to 7* 235
Figure 143 The camera elevates at the end of the opening sequence in *La La Land* 242
Figure 144 The camera tilts up for *La La Land*'s title card 242
Figure 145 The camera tilts back down to the cars after *La La Land*'s title card 242
Figure 146 Mia and Seb dance with a pendulum contraption in *La La Land* 243
Figure 147 Seb looks directly in the camera in *La La Land* 244
Figure 148 Mia looks directly in the camera in *La La Land* 244
Figure 149 Seb and Mia float among stars and planets in *La La Land* 244
Figure 150 Mia begins in a wide shot in the audition room in *La La Land* 246
Figure 151 The lights begin to dim in the audition room in *La La Land* 246
Figure 152 The camera tracks around Mia's face in the audition room in *La La Land* 246
Figure 153 A sample of the oscillating final shot from *The Conversation* 254
Figure 154 A sample of the oscillating final shot from *The Conversation* 254
Figure 155 The oscillating final shot with Harry in frame from *The Conversation* 255
Figure 156 A sample of the oscillating final shot from *The Conversation* 255
Figure 157 A sample of the oscillating final shot from *The Conversation* 256
Figure 158 The T-1000 rams an iron girder in *Terminator 2: Judgment Day* 262
Figure 159 The Terminator about be hit by iron girder in *Terminator 2: Judgment Day* ... 262
Figure 160 The title card for *Eyes Wide Shut* 264
Figure 161 A shot of a Manhattan street in the opening of *Eyes Wide Shut* 264
Figure 162 Bill locating his wallet in the opening of *Eyes Wide Shut* 265
Figure 163 Bill switches off the music player in the opening of *Eyes Wide Shut* 265
Figure 164 Louis outside his store in *À Nous la Liberté* 269
Figure 165 Louis resets his phonograph in *À Nous la Liberté* 269
Figure 166 Louis' phonograph crossfades to a factory in *À Nous la Liberté* 270
Figure 167 The outside of a factory in *À Nous la Liberté* 270

Figure 168 The face of the zither in the title sequence of *The Third Man* 274
Figure 169 The title card for *The Third Man* on the face of a zither 274
Figure 170 An establishing shot of Vienna after the title sequence of *The Third Man* 275
Figure 171 Mozart asks Salieri for the score of Mozart’s Requiem in *Amadeus* 277
Figure 172 Mozart cues the Requiem in *Amadeus* 277
Figure 173 A shot of a horse carriage in *Amadeus* 278
Figure 174 Constanze travels home in the horse carriage in *Amadeus* 278
Figure 175 Constanze greets Salieri as she returns home in *Amadeus* 279
Figure 176 Constanze greets Mozart as she returns home in *Amadeus* 279
Figure 177 The scientists in *Close Encounters of the Third Kind* 283
Figure 178 The spaceship in *Close Encounters of the Third Kind* 283
Figure 179 The spaceship takes off in *Close Encounters of the Third Kind* 284
Figure 180 Harry notices Hagrid in *Harry Potter and the Philosopher’s Stone* 286
Figure 181 A wide shot of Hagrid in *Harry Potter and the Philosopher’s Stone* 286
Figure 182 Harry, Ron, and Hermione run towards Hagrid in *Harry Potter and the Philosopher’s Stone* 286
Figure 183 Harry, Ron, Hermione confront Hagrid in *Harry Potter and the Philosopher’s Stone* 287

Introduction

> "When the favourite [music] passages are extracted from their original context and presented in the concert hall, almost invariably it sounds as though something is missing."[1]

After the theatrical release of *La La Land* (Chazelle, 2016), I experienced the film at a concert hall with a live orchestra performing the soundtrack. *La La Land* is a film musical, which Rick Altman has influentially argued are films with music that emanates from the fictional film world created by the film.[2] However, in my experience in the concert hall that night, the film's music emanated instead from within the space of my world, that is, the space of the concert hall. As I watched the film, I became aware that the music from the orchestra pit of the concert hall appeared to be manipulating the actions unfolding on screen. Of course, music cannot physically do this, and as *La La Land* was filmed and constructed years prior to my concert hall experience, there was no way in which the music could have enacted such an effect upon the film. And yet, in this process of experiencing *La La Land* in a concert hall with the music source physically and visually separated from the screen (in the orchestra pit), I came to realise that film music and the visual image were interacting through my holistic experience of the film. The film and its 'film music' were interlocked in one image of experience. Thus, when Roger Scruton, a leading musicologist and theorist, suggests that "something is missing" when film music is extracted and heard separately from the film, I argue that it is the *film itself* that is missing.

What is film music and how should it be analysed? That is a question I have asked myself ever since my fascination with film began. How does one describe or analyse what film music is and what it is doing in a film at a certain time? Defining film music is a simple yet complex task that has attracted much scholarly debate. Put simply, film music is music in film. In 1950, Raymond Spottiswoode defined music 'in' film as: "A musical score, either already existing or specially composed…synchronised with the visual film."[3] The two key claims in this definition are that film music is *composed* for film and *synchronised* with film. Film music, then, exists totally for film. However, Martin Marks in 1979 argues that film music "exists only as an accompaniment to a film", inferring that it is separate and discrete.[4] How can film music both accompany and synchronise with film?

---

[1] Roger Scruton, *Music as an Art* (London; New York; Oxford; New Delhi; Sydney: Bloomsbury Continuum, 2018), 180.

[2] Rick Altman, *The American Film Musical* (Bloomington; Indianapolis: Indiana University Press, 1987), 12.

In a visit to the Museum of the Moving Image in Queens, New York in 2019, I stumbled across a film music exhibit that invited the museum goer to match the 'correct' film music excerpt with a film sequence on a computer. There were already two children participating under adult supervision, so I had to wait my turn and watch on. The selected sequence was from the film, *Twister* (Jan de Bont, 1996). The two children were given a few different musical excerpts to choose from and pair with the film sequence. It was not until their final attempt with the final available musical excerpt that the screen told them they had found the correct match. I was not familiar with the film, and so I had as little idea as the children of the 'correct' match. What struck me was that for every attempt the children made, I was satisfied they had made the correct match. Every musical excerpt was 'fitting' with the sequence in some way. However, each audio-visual matchup created a completely different sequence. This experience seems to me to reflect Michel Chion's notion of the "forced marriage" of sound and image. In an echo of my experience, Chion describes experiencing a film sequence that has not previously been heard or seen where various musical excerpts played over the image generate radically different synchronisations.[5] Could we not take Chion's notion even further? That is, to suggest that this "forced marriage" is in fact the situation of every audio-visual relationship, with each bearing a different meaning?[6]

---

[3] Raymond Spottiswoode, *A Grammar of the Film* (Berkeley: University of California Press, 1950), 49.

[4] Martin Marks, "Film Music: The Material, Literature, and Present State of Research," *MLA Notes* 36, no. 2 (1979): 283.

In his brief analysis of Lufthansa's *New York/New York* TV advertisement, which shows the same montage sequence twice but with different soundtracks, one with free jazz/noise, the other with Pachelbel's Canon in D, Oliver Vitouch makes a profound observation. Although the sequence is followed by the slogan "you see the world the way you fly," Vitouch suggests that "we sometimes see the world the way it sounds."[7] The same visual clip repeated with two different pieces of music played over the top completely changes the experience of the sequence. When I showed this clip to film students in 2020, one student was convinced that the second sequence (the version with Pachelbel's Canon) was edited more slowly (with a longer shot duration) than the initial jazz music version, even though both sequences have the same duration (both in shots and sequence). Why is this so? The discussion following this clip could focus on the changing tempo and mood of the music or analyse the rhythmic montage of the images of Manhattan streets. But couldn't we also find some way to discuss both the aural and visual language as one image of experience? If so, it would be inadequate to treat the music in both the museum exhibit and Lufthansa's *New York/New York* TV advertisement as 'accompaniment'. When different stimulus is heard, different visual stimulus is 'seen', and so different stimulus is *felt*. What encompasses these stimuli is one image of experience.

---

[5] Michel Chion, *Audio-Vision: Sound on Screen*, trans. Claudia Gorbman (New York: Columbia University Press, 1994), 188.

[6] Ibid., 189.

[7] Oliver Vitouch, "When Your Ear Sets the Stage: Musical Context Effects in Film Perception," *Psychology of Music* 29 (2001): 81. See, "Lufthansa New York," Springer and Jacoby. YouTube video, https://www.youtube.com/watch?v=VPjmJOzsP28., for *New York/New York* spot.

Why use language such as ‘accompaniment’ if film music seems inseparable from film experience? Do we have the liberty to pick and choose which music to ‘accompany’ a film? No, the film is set; *Twister*, for example, is more than a film image, it is a film with exact audio-visual parameters. This traditional separation of film and music derives from a problematic history of film music scholarship dominated by the two perspectives of Film Studies and Musicology. My research project began by surveying perspectives of film music, and varying results were collected due to different and competing methodologies to analyse film music. Film music is often neglected in studies of film due to its inherent musicological qualities and characteristics; and yet film music is often neglected in musicological studies due to its inherent filmic qualities and characteristics. When the two disciplines engage in an analysis film music, there is commonly a *separation* of film from music, and music from film.

In this thesis, I argue that there needs to be a new way of thinking about film music to interlock ‘film’ and ‘music’ as one image of interpretation and experience to encourage new film music analysis. I call this new way of thinking the music-image. The music-image interlocks the semiotic systems of the visual image and musical codes, holding music and image together and in a relationship. This is a philosophical and theoretical model that encourages to change how music is thought to function in film rather than being an applied approach for any film music analysis. This is an original concept, but not an original name. I first came across the term ‘music-image’ in Claudia Gorbman’s *Unheard Melodies*, in which Gorbman uses the terms ‘music-image’ and ‘music-narrative’ to find a “mutual implication” between the two word pairings, ‘music’ and ‘image’, and ‘music’ and ‘narrative’.[8] The two words have been joined together to create one meaning that is different from the two words with two separate meanings.[9] I borrow this semantic terminology to construct a semiotic relationship between the two ontologies of aural and visual form. Benjamin Nagari used the term music-image, but to emphasise the individuation of the musical source, and thus again to make the separation of music from the film image visible in analysis. Nagari challenges film music's subordination (in Film Studies scholarship) to the visual image, arguing that music can *also* be an image (like the visual image), becoming "an image in its own right."[10]

---

[8] Claudia Gorbman, *Unheard Melodies: Narrative Film Music* (Bloomington; Indianapolis: Indiana University Press, 1987), 15.

Gorbman and Nagari use 'music-image' in very different ways, but there is still valuable insight from each, namely that music is not just important but intrinsic to film analysis and experience. The model I propose is inspired by Gilles Deleuze's argument in *Difference and Repetition* that ideas and concepts with prior meaning can be changed into new ones with the old concepts and ideas still existing. Applying this principle to film music, I would argue that if we attempt to 'isolate' music in a film (a common practice and analytical terminology), that music is automatically separated from the visual form of the film, and thus analysed in and of itself. But if aural and visual form is *combined*, *interlocked, and made interdependent* as a singular image form – that form creates a new stratum of meaning that exceeds or expands upon the insularity of musical and visual form. The interlocking of music and image is suggested by Gorbman in her discussion of the French filmmaker, Jean Cocteau who would take Georges Auric's music written for particular scenes in Cocteau's films and deliberately "apply the wrong music to the wrong scenes", in what Cocteau called "accidental synchronization."[11] "Accidental synchronization" unlocks the structure of a holistic image of film experience. The model of the music-image I propose takes an excursion from Gorbman's pivotal consideration of music and image as interlocked in mutual implication to construct the concept that underpins this thesis. I use this model, which I argue is a new or ‘strange’ way of thinking about film music to bridge the ontological gap between filmic and musical form, to bridge the methodological gap between Film Studies and Musicology in film music analysis.

---

[9] Kia Afra also uses 'music-image' in a similar sense to Gorbman (as well as the title for the article), as the "correlation" between music and image. See Kia Afra, "'Vertical Montage' and Synaesthesia: Movement, Inner Synchronicity and Music-Image Correlation in Alexander Nevsky (1938)," *Music, Sound, and the Moving Image* 9, no. 1 (2015): 34.

[10] Benjamin Nagari, *Music as Image: Analytical Psychology and Music in Film*, Research in Analytical Psychology and Jungian Studies Series (London; New York: Routledge, 2016), xiii.

[11] Gorbman, *Unheard Melodies: Narrative Film Music*, 15.

The second aspect of the music-image concerns the subjectivity of experience. In the analyses that structure this thesis, while I place myself at the centre of each film and film music reading, I attempt to foreground the radical subjectivity and changeability of filmic experience. Rather than searching for fixed meanings within sequences, I begin with the ‘meanings’ that attach to the phenomenon of film experience and that, in David Bordwell’s terms, are constructed through a new way of thinking.[12] Rather than applying a broad theory to locate a fixed meaning of a film or film sequence, I start with the film and attempt to reconcile form with the experience of the spectator. The analyses in this thesis are necessary to tease out the philosophical and theoretical framework of the music-image.

To conceptualise the experience of film as one audio-visual image and language, I begin by surveying the methodological landscapes of Musicology and Film Studies in Chapter 1, which I name ‘Film Music Studies’.[13] After this investigation of the field, I define the music-image in three parts: *what* is the music-image, *when* is the music-image, and *where* is the music-image. In this three-part model, I present close analyses of many film sequences and music excerpts from a wide range of film eras and traditions to unpack the ontology of the music-image, the temporality of the music-image, and the spatiality of the music-image.

---

[12] See, David Bordwell, and Noël Carroll, "Introduction," in *Post-Theory: Reconstructing Film Studies*, ed. David Bordwell and Noël Carroll (Madison, Wisconsin: The University of Wisconsin Press, 1996).

[13] I use the title of the book *The Oxford Handbook of Film Music Studies* to name the disciplinary field to honour the history and trajectory of film music scholarship. See, *The Oxford Handbook of Film Music Studies*, (Oxford; New York: Oxford University Press, 2014).

Chapter 2 is the outline of my music-image model and the first of three chapters dedicated to 'what' is the music-image. I employ the methodology of Gilles Deleuze to unveil the music-image model, however this is not a model solely based on Deleuzean theory, and this is not a thesis dedicated to Deleuze.

Chapter 3 continues to build an understanding of the music-image in relation to film diegesis.[14] I investigate whether film music can and does 'underscore' film narrative, time, and space, as film music scholarship suggests, through analysis of audio-visual harmony, dissonance, time, tone, and space.[15] Chapter 4 acts as the second part of Chapter 3, being an extended music-image analysis of one film, *Jaws* (Spielberg, 1975).

Chapter 5 moves into the second of three sections, the notion of 'when' the music-image exists in film. This chapter begins by investigating the philosophical question of whether film music can ever really be 'unheard' and moves to examples of when music-images occur in film.

Chapter 6 begins the final part of the three-part structure of the music-image model, with 'where' music-images exist in film. This chapter examines how film music acknowledged as existing within the film world or space is part of the construction of the film image.

With the exploration of what, when, and where the music-image is, Chapter 7 challenges the legitimacy of the music-image.

Approaching a new way of thinking about film music raises the question of how the films chosen for analysis are selected. To avoid going down the path of a 'Grand Theory' of film music, I have avoided stating ‘this [model] is my argument, and I will apply this [model] argument to these films’.[16] This research project attempts to highlight that the music-image is a new methodology with which to approach the analysis of any film and way of thinking about film, not just a few. In Kristin Thompson’s *Storytelling in the New Hollywood*, one of the reasons she gives for deciding on 10 particular films to analyse is because she “had to like the films well enough not to get tired of them in the course of the repeated viewings and close scrutiny that such analysis entails.”[17] In Thompson’s other book *Breaking the Glass Armor*, she points out that film scholars tend to analyse films to *confirm* the method.[18] Instead, *any* film should be analysed “…because it is intriguing…there is something about it which we cannot explain on the basis of our approach’s existing assumptions.”[19] Instead of choosing a few films or film composers to conform to a particular theory, I will analyse films that I have already seen, been intrigued by, and find suitable for the various film music topics explored in this thesis.[20] These tend to be films I watch or have watched in my spare time. For instance, most, but not all, of the films I analyse in this research project are from the USA as these are the films I typically watch; however, this does not limit the music-image to films from the USA.[21] Not every cinemagoer sees the same film, and not every cinemagoer sees the same film the same way, and this is the case wherever the film is made. The films and subsequent analyses should be seen as examples from my own subjective film experience.

---

[14] Diegesis is, according to Claudia Gorbman, "the narratively implied spatiotemporal world of the actions and characters." Claudia Gorbman, "Narrative Film Music," *Yale French Studies* 60 (1980): 195.

[15] This investigation of nondiegetic film music 'underscoring' the film scene is spawned from James Buhler and David Neumeyer's discussion of nondiegetic film music in *The Oxford Handbook of Film Music Studies*. See, James Buhler, and David Neumeyer, "Music and the Ontology of the Sound Film: The Classical Hollywood System," in *The Oxford Handbook of Film Music Studies*, ed. David Neumeyer (Oxford; New York: Oxford Univeristy Press, 2014), 32.

---

[16] For a definition and discussion of what Bordwell terms ‘Grand Theories’, see, Bordwell, "Introduction.". For consistency with Bordwell’s use of the term, I will use a capital ‘G’ and ‘T’ for ‘Grand Theory’ throughout this thesis.

[17] Kristin Thompson, *Storytelling in the New Hollywood: Understanding Classical Narrative Technique* (Cambridge, USA; London, UK: Harvard University Press, 1999), 44.

[18] *Breaking the Glass Armor: Neoformalist Film Analysis* (Princeton, New Jersey: Princeton University Press, 1988), 4.

[19] Ibid., 4-5.

[20] To be clear, this is a departure from Kristin Thompson’s precise parameters for her studies. Thompson has chosen films within a particular era and style of Hollywood cinema in *Storytelling in the New Hollywood*, whereas I have chosen films from my own film viewing and experience.

[21] Out of the roughly 130 films mentioned in this thesis, roughly two-thirds are from the USA.

My task is to inspire and promote analysis and thinking about film music in this new way to other cinemagoers and the films they choose to analyse or experience.

It is important to note that this thesis is not an historical account of film music, nor is it an analysis of film genres such as the film musical.[22] I will not pay particular attention to certain film composers or directors for reasons other than that these are the films I have chosen to analyse. There are certain aspects of film music and film theory that I will not be analysing, such as pre-existing music with intertextual meaning (for example, the use of Richard Wagner's *Ride of the Valkyries* in *Apocalypse Now* (Coppola, 1979)), documentary cinema, 'temp' music, and music of avant-garde cinema.[23]

Another crucial point to make is the acknowledgement that film music belongs to the larger family of film sound. Michel Chion, a pioneer in Film Sound Studies, uses the term 'sound-image' throughout *Audio-Vision*. In a similar sense to the music-image, Chion investigates the relationship between sound and film image, sound's relationship to the frame and the contents of the frame, and thus the film world.[24] When naming principles of sound's relation to the image, Chion includes music of the frame and in the frame.[25] Based on this shared principle and auditory nature, the music-image does not dismiss or neglect film as a sound image while discussing the music-image, especially as both diegetic sound and diegetic music were introduced when sound was introduced to film in the 1920s. This thesis attempts to reconcile film music with film, much like Chion reconciles sound with film (which includes music). Having said that, film music does bear a different ontology to other aspects of film sound such as voiceover narration, ADR (automated dialogue replacement), foley or ambience, and generates different meanings. Even though the music-image departs from

---

[22] I mention the film musical here as I reference the film musical extensively in Chapter 6, but it is not a genre study.

[23] This is not to say I am not interested in these missing topics; however, they are subjects for another volume.

[24] Chion, *Audio-Vision: Sound on Screen*, 68.

[25] Ibid.

Chion's audio-vision, the music-image would not exist without sound and exists alongside film sound.

Moving towards a new way of thinking about film music and what I call the music-image comes from a deep love and defence of film music. This thesis draws on my own film experience but is also a response to the convoluted history of Film Music Studies. Beginning with an overview of the existing academic literature and the schism between Musicology and Film Studies, I will build the case for a new model to bridge this gap.

Chapter 1: Towards a New Way of Thinking about Film Music

> "Music and image are an audio-visual unity to which violence is done when theoretically separated."[26]

Analysing the score for *The Best Years of Our Lives* (Wyler, 1946), Ben Winters came to this conclusion, finding that film music is an intrinsic part of the film image; he therefore argues against the theoretical separation of music and visual filmic analysis. Over the course of Film Music Studies' history, a split has occurred in methodological approaches to analysing film music, mainly from the competing perspectives of Musicology and Film Studies. In *Music as an Art*, philosopher and musicologist Roger Scruton argues that symphonic film music, particularly in Hollywood cinema, is a heavily criticised artform, often neglected due to unfavourable comparisons with concert-hall listening.[27] He argues further that when film music is extracted from film and is heard again in a concert-hall setting, something in the music feels absent.[28] This extraction also removes film music as a source of analysis from 'film' analysis, associating it with the analysis of concert-hall music and its distinctive musicological history. Scruton's notion of 'absence' in the above is not to suggest that something is absent from the music, but rather that the music is noticeably bereft of something that accrues from the film. This lacuna between film and music has been a problem in academic literature that many scholars have tried to address, to my mind with limited success, with the methodological approaches to analysing film music and image commonly viewed as separate entities and rarely brought into meaningful relation.

---

[26] Ben Winters, "Musical Wallpaper?: Towards an Appreciation of Non-narrating Music in Film," *Music, Sound, and the Moving Image* 6, no. 1 (2012): 51.

[27] Scruton, *Music as an Art*, 179.

[28] Ibid., 180.

When the two methodologies are bridged, I argue that music can be reconciled with the image to be analysed as a single, *whole* image. Theorists who have attempted this 'bridging' have leaned towards creating unified theories of musical score and visual form, addressing in the main the mechanics of how music and images *work*. However, such models rarely consider the spectator's experience as part of the puzzle. In this thesis, I argue that we need a new *way of thinking* about film music to interlock music and image in one analysis to consider a whole audio-visual *experience*. I term this ontological interlocking of music and image, quite literally, the music-image, which is a term I utilise for the duration of this thesis.

The chapter will begin by outlining the problematic and limited history of Film Music Studies as a discourse that has separated music from the image in formal analysis. Following this outline, the importance of bridging the methodologies of Musicology and Film Studies will be discussed as I explore the music-image as an irreducible 'whole'. This discussion will emphasise the importance of avoiding 'Grand Theories' that ultimately fall into the trap of separating music from the image; that is, in Film Music Studies, there is always the risk of employing a 'top-down' view of film music that formulates a theory without analysing the film that must validate that theory.[29] This chapter will instead propose an analytical 'bottom-up' method which begins with the analysis of film and music as form.

The problem with Film Music Studies in academic literature

Film music and its related scholarship have had a long history since the wide introduction of film sound in 1927. Film Music Studies has resulted from a vast and convoluted history much older than Film Studies as a discipline, its convolution partly a result of the many conflicting methodologies brought to the analysis. The three dominant existing methods of studying film music involve: 1. analysing the musical score; 2. analysing film music as a subsidiary commodity production of the film; and 3. Analysing the way in which music interacts with everything the spectator sees on screen. The fact that there are multiple approaches and methodologies brought to the study of film music reflects the general disagreement over a definitive, holistic method.[30]

---

[29] See, David Bordwell, "Contemporary Film Studies and the Vicissitudes of Grand Theory," in *Post-Theory: Reconstructing Film Studies*, ed. David Bordwell and Noël Carroll (Madison, Wisconsin: The University of Wisconsin Press, 1996).

Unlike the clean designation ‘Film Studies’, ‘film music’ involves music and film, each housing its own rich conceptual and methodological discourses. This is to say that the various studies of music and film tend to inhabit separate disciplinary fields, Musicology and Film Studies, respectively. This duality in form and expressive modalities of music and film has therefore drawn the analytical attention of scholars from both fields, often using different methods for examining ‘things’ that intersect and inter-relate across a vast field of signification. In the history of Film Music Studies, scholarly debate has unveiled how rare it is for Musicologists to incorporate the analytical methods of visual analysis, narrative, genre, or other methods commonly utilised in Film Studies; similarly, scholarly debate has noted the young discipline of Film Studies rarely incorporates music theory.[31]

---

[30] This is not to say that every scholar desires a definitive method to study film music, but the fact that there are multiple means that there is an inferred disagreement.

[31] James Buhler and David Neumeyer are both musicologists from The Butler School of Music at The University of Texas. Both Buhler and Neumeyer have written extensively (and sometimes together) on film music and are rare examples of music scholars attempting to incorporate film theory into their film music studies. Buhler is aware of the methodological gap, as, in his review of musicologist Annette Davison’s *Hollywood Theory, Non-Hollywood Practice*, he remarks the study of film music “...is now rapidly becoming a study of music in film. Until recently, research on film music routinely ignored the large and very rich literature of film studies....” See, James Buhler, "Review of Hollywood Theory, Non-Hollywood Practice: Cinema Soundtracks in the 1980s and 1990s, by Annette Davison," *Twentieth-Century Music* 3, no. 1 (2007): 145. This is a huge claim to track the history of Film Music Studies, but Film Studies is a relatively new academic discipline, as it has largely existed as a multidisciplinary field consisting of “literary studies, art history, musicology, and many other disciplines within the humanities.” See, Bordwell, "Contemporary Film Studies and the Vicissitudes of Grand Theory," 29. The study of music ‘in’ film was bound to occur in the early years of multidisciplinary film theory. The inclusion of Musicology in the field of Film Studies brought with it “music theory”, William H. Rosar argues, which film academics have criticised (namely musicologists) as unqualified to study film music. See, William H. Rosar, "Film Studies in Musicology: Disciplinarity vs. Interdisciplinarity," *The Journal of Film Music* 2, no. 2-4 (2009): 102. Musicologists are criticised because their “American” music

David Neumeyer argues that Film Music Studies became more proactive with film music analysis with the introduction of the VCR player in the 1980s, which provides the ability to record television and play back films. However, the study of film music has been around since scholars have been able to interact with film music.[32] Kurt London's *Film Music: A Summary of the Characteristic Features of its History, Aesthetics, Technique, and Possible Developments*, was the first book published on film music, in 1936, only nine years after the wide introduction of film sound in 1927, although it is rarely cited within the field.[33] There are many works on film music included in larger volumes devoted to film as a whole published earlier than London's book, and although Leonid Sabaneev's *Music for the Films* is regarded as the first major work entirely devoted to film music, published a year earlier in 1935, it is "a handbook for composers and conductors", unlike London's book which is the first study of film music.[34]

Theodor Adorno and Hanns Eisler's controversial yet profound monograph, *Composing for the Films*, was the first notable entry to criticise the practice of composing film music. Adorno and Eisler introduce the influential idea that film music is a manipulated, reproduced, and therefore consumer's art, lacking any real aesthetic value.[35] Eisler (a film music composer in Hollywood in the 1940s) and Adorno effectively dismiss film as 'art'. Instead, they argue that a film's technical elements are commodities of a cultural object that serves mass production and reproduction. Eisler argues: "The function of music in the cinema is one

---

theory draws a line between the study of music in film and study of film music. It is unclear (because of this line drawn) how separate the methodologies of Musicology and Film Studies are in film music analysis, and how much music theory and film theory are included or excluded in analysis. However, the fierce scholarly debate between musicologists and film scholars on this topic of film music analysis methodology shows both parties are approaching film music analysis from opposing scholarly traditions.

[32] David Neumeyer, "Overview," in *The Oxford Handbook of Film Music Studies*, ed. David Neumeyer (Oxford; New York: Oxford University Press, 2014), 3-4.

[33] See Kurt London, *Film Music: A Summary of the Characteristic Features of Its History, Aesthetics, Technique; and Possible Developments* (London: Faber & Faber, 1936).

[34] Scott Goddard, "Review of Music for the Films, by Leonid Sabaneev," *Music & Letters* 17, no. 2 (1936): 162. See also, Leonid Sabaneev, *Music for the Films: A Handbook for Composers and Conductors*, trans. S. W. Pring (London: Pitman, 1935).

[35] Hanns Eisler, *Composing for the Films* (London: Dennis Dobson LTD, 1947), x.

aspect - in an extreme version - the general function of music under conditions of industrially controlled cultural consumption."[36] This position was revolutionary for the study of film, and challenged the status and validity of Film Music Studies for decades. Film Studies was in full swing by the mid-1960s, but Film Music Studies remained largely silent until the 1970s.

It was not until Roy Prendergast reinitiated interest in Film Music Studies with his seminal work *Film Music: A Neglected Art,* which offered a nostalgic view of film music as an 'art' interpreted wrongly, that Eisler and Adorno's position, which held sway for decades within the disciplines of Film Studies and Musicology, was challenged.[37] Prendergast directly disagreed with Eisler and Adorno, offering a broad survey of Hollywood film music aesthetics. While a significant and courageous intervention, there were nonetheless early signs of limitations in his analytical methodology. Prendergast offered readings of his appreciation for film music, but openly described music merely as 'background filling', agreeing with famous classical composers such as Aaron Copland that film music could be extracted and performed, and thus *exist*, in isolation from the visual image.[38] Even though Prendergast's work is fundamentally a study of film music's history, aesthetics and musicological techniques, his use of the word 'accompany' to describe the score's function restricts the possibilities of this function.[39] And although influential, I would argue that his contribution was limited because it galvanized a field of musicological analysis that resisted analysis of an integrated, holistic 'film music image'.

Claudia Gorbman's book *Unheard Melodies*, released a decade after Prendergast's, added a new scholarly treatment to film music. As the adjective 'unheard' in the title suggests, she argues for music's utmost importance within the growing field of Film Sound Studies.[40] Michel Chion's several publications exemplify this field, which include his *Audio-Vision: Sound on Screen; Film, a Sound Art;* and *The Voice in the Cinema*. As these book titles suggest, Chion attempted to overturn the visual bias associated with Film Studies and emphasise film as an audio-visual medium. Gorbman approached film music from a narrative perspective, arguing that "ultimately it is the narrative context, the interrelations between music and the rest of the film's system, that determines the effectiveness of film music."[41] Reconciling music with the image reignited the interest of film scholars who were initially afraid to discuss the musicological aspects, feeling more comfortable analysing it in relation to other elements of film form. Gorbman essentially argued that music functions with "pure musical codes" when heard independent of any other activity; for example, when listening to a Bach Fugue on a CD.[42] This form of listening refers only to the musical structure for a source of meaning or analysis, generating musical discourse.[43] Film music, Gorbman argues, operates within an image with its own visual and cultural codes, losing its "pure" status, and should not be analysed as such.[44]

---

[36] Ibid., 20.

[37] See Roy M. Prendergast, *A Neglected Art: A Critical Study of Music in Films* (New York: New York University Press, 1977).

[38] Ibid., 205.

[39] Ibid., vii.

The adjectives 'neglected' and 'unheard' used by Prendergast and Gorbman respectively sparked conversation in academic fields interested in the topic, namely Musicology and Film Studies, about 'what and why is film music neglected?' and 'what is unheard?'[45] These two works, and especially Gorbman's, opened the floodgates for multidisciplinary approaches which broadened but also convoluted the field. In the late 1980s and early 1990s there was a boom in critical works published, including Caryl Flinn's *Strains of Utopia*, Kathryn

---

[40] Neumeyer, "Overview," 3.
[41] Gorbman, *Unheard Melodies: Narrative Film Music*, 12.
[42] Ibid., 12-13.
[43] Ibid.
[44] Ibid., 13.
[45] Neumeyer, "Overview," 3.

Kalinak's *Settling the Score*, Royal S. Brown's *Overtones and Undertones*, and George Burt's *The Art of Film Music*. However, the critical positions of these authors did not fully align with Prendergast or Gorbman.

Caryl Flinn's book was a response to the works of both Prendergast and Gorbman. Flinn claimed that using the word 'neglected' is overstating the case. She suggested that film music should be removed from the film before it could be critically analysed, and that only then could 'film music' be classified as an artform.[46] Her book does not discuss music's interaction with the image, which runs against Gorbman's first principle. Flinn accepts the dichotomy of high and low art (similar to Adorno and Eisler's position), with classical music in a concert hall setting constituting high art and film music constituting low art.[47] Flinn argues that music for an aesthetician has a 'utopian' function of being 'pure', and for others only as an emotional expression, but still they appreciate the fullness of experience that it offers or its ability to "return them to better, allegedly more 'perfect' times and memories."[48] The fundamental problem with this style of analysis is that it can be used for any form of music outside of film, and inclines toward the far more problematic position that music is *only* 'pure'. The book essentially addresses the object of film music as if the film existed in isolation from, and prior to, the music; as an accompaniment rather than part of, and thus intrinsic to, the image.

Kathryn Kalinak's *Settling the Score* extends Prendergast's argument of a neglected art and counters Flinn's. For Kalinak, film music is a "fundamental part of the filmic experience."[49] She aligns with Gorbman's position that film music accompanies narrative, but in the main addresses the visual image separately from the music. Her analysis of the opening title cards of *Vertigo* (Hitchcock, 1958) cannot help but provide a musicological reading of music *accompaniment*: "first...hear it as music; then it becomes necessary to analyse it as a part of a larger construct, that is, to analyse film music in conjunction with the image and sound track it accompanies."[50] Even though the intent of analysing film music as part of the image is present, ironically there is still the separatist approach endorsed by Flinn.

---

[46] Caryl Flinn, *Strains of Utopia: Gender, Nostalgia, and Hollywood Film Music* (Princeton, New Jersey: Princeton University Press, 1992), 3-4.

[47] Ibid., 4.

[48] Ibid., 9.

[49] Kathryn Kalinak, *Settling the Score: Music and the Classical Hollywood Film*, ed. Donald Crafton David Bordwell, and Vance Kepley Jnr, Wisconsin Studies in Film (Madison, Wisconsin: The University of Wisconsin Press, 1992), xiii.

George Burt's *The Art of Film Music* presents itself as a textbook for filmmakers and composers for how to approach composing film music as something distinct from concert-hall music. Burt labels film music as a story (narrative)-based phenomenon, like Gorbman, and discusses the way in which film music guides the spectator. Burt elaborates on this relationship:

> "...music being a temporal art, an art that takes place in time (as does film), it can have an enormous impact on the pacing of events, moving things along when needed, dwelling on something that requires attention, accenting this or that instant or event to help bring out the various connections and divergent points of view…[it is] difficult to recall a dramatic film...that does not employ music to dramatic advantage to some extent."[51]

However, it is unclear what Burt means by 'art', and how he frames his definition of temporality in relation to music and image. He explicitly aligns film music with the function of storytelling, and in several chapters analyses a film's sheet music (musical composition on paper), marking cues in the film related to narrative points. His examination of Hugo Friedhofer's score for *The Best Years of our Lives* (Wyler, 1946) is indeed a detailed analysis, but Burt stops short of acknowledging that the score is written *for the film*, and that the relational aspect of film music had significant import.[52]

---

[50] Ibid., 14.

[51] George Burt, *The Art of Film Music* (Boston: Northeastern University Press, 1994), 4.

This initial scholarly boom in the study of film music after Gorbman's influential work left the field wide open. By the end of the 1990s, there was enough academic literature for future scholars to embark on "thorough critical reviews", such as writing critical histories.[53] Stephen Wright and Jeannie Pool, researchers in the field, observed in the early 2000s that Film Music Studies had exploded.[54] Case studies proliferated and essay anthologies and textbooks of film music history were written during this period, with key writers such as James Buhler, Caryl Flinn, K. J. Donnelly, Robyn Stilwell, Mervyn Cooke, and James Wierzbicki. There are even journals devoted to the field such as *Music and the Moving Image*; *The Journal of Film Music*; *Music, Sound, and the Moving Image*; and *The Soundtrack.*[55] Film Music Studies had widely established itself in the humanities.

What Prendergast set out to achieve with his seminal work, calling the art of film music 'neglected', had in large part been achieved. And yet, and perhaps ironically, in these often-conflicting methodologies, the analysis of film music had become disordered. As the musicologist Roger Scruton stated, when listening to film music in a concert hall setting, "something feels missing". I could say the same of a great deal of contemporary film music analysis, that is, that the film is missing.[56] One of the major volumes to emerge from the field within the last decade, *The Oxford Handbook of Film Music Studies*, similarly neglects this fundamental problem. Marianne Kielian-Gilbert, who is the only author in the handbook to identify this problem, argues that “listening in film becomes a problem for analysis.”[57] Kielian-Gilbert recontextualises music in a cinematic setting and examines how it interacts with experiences of the combined temporal unfolding of music and image, and investigates whether this differs in different media settings.[58] Should then a film or a film sequence with music be analysed with all senses instead of just ‘listening’ or ‘watching’? Analysing the same film sequence with different sensory tools such as listening, watching, or experiencing (combining all senses) ultimately changes the results of the analysis.

---

[52] See ibid., 143-66., for a detailed analysis of the film score for *The Best Years of our Lives*.

[53] Neumeyer, "Overview," 4.

[54] Ibid.

[55] Emilio Audissino, *Film/Music Analysis: A Film Studies Approach*, ed. K. J. Donnelly, Palgrave Studies in Audio-Visual Culture (Cham, Switzerland: Palgrave Macmillan, 2017), 2-3.

[56] Scruton, *Music as an Art*, 180.

The diversification of methodologies that separate the analysis of music from image and image from music remains problematic in Film Music Studies. Even though film scholars such as Gorbman, Kalinak, and Flinn launched Film Music Studies after Prendergast, the field has attracted multiple disciplines. For example, books on film music compiled for this research project, published between 2000-2020, were in the main located in the music library at The University of Sydney rather than the Film Studies section of the main university library, and the primary contributors to such volumes were musicologists.[59] Conferences on film music tend to be organised and hosted by Music departments rather than Film Studies departments.[60] When scholars from Film Studies contribute to the discourse, they are often critiqued for their lack of musicological background. For example, film scholar Anahid Kassabian’s *Hearing Film: Tracking Identifications in Contemporary Hollywood Film Music*, was critiqued by musicologist James Wierzbicki for her perceived lack of musical terminology, which Wierzbicki attributes to her 'benighted' understanding of music as a film scholar.[61] Wierzbicki's more general disposition toward the musicological capacity of the film scholar results in a reluctance among film scholars to discuss the musicological elements of film music.[62] Emilio Audissino points out that a book such as *The Cinema of Steven Spielberg: Empire of Light* (a book dedicated to the *films* of Steven Spielberg with a chapter per film), John Williams, Spielberg's frequent film composer, is only mentioned in passing a handful of times.[63] In extensive analyses of *Jaws* (Spielberg, 1975) and *Close Encounters of the Third Kind* (Spielberg, 1978) (a film with music at the heart of the narrative), the music is only ever acknowledged but not discussed.[64]

---

[57] Marianne Kielian-Gilbert, "Listening in Film Music: Music/Film Temporality, Materiality, and Memory," in *The Oxford Handbook of Film Music Studies*, ed. David Neumeyer (Oxford; New York: Oxford University Press, 2014), 501.

[58] Ibid.

[59] To be clear, this claim cannot be representative of all university libraries, but my experience at The University of Sydney is still striking. The sources needed to conduct research for this research project were mainly sourced from The Conservatorium of Music, The University of Sydney’s main music campus. Even works by film scholars associated with Film Music Studies, were found in the University’s music department.

[60] For example, the conference for *Music and the Moving Image*, a top journal for publishing works in Film Music Studies, was held at the Steinhardt School of Music, New York University, and the *Music for Audiovisual Media* conference was held at the School of Music, University of Leeds. See, Audissino, *Film/Music Analysis: A Film Studies Approach*, 3.

Closing the (methodological) gap between Music(ology) and Film (Studies)

I recognise that the opposing disciplinary studies, namely of Musicology and Film Studies, have promoted the thinking of film and music as two separate and competing entities.[65] This analytical separation breaks apart aural and visual stimulus when analysing a primarily audio-visual phenomenon.[66] Following Gorbman, I argue that film music cannot function independently of a range of other activities, that is, the music of film music can never present as, to use Gorbman's terminology, a *pure musical code*.

---

[61] See James Wierzbicki, "Anahid Kassabian: Hearing Film: Tracking Identifications in Contemporary Hollywood Film Music," *Journal of Film Music* 1, no. 4 (June 19, 2001). Anahid Kassabian was the co-founder of *Music, Sound, and the Moving Image*.

[62] As seen so far in this chapter, this is also the case with, and is countered by, the reluctance of musicologists to discuss film theory.

[63] See, Audissino, *Film/Music Analysis: A Film Studies Approach*, 4., where Audissino raises the lack of mention of John Williams or his music in a book dedicated to the films of Steven Spielberg. Upon further investigation, I discovered in the chapter on *Jaws* (Spielberg, 1975), John Williams' score is only mentioned twice: once for Williams winning an Academy Award, and another for Steven Spielberg acknowledging that "Williams' score was central to the movie's success." Ironically, his score was not central to this chapter. See Nigel Morris, *The Cinema of Steven Spielberg: Empire of Light* (London; New York: Columbia University Press, 2007), 45-46.

[64] In the chapter on *Close Encounters of the Third Kind* (Spielberg, 1977), Williams is never mentioned, nor is his music in any detail. See, *The Cinema of Steven Spielberg: Empire of Light*, 8-19.

[65] Audissino, *Film/Music Analysis: A Film Studies Approach*, 35.

[66] Separatism in film music scholarship is when the 'aural' and the 'visual' image are separated in the analysis of film music. An example of this is micro-analysing the film score without considering the visual image, pointing instead to cues in the sheet music. Another example is using language that suggests the aural and visual are separate entities, combined, complementing, or accompanying the other, instead of 'experienced' at the

Since theorists such as Prendergast and Gorbman identified the scale and coordinates of the problem, several film sound scholars, including Michel Chion, Rick Altman, and K.J. Donnelly, have addressed the analytical schism reflected in the separate disciplines of Film Studies and Film Music/Musicological Studies.[67] Chion, in a radical gesture, calls film a *sound art*, and addresses all sound as part of the ontological field of the image.[68] Chion writes:

> "There are many consequences of this negative claim I have been making that 'there is no soundtrack.' First, it necessarily explodes any pretense of establishing an overall theory of film composed of two complementary elements, image and sound, and instead leads us to an incomplete and fluctuating model of cinema that does not allow for simply transposing the technical model (one channel for images and one channel

---

same time as one (analysed) image. Audissino in his book *Film/Music Analysis* seems to be the only scholar who coins the term 'separatism', but it has been a concept many scholars have critiqued since the early 1990s starting with Nicholas Cook's *Analysing Musical Multimedia.*

[67] Rick Altman in the 60th edition of *Yale French Studies* published in 1980, titled 'Cinema/Sound', introduces the edition as a reconsideration for film sound by scholars in Musicology and Film Studies. This reconsideration comes from Altman highlighting a visual bias in Film Studies, or as he articulates, there is a 'hierarchy' of film form. He writes: "Instead of treating sound and image as simultaneous and coexistent, the historical fallacy orders them chronologically, thus implicitly *hierarchizing* them. Historically, sound was added to the image; [therefore] in the analysis of sound cinema we may treat sound as an afterthought, a supplement which the image is free to take or leave as it chooses." He further argues: "Another argument was called for, a strategy tied not to film's history but to the medium's very essence. Thus, the ontological fallacy was born. The version of the ontological fallacy regularly applied to cinema claims that film is a visual medium and that the images must be/are the primary carriers of the film's meaning and structure." See Rick Altman, "Introduction," *Yale French Studies* 60, no. Cinema/Sound (1980). K.J. Donnelly, in 2014, claims there is still a separation of sound and image in scholarship and refers to Gestalt Psychology as a method to join together sound and image stimuli. Donnelly writes that the separation is due to the assertion that film is a visual medium: "We should remember that it is problematic to deal with sound and image as separate entities—and thus with the commonplace assertion that film is a visual medium. The synergetic effect of the sound and image amalgam is the heart of cinema, and indeed it is why its illusory character can be so convincing." See K.J. Donnelly, "Synchronisation," in *Occult Aesthetics* (Oxford: Oxford University Press, 2014), 27.

[68] See, Michel Chion, *Film, a Sound Art*, trans. Claudia Gorbman (New York: Columbia University Press, 2009).

> for sounds, distinct yet parallel) and where a radical break is produced between the technical level and the levels of perception, discourse, effects, and theory."[69]

For Chion, sound and visual material operate as a single integrated channel rather than 'complementing' elements. This way, cinema can be acknowledged as a 'sound art'.[70]

Grand Theories and 'communications' models

A great deal of the 'separatist' position on Film Music can be attributed to the imposition of what David Bordwell calls "Grand Theories". Bordwell defines a 'Grand Theory' as an "abstract body of thought" and "ethereal speculations" that impose risky 'top-down' views rather than analytically motivated 'bottom-up' analysis.[71]

Emilio Audissino argues that Gregg Redner's *Deleuze and Film Music* is an example of a "risky" Grand Theory applied to film music.[72] I would argue that Redner has found an encouraging way of bridging music theory and film theory. Redner argues that music and image are separate entities that can be bridged by a Deleuzian theoretical framework. After an extensive historiographical analysis, he concludes that there is a theoretical/methodological gap between musicological theory and film theory. The several chapters of the book then attempt to 'bridge' the methodologies through a reading of Deleuzean theory. Each chapter applies a different Deleuzean 'theory' to a film (each film is culturally and musically different from the previous). After contextualising each film composer's philosophy and approach to the screen, Redner applies the Deleuzean theory and proceeds to analyse the film/score. In the first chapter dedicated to a film, Jean Vigo's *L'Atalante* (1934), Redner considers film music's aesthetic plane of composition as involving 'sensation'.[73] The philosophy of sensation in art is not original to Deleuze, as he is inspired by the artistic works of Francis Bacon in *Francis Bacon: the logic of sensation*, and biology.[74] Sensation, according to Deleuze and Guattari, is the compound of percepts and affects in art:

---

[69] Ibid., 229-30.

[70] In his book *Audio-Vision*, Chion mentions that music can directly express participation in the feeling of a scene, which he called 'empathetic music', that can take on "the scene's rhythm, tone, and phrasing." See, Chion, *Audio-Vision: Sound on Screen*, 8. The use of this language of 'empathy', however, comments on and engages a plurality of subject matter. Chion situates music as being its own autonomous entity that comments on the medium of film, rather than being part of it. Even when this separation is not intended, there is a vulnerability attached. Alternatively, Rick Altman, a scholar also championing film as a sound art, cannot see sound in the cinema without being separate from it: "the playback permitted immediate capitalization on the sound film's fundamental lie: the implication that the sound is produced by the image when in fact it remains independent from it." See, Altman, "Introduction," 6. Altman considers sound on the same channel as the visual, but creates a gap between them in the process. The methodological gap between Musicology and Film Studies has created a phenomenological gap between film music and film.

[71] Bordwell argues that Film Studies scholarship can operate without employing this "indispensable frame of reference for understanding all filmic phenomena." See, Bordwell, "Introduction," xiii. Instead, recent scholarship employs an inquiry that moves easily from evidence to more 'general arguments and implications'. See, ibid. For Bordwell, Grand Theories are the 'top-down' risk that a theory comes before a film is selected for analysis, so the film selected conforms to that theory. Kristin Thompson agrees with the risk of scholars using this method of analysis to test their theory, choosing a film to confirm its validity. See, Thompson, *Breaking the Glass Armor: Neoformalist Film Analysis*, 4. For example, theorising Hollywood as a patriarchal system orchestrated through the male gaze promotes the analysis of films based on these gaze dynamics. This example is a top down view as it searches films for oedipal complexes, giving it a limited number of films to analyse when not all Hollywood films conform to this theory, weakening the initial argument.

[72] Audissino, *Film/Music Analysis: A Film Studies Approach*, 52.

> "It is independent of the creator through the self-positing of the created, which is preserved in itself. What is preserved - the thing or the work of art - is a bloc of sensations, that is to say, a compound of percepts and affects."[75]

Percepts are not perceptions, and affects are not affections; percepts and affects (and sensations) are:

---

[73] Gregg Redner, *Deleuze and Film Music: Building a Methodological Bridge Between Film Theory and Music* (Bristol; Chicago: Intellect, 2011), 33.

[74] See Deleuze's developed philosophical concept of 'sensation' inspired by the paintings of Francis Bacon in, Gilles Deleuze, *Francis Bacon: The Logic of Sensation*, trans. Daniel W. Smith (London; New York: Continuum, 2003).. Deleuze, and colleague, Felix Guattari, in their book, Gilles Deleuze, Guattari, Felix, *A Thousand Plateaus: Capitalism and Schizophrenia*, trans. Brian Massumi (Minneapolis; London: University of Minnesota Press, 1987)., developed the philosophical concept of the 'rhizome', based on the botanical rhizome, different to roots and radicles, a building block, or stem, if you will, can branch off into different biological strands and life, but also ideas and concepts, repeating the foundation of the rhizome. The metaphor of the rhizome is simply the tree is the rhizome, and the rhizome is the tree. "Any point of the rhizome can be connected to anything other, and must be." See ibid., 7. The rhizome is essential in understanding 'sensation', as I address later.

[75] Gilles Deleuze, Guattari, Felix, *What is Philosophy?*, trans. Hugh Tomlinson and Graham Burchell (New York: Columbia University Press, 1994), 164.

> "…*beings* whose validity lies in themselves and exceeds any lived...exist[ing] in the absence of man because man, as he is caught in stone, on the canvas, or by words, is himself a compound of percepts and affects. The work of art is a being of sensation and nothing else: it exists in itself."[76]

Redner develops the philosophical concept by Deleuze and Guattari before treating the score to *L'Atalante* to make it clear that the film score is a "pure product of sensation" which allows the free movement between nondiegetic and diegetic spaces.[77] Redner further creates a 'wheel' of percepts, affects, diegetic and the nondiegetic 'rhizomes' in a visual diagram, relating to each other through the bloc of sensation around which they revolve.[78]

The outcome is thus to use a Deleuzean theory to proffer a new form of musical experience that extends beyond the conventions of vibration and frequency to incorporate emotion in the sensation of human interaction. For Redner, "it becomes possible to address not only musical issues but also narrative, mise-en-scene and dramatic issues simply by moving freely through the concepts addressed above."[79]

My argument is that, while the notion of sensation in art is valued and well-conceived in Redner's analysis, the relationship music has to filmic *form* is loosely realised.[80] This invariably leads to promoting film music's independence from the image: "...film music is at its best when...it does not mimic or 'Mickey Mouse', resorting to mere mimicry (imitating something). Perhaps film music functions most authentically when its functions are independent but also collaborative in their interaction with the narrative."[81] This way of

---

[76] Ibid.

[77] Redner, *Deleuze and Film Music: Building a Methodological Bridge Between Film Theory and Music*, 34.

[78] Redner uses the concept 'rhizome', referenced in my earlier footnote, as the diegetic and nondiegetic act as the rhizomes that create the blocs of sensation that are percepts and affects. See ibid., 35.

[79] Ibid., 37.

[80] Audissino criticises Redner for referring to film as 'mise-en-scene', but the main issue, I argue, is how Redner relates the *music* to film. See, Audissino, *Film/Music Analysis: A Film Studies Approach*, 52.

[81] Redner, *Deleuze and Film Music: Building a Methodological Bridge Between Film Theory and Music*, 39.

thinking also promotes the *analysis* of 'independence' in the functionality of the score, or the visual image, when film is simplified to narrative. The 'collaboration' in the interaction of music and image does not consider a sensation that is interlocked from 'existing' sensation that Deleuze and Guattari refer to in their reading on preservation and memory:

> "...every work of art is a monument, but here the monument is not something commemorating a past, it is a *bloc of present sensations that owe their preservation only to themselves and that provide the event with the compound that celebrates it*...We write not with childhood memories but through blocs of childhood that are the becoming-child of the present. Music is full of them."[82]

Redner's application of Deleuzean theory to film music analysis is encouraging, however his reference to Film Studies concepts, terminologies, and methodological principles are inept and muddle the principle of bridging music theory and film theory.[83]

Although Redner's idea of how to bridge methodologies is intriguing and has enormous potential, it nonetheless demonstrates a tendency to apply a concept or set of concepts onto a base level of filmic materials. One can then ask: would this functioning of the Deleuzian paradigm apply to all films? If not, what makes one film more suited to this schema than another?[84]

---

[82] In Deleuze, *What is Philosophy?*, 168., the idea of 'preservation' and memory is similar to what Bazin philosophised in his essay *The Ontology of the Photographic Image* about the Ancient Egyptians preserving human 'existence' through the body in mummification, and on top of preserving the memory of others and places, Deleuze argues that the perceptions and affections of these preservations have been wrested blocs of affects and percepts, in that they exist in themselves.

[83] For example, in one section of the chapter on sensation, Redner criticises Claudia Gorbman's *Unheard Melodies* for denying "on one level the very essence of the musical score, which is to be an auditory participant in the filmic universe." See, Redner, *Deleuze and Film Music: Building a Methodological Bridge Between Film Theory and Music*, 31. However, Redner fails to capitalise on his own critique, as ironically he does not extend past the film music's narrative input in each film he surveys. He argues, "Claudia Gorbman['s] analyses often result in diminishing the score's place as a coequal within the greater being that is the film," but is not that the point of sensation? See, ibid., 45. That it coexists and 'coequals' alongside the visual? The blocs of percepts and affects function within themselves both aurally and visually to create an interlocked sensation. Redner states "By embracing the concept of sensation for the study of film music, we are left to hear the score in its purest form," but Gorbman argued against the idea that film music functions in pure musical codes, or as Redner puts it, 'in its purest form', as the music serves a context, it is not functioning independent of any other activity. See, ibid. This contradicts Redner's stance on the methodological 'gap' from multiple disciplines working in film music scholarship. The main problem with Redner's book is not this gap, but its foundation of music and image separatism, a result of utilising the Grand Theory approach to bridge the methodological gap between Musicology and Film Studies when analysing film music.

As well as Grand Theories, the 'communications model' is another separatist approach to film music analysis, in that a message is being communicated. The three components of the model are the "sender, medium, and receiver," with the main activity being communicating "a message from sender to receiver through the medium."[85] The model is more concerned with what is being communicated rather than how and why in an analytical way. Not considering how and why a message is communicated prioritises interpretation over analysis of what film music is. When music and image are combined, it creates a meaning, according to this model, instead of analysing what this relationship between music and image is. An example of this is Bordwell describing what music communicates when 'added' to the image: "If I want you to feel sad about what is onscreen, I can insert sad music."[86] This favours the visuals and denotes music as added value with the sole purpose of communicating an emotion. Dominique Nasta in her book *Meaning in Film: Relevant Structures in Soundtrack and Narrative* also argues that films are principally visual, moreover she favours interpretation over analysis, where music could communicate content, promoting music's separation from the image.[87] Nicholas Cook's *Analysing Musical Multimedia* aimed to debunk the analytical favouritism of music functioning only with pure codes. Cook argues: "Outside the imagination of aestheticians and analysts, *music never is alone*."[88] Although

---

[84] This is not to say that all separatist approaches to film music are Grand Theories. Adorno and Eisler's *Composing for the Films* is both separatist and a bottom-up view on film music analysis. There is no Grand Theory attached to their critique, rather they are responding to the state of Hollywood film music with their Marxist assessment. They maintain their autonomy and critical consciousness capable of retaining that ideological dialecticism. See, Gerald Sim, "The Other Person in the Bathroom: Mixed Emotions about Cognitivist Film Music Theory," *Quarterly Review of Film and Video* 30, no. 4 (2013): 321.

[85] Thompson, *Breaking the Glass Armor: Neoformalist Film Analysis*, 7.

[86] David Bordwell, *Poetics of Cinema* (New York; London: Routledge, 2008), 125.

[87] Dominique Nasta, *Meaning in Film: Relevant Structures in Soundtrack and Narrative* (Berne; Berlin: Peter Lang, 1991), 43.

[88] Nicholas Cook, *Analysing Musical Multimedia* (Oxford: Clarendon Press, 1998), 265.

Cook attempts combining the analysis of music and image across different multimedia, he begins with the notion that music and image are two separate entities that *combine* to "channel communication" and 'participate' in a *construction* of sorts.[89] Cook argues that music *interacts* with image to create three different effects: 'amplification', 'projection', and 'dominance'.[90] However, there is a problem in the methods he chooses to declare a lack of 'separation'. Throughout his book, Cook's thesis is an example of a theorist *against* separatism but uses a communications model that promotes a musical accompaniment that only exists if the music and visual were initially separate.[91] What can be seen by these attempts at bridging the methodologies of Musicology and Film Studies to analyse film music is that they have not been able to escape the separatist notion of music and image as elements that are *combined* to create new meaning through their connection. What about film music as an absolute image that can be analysed? There needs to be a language used that analyses music and image simultaneously to avoid this analytical and methodological separatism. There needs to be a language for the analysis in the bridge itself.

Analysing film music through film experience

Beginning film music analysis with film experience – or what I am proposing is a 'music-image' – might begin with the formulation of a problem (which I've attempted to do in this chapter), and an attempt to solve this problem through analysis, interpretation, reflection, and empirical research.[92] I find this method admirable as the operation of any theory is predicated on the response to a problem in cinema. To get to this response, the theorist is required to *experience* a film. For Nicholas Cook 'experiencing' music is a psychological activity.[93] On a basic level the listener understands the music based on a personal experience. The listener can afterward investigate why they may have experienced the music in a particular way, but this analysis follows from an initial experience. In the model of film music analysis I would propose, the spectator must experience a film (or many films) to formulate a response, and out of that and other processes a theory may emerge. Grand Theories, for Bordwell and Carroll (with whom I agree) are the inverse: created through speculation in isolation from experience.

---

[89] Ibid., 261.

[90] Amplification is when the music parallels what is happening visually, what many scholars call 'mickey-mousing'. Projection is music adding something to the visual to extend its original meaning. Dominance is when music antagonises, resists, and struggles against the meaning presented by the visuals. See, ibid., 112.

[91] For example, in Nicholas Cook's outline for the 'Three Basic Models' of multimedia, the use of the words 'conformance', 'complementation', and 'contest' to express the meaning music and multimedia create when combined implies that music has a role to combine with the image. If music combines, it can be analysed separately when not combined which defeats the purpose of one whole image. See, ibid., 98.

[92] Bordwell, "Introduction," xiv.

When the spectator experiences a film, they are not sitting in a room with the live orchestra performing the score, reading the flow of music on a sheet (not even classical music concerts supply the audience with the score). Rather, the visual stimulus is the film with an aural component. The spectator cannot see the flautist blow into their flute to perform the 'Concerning Hobbits' theme for *The Lord of the Rings: The Fellowship of the Ring* (Jackson, 2001) or Itzhak Perlman using vibrato on his violin while performing 'Jewish Town' for *Schindler's List* (Spielberg, 1993); rather, the spectator *hears* a stimulus. As Stanley Cavell suggests in *The World Viewed*, a heard flute or a violin does not necessarily imply that the instrument (sound source) is physically present. Rather, the sound has been recorded in a distinct space to evoke its presence in another.[94] If a spectator were to watch the famous scene from *Gilda* (Charles Vidor, 1946) when Rita Hayworth is revealed onscreen for the first time, it is less paradoxical to say "that is Rita Hayworth" than hearing what you believe to be a violin playing while watching *Schindler's List* and saying "that is a violin playing vibrato". These examples are proof of the importance of an argument beginning with the analysis itself. If the claim were, “When experiencing a film the spectator can synaesthically see instruments play the film music,” there would need to be a search for a film that conforms to that claim, leaving no room for revision.[95] This approach is problematic from the outset, as what if a spectator does not confirm this position? It is a variable that depends on the individual response of a spectator. For instance, a spectator with little orchestral understanding may struggle to tell the difference between an oboe and a clarinet being played unless it was prompted diegetically in the example of Gabriel playing his oboe in *The Mission* (Roland Joffé, 1986). In this scene, the camera cuts between shots of Gabriel playing, wide shots of the jungle landscape, close-ups of natives closing in on Gabriel, and medium shots of the natives meeting in the same camera space as Gabriel (See Figures 1-4). When the spectator hears the theme ‘Gabriel’s Oboe’ playing in the shots where Gabriel is not present, the spectator may picture Gabriel playing his oboe to fill in the image.[96]

---

[93] Nicholas Cook, *A Guide to Musical Analysis* (London; Melbourne: J. M. Dent & Sons Ltd, 1987), 67.

[94] Stanley Cavell, *The World Viewed: Reflections on the Ontology of Film* (Cambridge, USA: Harvard University Press, 1979), 18.

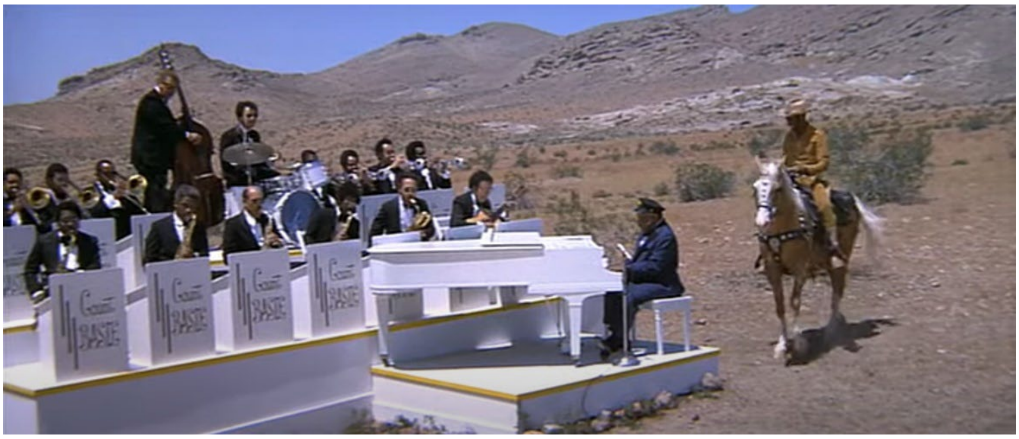
---

[95] Synaesthesia is the merging of senses that are not normally connected, for instance, stimulation of one sense can cause a reaction from another, like hearing colours or seeing sounds. See Afra, "'Vertical Montage' and Synaesthesia: Movement, Inner Synchronicity and Music-Image Correlation in Alexander Nevsky (1938).", for an in-depth analysis of Eisenstein’s example of synaesthesia in film music.

[96] It is not assumed that the spectator knows the name of the musical piece as listed in the film’s soundtrack. For instance, here, it is not assumed that the spectator knows that Gabriel is playing the oboe because they know that the piece of music is called ‘Gabriel’s Oboe’.

Figures 1-4. The musical theme ‘Gabriel’s Oboe’ plays ‘diegetically’ for the duration of these four instances, when Gabriel and his oboe are on and off-screen.

If the above argument were to be applied to another randomly selected film sequence such as the shower scene in *Psycho* (Alfred Hitchcock, 1960), the spectator may relate the famous cluster of shrieking accented chords played on the violin to a stabbing knife rather than visualise the violins. Mel Brooks played on this notion in *Blazing Saddles* (Mel Brooks, 1974) when Bart rides his horse over the hill to sweepingly melodic, heroic orchestral music that is supposedly nondiegetic (not located within the narrative space). Then the camera zooms out and pans to show that the orchestra pit is in fact in the same space; this is an audio-visual gag that questions the synaesthetic *appearance* of film music (see Figures 5-6). On the other hand, if the claim began with a formal analysis of a film or a group of films and identified a problem (to reiterate Bordwell and Carroll) when experiencing them, thoughts can be written down to conjure up an argument. This is the bottom-up approach to film music analysis. Instead of relying on existing structures such as Grand Theories to conjure meaning, the experience of the spectator is given priority, which relates to a certain scepticism associated with post-structuralist thought.[97] The analysis is ongoing and can be left to further criticism, whereas the Grand Theory leaves little or no room to amend. As Bordwell agrees: "Grand Theories will come and go, but research and scholarship will endure."[98]



Figures 5-6. Music accompanies Burt as he rides over a hill, seemingly nondiegetic, until the camera suddenly zooms out and pans to reveal an orchestra.

---

[97] Structuralism is the belief that "phenomena of human life are not intelligible except through their interrelations. These relations constitute a structure." Post-structuralism functions as a response to this, as it adds an "interest in their origins in relationships of power, or in the unconscious…It provides one manifestation of the skeptical stance of postmodernism by refusing any concepts of objectivity, reality, and truth." The bottom-up analysis model is then considered in this realm of rejecting 'objective' experience of a film. See Simon Blackburn. *Oxford Dictionary of Philosophy.* Third ed. s.v. "Structuralism." Oxford, UK: Oxford Univeristy Press, 2016., and *Oxford Dictionary of Philosophy.* Third ed. s.v. "Post-Structuralism." Oxford, UK: Oxford Univeristy Press, 2016.

[98] Bordwell, "Contemporary Film Studies and the Vicissitudes of Grand Theory," 30.

Eisenstein’s Vertical Montage and towards a new way of avoiding separatism

To avoid separatist conceptions, we must adopt terminology that does not conform to the idea. In this chapter, it has been argued that language that conforms to separatism is A + B + C, where A is the music, B is its mode of operation, and C is the film stripped of the music. Separatism believes that A + B + C (notice that each variable is isolated) = ABC (notice that each variable seems to be stitched together), but algebraically, this is incorrect, as A + B + C still results in A + B + C, or, a new variable, D.[99] Instead of the ABC model, the result needs to be a distinct and independent variable to analyse at face value, that being D, which would comprise A, B, and C operating simultaneously, and in constant relation. Audissino’s ‘fly in the soup’ analogy is quite helpful to explain this. A non-separatist conception would be to see film as an integrated system of elements, the soup, and the analysis would be something like chemical analysis, breaking down the soup into ingredients, with music constituting one of those ingredients. The separatist model views music as a fly in the soup. It is an insoluble element that does not blend in, but rather floats at the surface, not going deep in the soup solution.[100] An original variation on this is to visualise the elements of a film as a robot that, if any part of the robot is removed, the robot cannot operate, and each part cannot operate separately. The music can represent, for example, an arm of a humanoid robot. The robot needs the arm to operate as a fully functional humanoid robot, so it cannot be added to operate and it cannot be dismembered; that is, it is a closed system.

The great Soviet filmmaker Sergei Eisenstein, after making the film *Alexander Nevsky* (Eisenstein, 1938), attempted to analyse ‘The Battle on the Ice’ from the film as one audio-visual image. When he and composer Sergei Prokofiev prepared the scene, they attempted to match the music to the visual sequence, describing the preferred final audio-visual image as 'polyphonic' in structure, welding the music to the existing lines of visual montage into a 'super-structure'.[101] When analysing the sequence, Eisenstein referred to this as 'Vertical Montage', music's vertically synchronous relationship with the horizontal lines of existing montage. Eisenstein promoted this way of thinking in the following passage:

---

[99] In order to obtain ABC, the entities would need to be multiplied together like so, A x B x C. However, ABC still appears as separate entities stitched together, ‘ABC’, not a new, singular variable.

[100] Audissino, *Film/Music Analysis: A Film Studies Approach*, 55.

> "visualise it [the sequence] as two lines, keeping in mind the *whole* complex of a many-voiced scoring…matching both picture and music to the general, complex imagery produced by the whole…Unlike the horizontal succession of pictures in montage, there is a [simultaneous] 'super-structure' erected vertically equally in tonal length. Pieces of sound do not fit into the picture pieces in sequential order, but in simultaneous order."[102]

This is an original way of thinking and approaching the problem of analysing film music at face value. However, Eisenstein's model comes across as naïve as it does not consider anything musicological, for example, tonality, harmony, rhythm, and volume. It bases the geometry of the mise-en-scene with pitch only and draws a geometrical shape between them (see Figure 7).[103] Nonetheless, the theory is anchored on his own film, and he was in close collaboration with the composer Prokofiev, so perhaps this was the intention with his use of a limited analysis of musical form.

---

[101] Polyphonic structure in music consists of two or more opposing melodies played simultaneously, as opposed to monophony structure of one dominant voice. See Sergei Eisenstein, *The Film Sense*, trans. Jay Leyda. (New York: Meridan Books, 1957), 76-77.

[102] Ibid., 78.

[103] Figure 7's 'Vertical Montage' image was retrieved from Afra, "'Vertical Montage' and Synaesthesia: Movement, Inner Synchronicity and Music-Image Correlation in Alexander Nevsky (1938)," 42.

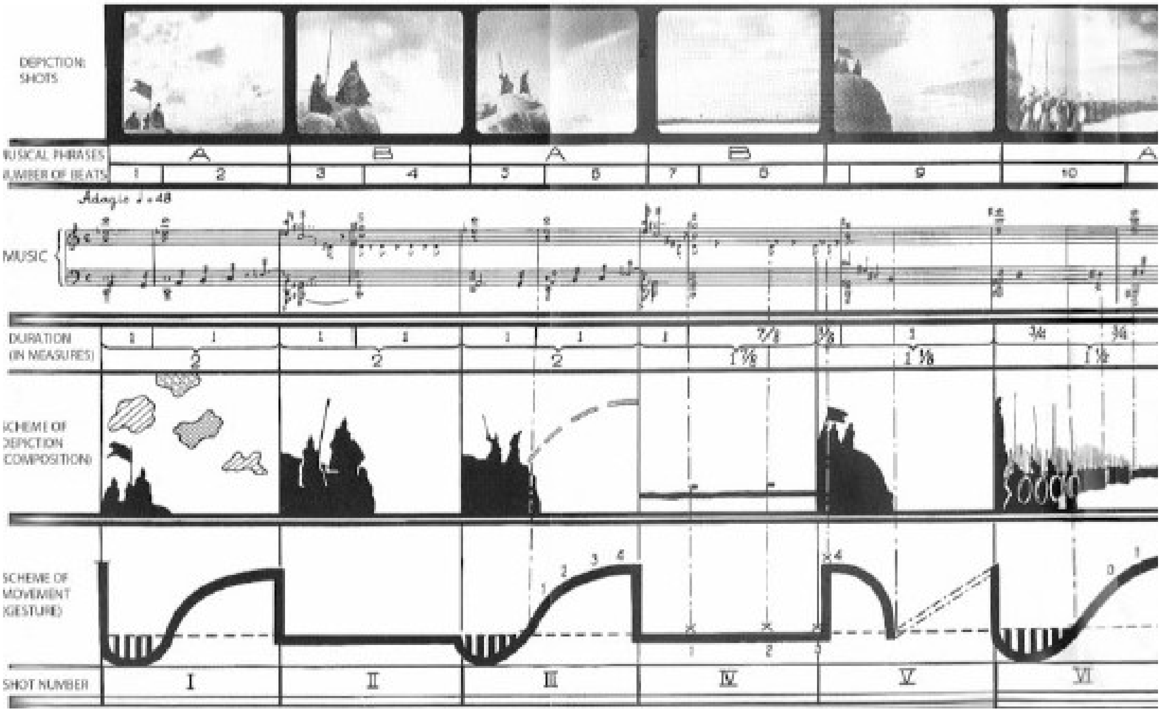

DEPICTION: SHOTS
USICAL PHRASES
UMBER OF BEATS
A
B
A
B
A
1
2
3
4
5
6
7
8
9
10
Adagio ♩= 48
MUSIC
DURATION (IN MEASURES)
CHEME OF EPICTION COMPOSITION)
SCHEME OF MOVEMENT (GESTURE)
SHOT NUMBER
I
II
III
IV
V
VI

Figure 7. Sergei Eisenstein's 'Vertical Montage'.

Adorno and Eisler criticised Eisenstein's Vertical Montage by stating that there is no guarantee that the spectator experiences the image in the way Eisenstein intends, from left to right like the musical score.[104] They also iterate that high tones, according to Eisenstein, do not necessarily have anything to do with the high registers of the image, like lower tones and lower registers.[105] For Eisenstein, Kia Afra argues, the eye can be trained to hear while the ear can be trained to see. In Eisenstein's earlier montage theory, based on conflict, and his ideas on 'synaesthesia' as seen in his vertical montage, the relationship between image and music, are theorised on a certain degree of *movement*: "...surely find those whose *movement harmonises* not only with the movement of the rhythmic pattern, but also with the movement of the melodic line."[106] Afra's main problem with vertical montage is that it does not consider many other factors in film, namely, movement *within* the image as well as between multiple images in montage. If the image physically moves, if there is movement within the frame, or if the camera is moving, how does the theory in the "melodic line and chord formations in the music, which find their respective correlations in the image", have universal application to multiple audio-visual stimuli?[107] Afra concludes that Eisenstein's model is merely an experiment, tested only on one sequence for which Eisenstein had his vertical montage in mind. He was attempting to find a language that analyses music and image simultaneously, as his model visualises. Eisenstein declares after heavy consideration:

---

[104] Ibid., 34.

[105] Ibid.

[106] Eisenstein, *The Film Sense*, 84.

> "To remove the barriers between sight and sound, between the seen world and the heard world…To bring about a unity and a harmonious relationship between these two opposite spheres…What an absorbing task! The Greeks and Diderot, Wagner and Scriabin - who has not dreamt of this ideal? Is there anyone who has made no attempt to realize this dream? Our survey must lead us to some method of fusing sound with sight, to some investigation of the preliminary indications leading towards this fusion."[108]

Although Vertical Montage is not entirely successful, it does at least attempt a method for bridging the methodologies of Musicology and Film Studies to analyse film music.

After surveying multiple approaches to the analysis of film music from Musicology and Film Studies, with some attempts at bridging those methodologies, the enterprise seems to be as divisive as the history of Film Music Studies itself. It has been difficult for scholars to avoid the separation of music and image when analysing the function of music in film, but Eisenstein's original vertical montage model is promising. It proves that not all attempts at analysing film music require the separation of music from image. To move beyond Eisenstein’s model, we require another definition of film music ontology.

---

[107] Afra, "'Vertical Montage' and Synaesthesia: Movement, Inner Synchronicity and Music-Image Correlation in Alexander Nevsky (1938)," 47-52.

[108] Eisenstein, *The Film Sense*, 87.

Chapter 2: What is the Music-Image?

In this chapter, I attempt to formulate a theory explaining the ontology and operation of what I call the 'music-image', which I hope will provide a model that will encourage analysing film music as a single, indivisible audio-visual image. This audio-visual image involves the cinematic spectator in a mode of film experience I adopt from Gilles Deleuze's concept of the 'spiritual automaton'.[109] My decision to integrate Deleuze's concept of the spiritual automaton is inspired by the work of Gregg Redner. Although in Chapter 1 I questioned Redner's broad-based application of Deleuzean film theory to music, in this chapter I utilise Deleuze's ideas of the automatic image. However, the music-image is not solely based on Deleuzean theory. The music-image is fundamentally a philosophical model that exists within the experience of the spectator and the film image, a new way of thinking to approach how film music functions. The music-image does not rely on certain Deleuzean theory to process analysis. In the following analysis, I attempt to expand the concept of the spiritual automaton to interlock the semiotic systems of music and visual image in a whole audio-visual system of meaning. The music-image coheres within a 'semiotic ambiguity' that interlocks music and image, and structures the meaning of film-music without fixed meaning. This philosophical and theoretical model of the music-image avoids top-down Grand Theorising, directly engaging spectator experience as intrinsic to the ontological form.[110] The model exists instead in the individual analysis of a music-image subject to its film. What this means is that every film bears its own music-images that are unique from any other film or music-image. Not every film image is a music-image, however, as music must be cued for the music-image to exist.[111] With the spectator as part of the music-image model, the same audio-visual cue in a film can result in infinite interpretations, meaning, there can be an infinite number of music-images processed from one audio-visual cue. This unique nature of a music-image governs the choice, or lack of choice, in films and music-images to analyse.

---

[109] Gilles Deleuze, *Cinema 2: The Time Image*, trans. Hugh Tomlinson and Robert Galeta. (Minneapolis: University of Minnesota Press, 1989), 156.

[110] See, Bordwell, "Contemporary Film Studies and the Vicissitudes of Grand Theory.", for a definition and study of Grand Theories in Film Studies.

[111] In Chapter 5, I discuss and expand on this notion of 'when' a music-image exists in a film.

The structure and function of the filmic ‘diegesis’ is integral to the semiotic ambiguity of the music-image. Diegesis is, according to Claudia Gorbman, the space-time universe and its inhabitants that are referred to by the principle of filmic narration; these ‘inhabitants’ – the content of the diegesis – will provide much of the structure of film music analysis for the remaining chapters.[112] The music-image’s ontology is grounded in the interpretation of the film being analysed, rather than a set of parameters that define what film music is. I also adopt Kristin Thompson’s Neoformalist methodology to analyse any film and its film-music images without recourse to a Grand Theory.[113] To bridge the two methodologies adopted by the separate disciplines of Musicology and Film Studies, I will seek to explain the importance of incorporating musical theory in music-image analysis. These musicological concepts affect the way spectators experience the music-image. From Film Music Studies’ convoluted history to the importance of bridging the methodologies of Musicology and Film Studies, analysing film music by employing the hermeneutics of semiotic ambiguity encourages a new way of thinking about film form with endless possibilities.

The difference and repetition of film music

In *Difference and Repetition*, Deleuze proposes that in a world where everything is indirectly connected, materialism and dialecticism are differentiated to create new forms, much like the ideas and thoughts that govern them. He states: “…thought thinks only by means of difference…The theory of thought is like painting: it needs that revolution which took art from representation to abstraction."[114] If applied to film music analysis, we can say that the analysis lies in the ideas that already exist to promote a way of thinking.

---

[112] Gorbman, "Narrative Film Music," 195.

[113] See Thompson, *Breaking the Glass Armor: Neoformalist Film Analysis*.

The choice to review *Difference and Repetition* is not an arbitrary one. Deleuze is a philosopher of *difference*, and if there were to be a new way of thinking about film music, something has to be *different*. What does Deleuze have to say about the 'new', but also, the origin from which the new derives? If a new way of thinking about film music is to be obtained, on the origin of the 'Idea', Deleuze insists that: "We may conclude that there is no true beginning in philosophy, or rather that the true philosophical beginning, difference, is in itself already repetition."[115] If there is no true beginning to an idea or way of thinking about the point of origin of film music, for that matter, is there repetition in a new way of thinking?

James Williams argues that for Deleuze, "things acquire an actual identity through repetition."[116] In searching for a new way of thinking about film music, as this thesis attempts, there is an ideal scenario, a light at the end of the tunnel, which is what Deleuze calls, the 'virtual'. A virtual is an aspect of reality that is ideal and exists as much as the actual reality. Deleuze outlines three different types of repetitions that situate the actual and the virtual; the third, which is most probable, "explains how things change in relation to virtual becomings, to difference in itself…repetition in habit and memory are only possible on a background of virtual differences."[117] Instead of an ontology of 'what there is', Deleuze wishes to account for 'how things change over time'; in so doing, he accounts for difference over time.[118] Knowledge is repeated, but there is a newness in each repetition. Difference in *every* repetition accounts for change. We can productively apply this model of difference (and change) in repletion to the semiotic changeability of the music-image. The first film score was composed based on musical codes which already exist; the composer repeated the act of composing music, then made a difference by interlocking the score to a film, in effect changing both film and score.[119] We can similarly say that any film is inherently different from any other, and yet many elements, or constituent parts, are *repeated*. If two film scores from separate films are interchanged, each film still has a score, and yet both films and both scores are now different and altered in a set of new relations.

---

[114] Gilles Deleuze, *Difference and Repetition*, trans. Paul Patton. (London: The Athlone Press, 1994), 276.
[115] Ibid., 129.
[116] James Williams, *Gilles Deleuze's Difference and Repetition: A Critical Introduction and Guide* (Edinburgh: Edinburgh University Press, 2003), 11.
[117] Ibid., 12.

When watching a film with film music, the spectator is listening at the same time. There may be many ways the mind translates this audio-visual phenomenon when it is experienced, and this may be different for each spectator. Sergei Eisenstein attempted an analysis of the visual and aural stimulus simultaneously to consider *one, whole stimulus*, but retired at the language barrier between the visual and the aural, realising that he could only describe them one at a time even if they *happen* at the same time. One method Eisenstein attempted was to combine words from a poem he wrote with music and colour. Here is an example:

> “WORDS: Sadly she wandered, loveliest of maidens.
>
> MUSIC: The notes of a flute, plaintive.

---

[118] A helpful thought experiment to break down this concept is to consider the construction of origami. Origami uses a piece of paper that can be folded and manipulated in an infinite number of ways. It can form and create different things and new meanings, like a bird or a butterfly, but it is still a piece of paper. Deleuze and many other philosophers, like Jacques Derrida through language and semantics, believed that the world and human nature are unstable, ontologically speaking. Every time the thought occurs of that bird crafted in paper, prior knowledge is repeated, but not all birds crafted in paper are exact copies of the previous ones.

[119] This thought experiment is not limited to film music, as, for example, considering the video-streaming platform Netflix, the inventor would have seen a DVD or a digital file, repeated the idea of a DVD or a digital file to view and store media, and then made a difference by inventing Netflix delivered over the internet. A DVD and the cinema both require a film, a production company, a distributor to function, and so does Netflix.

COLOR: Olive, mixed with pink and white."[120]

This method, however, is more comparative than analytical, relying on words to depict sound and vision. It would not work within cinema as it considers neither filmic form nor musical form. The closest Eisenstein got to a shared language to analyse film music was to analyse them visually through his 'Vertical Montage', previously discussed. Nonetheless, Eisenstein's model cannot deal with the inherent instability of film sound nor subjective interpretation of his own spectatorship.[121] Muting the film would disrupt Eisenstein's model because his model only accounts for one fully determined 'whole' image and experience under controlled conditions (the 'Battle on the Ice' sequence in *Alexander Nevsky* as it was conceptualised with his composer, Sergei Prokofiev).

The music-image automaton

I begin with the presumption that spectators experience films in different ways, and that each spectator's engagement is a unique experience.[122] Deleuze refers to the experience of film, or the automatic movement within the image, between the image and the spectator that compels the spectator to think, as the 'spiritual automaton'.[123] The term 'spiritual automaton' was first used by the Dutch philosopher Benedict Spinoza in his unfinished posthumous work *Tractatus de Intellectus Emendatione*, when referring to the mind:

---

[120] Eisenstein, *The Film Sense*, 89.

[121] For instance, what if a film were to be muted during a scene with film music, or what if another spectator did not experience a film the same way?

[122] A few notable works on cinema spectatorship and audience theory include: Judith Mayne, *Cinema and Spectatorship* (London; New York: Routedge, 1993)., Thomas Elsaesser, and Malte Hagener, ed. *Film Theory: An Introduction Through the Senses* (Florence: Taylor & Francis Group, 2009)., and, Carrielynn D. Reinhard, and Christopher J. Olson, ed. *Making Sense of Cinema: Empirical Studies into Film Spectators and Spectatorship* (New York; London; Oxford; New Delhi; Sydney: Bloomsbury, 2016).

[123] Deleuze, *Cinema 2: The Time Image*, 156.

> "This is the same as what the ancients said, i.e., that true knowledge proceeds from cause to effect—except that so far as I know they never conceived the soul (as we do here) as acting according to certain laws, like a *spiritual automaton*."[124]

The term may come across as robotic automation, but as Eugene Marshall suggests, this is not what Spinoza intended. Rather, Spinoza posits that the rational mind is a self-directed mechanism where the 'spiritual' is the mind and the 'automaton' is the self-directed, self-moving mechanism.[125] In Deleuze's adoption of the term, and if considering music as part of the image, the 'automaton' functions as part of this self-directed mechanism between film and spectator. Deleuze argues that unlike other artistic images (such as paintings or photographs), the cinematic image contains movement within itself as well as automatic movement between film image and spectator.[126] This automatic movement is not an automatous movement, as if the movement is a mind which thinks, but rather is a sensory "shock" that transports and "communicates vibrations to the cortex" so that the human mind, the spectator, engages with the image. The spectator is able to think of the image as a concept that generates action-thought.[127] The spiritual automaton creates a unity "between man and nature" through man's unity with film in automatic movement.[128]

The spiritual automaton relies on the concept of automatic movement, which is a movement that does not require a body or soul. It is a movement thought process within the image itself, a communication to the mind from the screen, a linkage between the thinker and the screen. Eisenstein, along with other Soviet Montage theorists such as Dziga Vertov, also promoted a phenomenological “whole” audio-visual image experience held together by movement.[129] The first movement, according to Eisenstein, goes from the image to thought, from percept to concept.[130] Recalling Gregg Redner’s discussion of Deleuze and ‘sensation’, here ‘percept’, one of the blocs of sensation, applies to the automaton.

---

[124] Benedict de Spinoza, *Tractatus de Intellectus Emendatione*, trans. W. Hale White (London: T. Fisher Unwin, 1895), 48-49.

[125] Eugene Marshall, *The Spiritual Automaton: Spinoza's Science of the Mind* (Oxford: Oxford University Press, 2014), 1.

[126] Deleuze, *Cinema 2: The Time Image*, 156.

[127] Ibid., 161.

[128] Ibid.. Deleuze uses the example of Sergei Eisenstein's montage cinema, in which the masses perceive, react, and act upon concepts that lie in their relation to the image, much like their relation to nature and the world. See, ibid., 156-63.

The ‘shock’ from the outside, as Deleuze describes it, influences the mind upon the percept’s entry. It shocks the mind like a nervous impulse and forces it to think, and to think of the ‘Whole’.[131] The whole can only be thought because it is an indirect representation of time which follows from movement; it relies on montage as a vehicle for this, as Deleuze argues. With the ‘whole’ as the concept derived from the percept, montage is then labelled as ‘thought-cinema’.[132] The elements of film form “enter in their own ways into suprasensory relations,” creating a shock wave of the nervous system.[133] Deleuze argues the spectator will no longer say “I see” or “I hear”, but “*I feel*” this totally physiological sensation.[134] These harmonies act on the cortex and give rise to thought, which move on to the second movement, which goes from the concept to the affect, a movement from thought back to image. The spectator understands the image based on the first movement from percept to concept by feeding the thought back into the image as affect. These movements are then in constant cycle, back and forth, animating the automaton in its movement regime. The music-image, like the Deleuzian movement-image, functions on this principle of the automaton. The spectator sees and hears the image, which moves to thought to be processed into one ‘whole’ image through image harmonies. The spectator then presents this thought back to the image in the second movement to understand the harmonic function music has in the image. To help understand this, Deleuze argues that metaphor functions on this principle. When shots collide in montage, the cortex reads this, as film scholar Christian Metz suggests, as a ‘language’ that feeds back to the image to actualise the harmonies that created it.[135] When the concept and image is identified, this is the third movement; concept is in the image, and the image is in the concept, an action-thought like “man’s relation to the world.”[136] To simplify these three movements: the first is perception (the spectator perceives the image), the second movement is affection (thought translating the image), and the third is the action-thought arising from the prior two movements. [see Figure 8]

---

[129] Ibid., 158.
[130] Ibid., 157.
[131] The idea of the ‘Whole’ comes from the branch of psychology known as the ‘Gestalt’, popularised by Max Wertheimer, Wolfgang Köhler, and Kurt Koffka. It is the gathering of external stimulus (that being the world external to the mind) through means of perception, behaviour, recollection, association, and thought, to create a ‘whole’ image. In film, if music and visual image were ‘perceived’ simultaneously, the mind would translate it as one ‘whole’ image, as Eisenstein theorised.
[132] Deleuze, *Cinema 2: The Time Image*, 158.
[133] Ibid.
[134] Ibid.

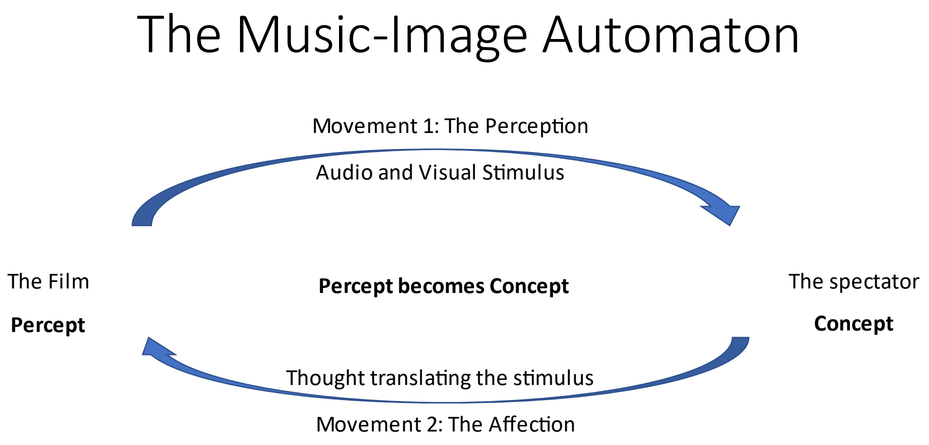

The Music-Image Automaton
Movement 1: The Perception
Audio and Visual Stimulus
The Film
Percept
Percept becomes Concept
The spectator
Concept
Thought translating the stimulus
Movement 2: The Affection

Figure 8. The spectator’s experience of film inside the model of an automaton.

---

[135] Christian Metz, "The Cinema: Language or Language System?," in *Film Language: A Semiotics of the Cinema* (Chicago: The Univeristy of Chicago Press, 1974), 46.
[136] Deleuze, *Cinema 2: The Time Image*, 161.

To tease out this complex model of the spiritual automaton, I will use the example of experiencing one of the time-travel sequences from *The Time Machine* (George Pal, 1960).[137] In this sequence, George, the time traveller, has pushed the lever forward on his time machine enabling him to travel into the fourth dimension, far into the future. After experiencing a nuclear explosion in the year 1966, while travelling through time (from 42:07 – 43:10), George becomes entrapped in molten rock until finally emerging from the rubble and stopping his machine at the year 802,701.[138] Despite an immobile camera and a lack of movement, there is a sense of movement in time created through the music that informs the visual image. George narrates throughout the sequence, worrying about his entrapment in the molten rock, but there is also a whirling, musical tune played by woodwind instruments lurking in the background that is not a sound effect within the diegesis.

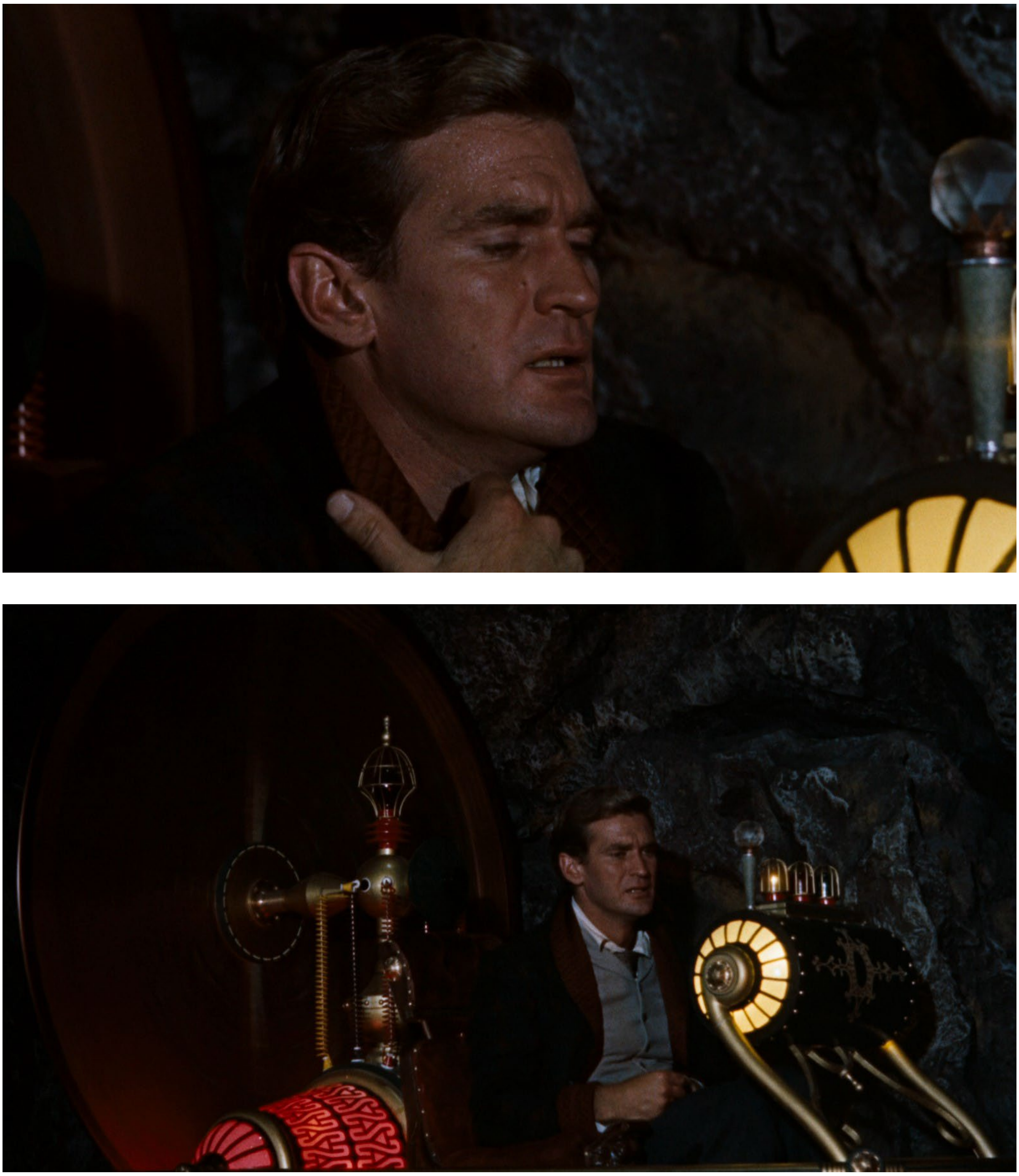
In what can now be called the music-image automaton model, derived from the Deleuze, I argue that the audio-visual stimulus is received as one channel. Visually, from 42:07 to 43:10 in *The Time Machine*, George is in a close-up and then a mid-shot, making just his upper body and part of the machine visible (See Figures 9-10). He shuffles from side to side, seated in his machine, sweating as he feels claustrophobic inside the cavern created by the molten rock, where the only light sources are the fitted bulbs surrounding him. In the close-up shot from 42:45 – 43:01, apart from George's narration, the only other audible source is the music track. A harp and woodwind instruments are heard ascending and descending the musical scale in a glissando. Glissando in music is when there are two notes at separate pitches

---

[137] Choosing this film as an example is not an arbitrary one. The choice comes from the fact that the film has no relation to any Film Music Studies literature, to prove that this model is built on a unique film experience rather than inspired from an existing theory. The choice is built off Kristin Thompson's choice in analysing the film *Terror By Night* Chapter 2 of her book, *Breaking the Glass Armor*, as it shows "Neoformalism can account for structures, materials, processes, and backgrounds in the average film…critical theory could apply just as easily to an average work." I agree with Thompson that choosing a film to analyse is difficult based on this notion, but there is certainly a goal here to demonstrate our shared idea. See, Thompson, *Breaking the Glass Armor: Neoformalist Film Analysis*, 49-50.

[138] Time code references are to *The Time Machine* (DVD), Warner Home Video, 2000.

(usually a significant interval between them) and the instrument played glides between the two notes creating a sliding, whirling sound. That whirling sound is distinctly present here. The choice of instruments is deliberate as well, as the harp and woodwinds when played together generates a tone that sounds mystical and fantastical, like the waving of a magic wand. When this audio-visual stimuli reaches the spectator, it is up to the spectator to piece together each fragment like pieces in a jigsaw puzzle to create a comprehensible image.



Figures 9-10. George in *The Time Machine* in a close-up then mid-shot.

There is no inherent sense of duration solely in the mid-shot of George sitting in his time machine. If this image were muted and there was no music, there is nothing to indicate that George is travelling through time. Without any narrative context, he could just be trapped in molten rock. With the mystical and fantastical layer of the whirling soundtrack, however, the lack of visual movement surrounding George suddenly appears to move in a way that is *out of this world*. The music gives a sense of this movement through the music-image automaton's automatic movement from percept to concept. The automaton explains that there is automatic movement from the film image to spectator in which the "thinker" (spectator) "can't escape the shock" the film image transmits.[139] Applying the automaton to this example from *The Time Machine*, both aural and visual stimuli are collected by the spectator in automatic movement between image and mind and are interlocked, translated into a conceptual thought or image. This new thought translates audio-visual stimulus into a mental image that the spectator comprehends, like "the relation between man and world."[140] The music stimulus, interlocked with the image of George positioned in close-up and a mid-shot trapped in molten rock, conceptualise into audio-visual movement, not just aural whirling movement. This concept, with the existing narrative knowledge of George travelling through time, generates a music-image of George *moving* through time.

In the second movement of the automaton, when the stimulus has been collected and returns to the image for the spectator to understand it, an *action* is created. If the visual and music stimuli are separated and played independent of each other, they could not be joined to create a new meaning through the process of automation. When George is immobilised in the molten rock, the camera is static, but the whirling tune in the music that gives a sense of duration, of time moving 'out of this world', *moves* the environment around him. The aural stimulus has affected the perception of the visual field, the physical space; the aural and visual stimuli have interacted, and fused, through the spectator's experience. As one image, with the two stimuli interlocked, it can be understood that George is travelling through time in the fourth dimension. The idea of the fourth dimension is fictional rather than ‘real’ science, and so the glissando performed on the harp and woodwinds create a fantastical sensory soundscape. Analysing the music that is intrinsic to the image creates a visual temporality, derived from the duration sensory prompts of the music. After the second movement of the automaton, when the action-thought of *believability* is created, only then can conclusions like ‘the music is autonomous’, ‘the music makes sense of the visual’, or, in the case of the example from *The Time Machine*, ‘the music makes George seem to travel in the fourth dimension’, be made.

---

[139] Deleuze, *Cinema 2: The Time Image*, 156.
[140] Ibid., 161.

Music and image as one image

Applying the framework of the spiritual automaton to film music is to situate music as part of the cinematic image and the analysis of film music as part of the music-image automaton. Claudia Gorbman alludes to this in her discussion of music’s relationship to film’s narrative. Gorbman argues that “[the] music in a film refers to the film…[and] bears [a] specific formal relationship,” which can seem separatist in nature, in that music and image are initially separate and then join.[141] While problematic for analysis, what she asserts about the ‘interlocking’ of music and film is quite profound. She argues that there is a limited amount of possible narrative-music relationships.[142] This means that music is specific to a narrative or image, and that they are interlocked; they have, according to Gorbman, “mutual implication.”[143] She asks whether in this case if any music is sufficient to ‘accompany’ a segment of film. Indeed, whatever is applied musically to an image will do *something*, it will have an effect, and Gorbman ostensibly agrees. To visualise this interlocking of music and image semantically in words, she suggests: "...Just as whatever two words one puts together will produce a meaning different from that of each word separately, because the reader/spectator automatically imposes meaning on such combinations."[144] This combinatory meaning functions similarly to what Deleuze argues in *Difference and Repetition*, that is, that words with prior meaning and context can be joined together to create a new one, but that the old, previously existing words still exist. Deleuze writes: "…difference not only appears in the contraction of the elements in general but also occurs…between two elements which are both determined and joined together by a relation of opposition."[145] The elements of 'music' and 'image' can oppose each other in relation but can be joined to create a difference that always already exists in nature.

---

[141] Gorbman, *Unheard Melodies: Narrative Film Music*, 13.
[142] Ibid., 14-15.
[143] Ibid., 15.

The attempt to identify the hermetic 'element' of music in a film enacts a *separation of that element from the image or narrative*. The interlocking of music and image – a music-image – is therefore not the suture of two elements previously prised apart but the music-image/music-narrative as a symbiosis that coheres as a combinatory form and meaning, but that nonetheless carries prior formal organisations and meanings ('music' and 'image') in the process. The same can be said about bridging the methodologies of Musicology and Film Studies when analysing film music: if only one methodology is applied to the process, music will inevitably be separated along with the separate methodological approach. This reconciles with Gorbman's argument that film music is not 'pure music' or a set of pre-existing musical codes and should not be analysed as such.

Arthur Custer, a film composer, agrees that film music should not be composed for meaning that coheres purely within the parameters of musical form. He writes:

---

[144] Ibid.

[145] Deleuze, *Difference and Repetition*, 77.

> “...there is the necessity to relate the music in a very specific way to visual and dramatic elements in the film. Others may see this as a loss of autonomy. I regard it as a particular benefit in the working out and shaping of music ideas. The composer of film music, by definition, is given the means to control his materials. He is presented with the incomparably precious gift of a framework for his musical ideas."[146]

Tom Armstrong writes similarly, suggesting that music for film is more grounded and interested in the musical language of emotion than any export of pure musical codes.[147] Custer argues further that “…the composer's creative sensibilities are meshed with those of the filmmaker and his associates, mutually enriching the final result. All are engaged in a process aimed at giving concrete meaning to highly subjective sonorous metaphors."[148] However, if further considering the spectator as part of this image process in the formation of a music-image automaton (which I argue is essential to its formulation as a theoretical object), the form or ‘language’ of the image is decanted through the cognitive and affective engagement of the spectator. And therefore, any notion of the music-image as ‘concrete’ or contained is misguided. And for this reason, also, the bridging of Musicology and Film Studies must be rooted in analysis occurring in constant relation to a set of formal elements, and cannot conform to a predetermined Grand Theory.

On the topic of mass art and its meanings, Noël Carroll agrees that audience activity is involved in the response to an image.[149] If we take this influential position as a starting point, we must acknowledge that the music-image is subject to change, fluctuation, and layers of semiotic ambiguity. Garry Schmidt uses the term ‘semiotic ambiguity’ when referencing scent, for being “invisible it appears to have components of non-materiality, suggesting perhaps even a substitution for the soul."[150] Adopting Schmidt's approach, I would suggest that the inherent ambiguity of the music-image begins with the semiotic relationship between music and image which is invisible in the sense that it carries 'components of non-materiality' that are formally determined by the cognition and affective life of the spectator. If music were to be played independent of any other activity, it would function in 'pure musical codes', and the resulting ambiguity (intrinsic to the abstract nature of musical codes) creates an open space for subjective intentionality. That is, ambiguity derives from the relationship between the signifier and the signified, and the interstitial encounter provided by the listener. Umberto Eco calls this semiotic relationship of signifier and signified a 'metaphor' that often elicits references to the visual from the aural signifier.[151] Since film music is not independent of other activity, as it operates within a visual landscape that has an independent semiology of its own, the signified 'metaphor' may not be what the spectator *feels* when experiencing a film. The two ambiguous 'codes' of independent aural and visual stimuli are interlocked to create a new audio-visual signifier in film music. However, the audio-visual interlocking requires the subjective thought of the spectator to operate the music-image automaton that generates analysis. It is up to the spectator to analyse the indexical interlocking of music and image, which makes the music-image semiotically ambiguous and radically open as a complexly signifying semiotic system.[152]

---

[146] Arthur Custer, "Comparing Concert with Film Music," *Back Stage (Archive: 1960-2000)* 20, 16, April 20, 1979, 34.
[147] Tom Armstrong, "Response: Music, image and the sublime," *Textual Practice* 22, no. 1 (2008): 71.
[148] Custer, "Comparing Concert with Film Music," 34.
[149] Noël Carroll, *Philosophy as Art: A Contemporary Introduction* (London; New York: Routledge, 1999), 33.

Film music, in its sonic output, is paradoxical in the Cavellian sense discussed earlier: that is, the music is only a 'likeness' to what is actually being performed. Film music is a recorded sound of a score in another space than that of the film, just as the screened world of the film is screened from the spectator.[153] Since film music shares that paradoxical 'likeness' from its aural dispersion with the projection of the visual image on a silver screen, there is a partnership, a shared transmission, an interlocking that moves to the cortex of thought. The visual and aural material constitutes one *shared* image experientially accessed as one image by the spectator. How does one interpret this shared image if language forbids its simultaneous translation?[154] In Deleuze's second movement of the spiritual automaton, the affection, the spectator reacts from the experience of feelings and thoughts brought from the first movement, that interlocked image stimulus. As seen in the example of *The Time Machine* earlier, the aural and visual stimuli collected by the spectator in the first movement of the automaton are collected as one stimulus, one image. This second movement, when the *comprehension* of the audio-visual stimulus has been fed back into the image as *experience* for the spectator, makes the spectator believe that Judy Garland is singing 'Somewhere over the Rainbow' in *The Wizard of Oz* (Fleming, 1939), Fred Astaire and Ginger Rogers are tap-dancing to the beat of the music in *Swing Time* (Stevens, 1936), and the camera is 'finding' the source of the music played at the start of *La La Land*'s (Chazelle, 2016) opening musical number; there is a semiotic relationship between aural and visual stimuli. In the example of Judy Garland singing 'Somewhere over the Rainbow' in *The Wizard of Oz*, for instance, if the spectator could not hear the song, or if the film were muted, there would be a different experience and comprehension of what occurred in the filmic space, just as there would if the spectator could hear the song but could not see Judy Garland. If the film were muted, Judy Garland would appear to be moving her mouth in a non-conversational way (see Figure 11). Judy Garland is not singing 'Somewhere over the Rainbow' while the film is muted because the spectator needs that aural stimulus that is interlocked with the visual material.

---

[150] Gary Schmidt, "Between Authors and Agents: Gender and Affirmative Culture in Das Leben der Anderen," *The German Quarterly* 82, no. 2 (2009): 235.

[151] Umberto Eco, *Semiotics and the Philosophy of Language* (Bloomington: Indiana University Press, 1984), 89.

[152] The 'indexical', a term coined by semiotician Charles Peirce, is an expression whose "reference on an occasion is dependent upon the context." So for music and image, they are indexes 'interlocked' based on needing the context of the other to exist in the music-image. See Simon Blackburn. *Oxford Dictionary of Philosophy.* Third ed. s.v. "Indexicality." Oxford, UK: Oxford Univeristy Press, 2016.

---

[153] Cavell, *The World Viewed: Reflections on the Ontology of Film*, 17-18.

[154] As seen in Eisenstein's assessment in Chapter 1. See, for reference, Eisenstein, *The Film Sense*, 87.

Conversely, if the screen was turned off but the speakers left on, the experience is solely based on pure musical codes, and thus contains a semiotic ambiguity subject to the spectator. The spectator perceives that Judy Garland is singing, processing the second movement of the music-image automaton in which the interlocked audio-visual stimulus in *The Wizard of Oz* was experienced. There needs to be the synchronisation of the song with the visual image of Judy Garland moving her mouth in that particular way for the spectator to experience "Judy Garland is singing 'Somewhere over the Rainbow'". The statement "Judy Garland is singing 'Somewhere over the Rainbow'" only exists in the experience of the spectator; it is something stated *after* the experience. The analysis, then, within the automaton, creates the subjective meaning of the image. Since the automaton exists within the autonomy of the 'context' (what is experienced), the third movement, the identity of the concept and image, is ambiguous, in that the semiotic interlocking is subject to the individual spectator's processing of the automaton. A *feeling*, or an *analysis*, of a given music-image, is subject to that individual automaton. It forms a basis for how one conducts an analysis of film music. And that interlocking whole that forms a semiotic ambiguity within the identity of concept and image *is the music-image itself.*

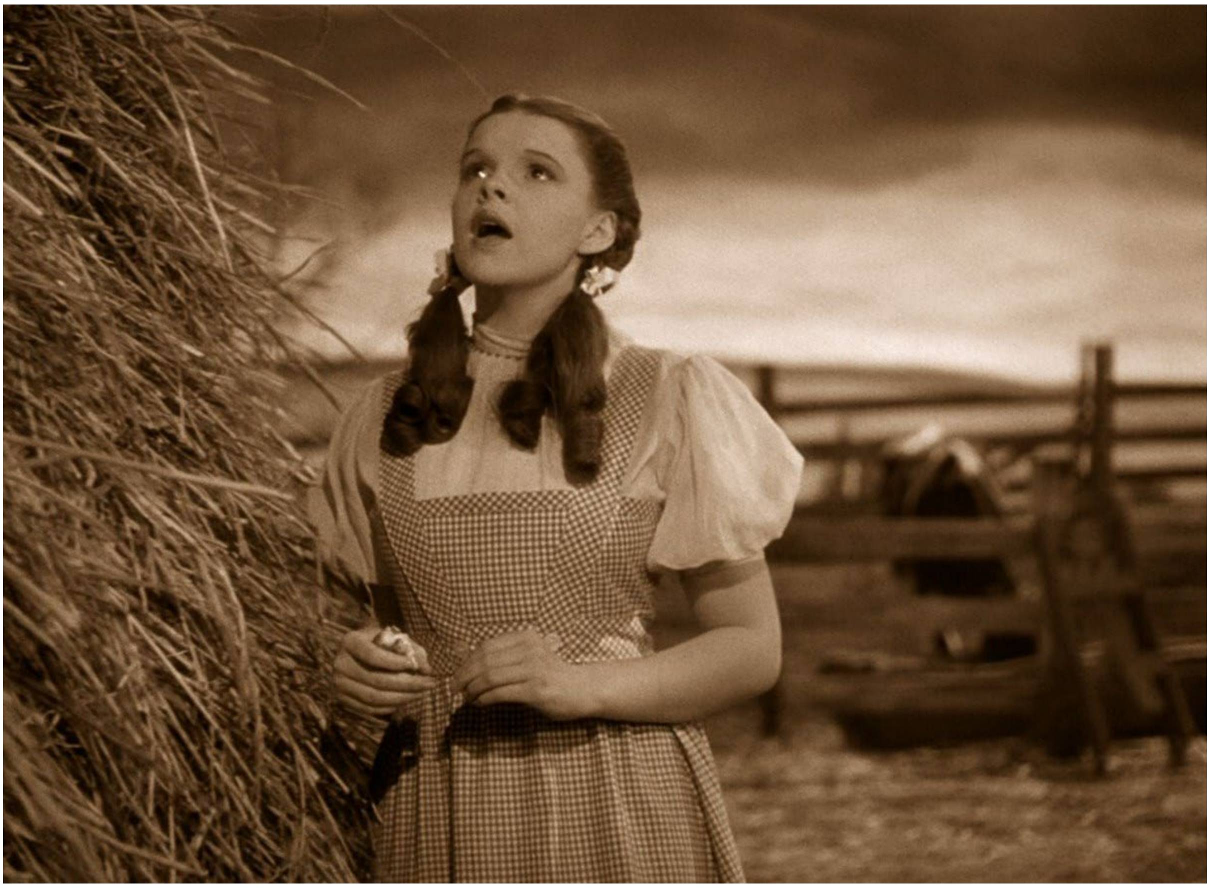


Figure 11. The moment Judy Garland begins singing ‘Somewhere over the Rainbow’, in *The Wizard of Oz*.

The music-image

I have argued that semiotic ‘ambiguity’ is intrinsic to the music-image, and transfers experiential and analytic agency to the spectator in the meaning-making process. But further, we can say that the music-image only exists within cinematic temporality, as both music and the cinema require duration to exist. If the film were to be paused, what is left is a 'visual-image'. If the film were to be resumed with the screen turned off, what is left is music operating in what Gorbman describes as 'pure musical codes'. Resuming the film, turning on the screen and letting the music synchronise to the visual image, achieves the gestalt of 'film music'.[155]

I employ a critical terminology – whole and part, the formation of the music-image as gestalt, semiotic ambiguity, and time (duration/temporality) – to underpin this ontology of the music-image. These terms also organise the analysis of the music-image into an experience.

Gestalt is a branch of psychology popularised by Max Wertheimer, Wolfgang Köhler, and Kurt Koffka based on form, shape, and configuration. It theorises how the human mind (perception, behaviour, recollection, association, and thought) organises external world stimuli into an experienced 'whole' in order to sustain a completeness, stability and balance.[156] Koffka famously used the example of a pianist performing a piece and even though the pianist may have played it to perfection, the audience and critics can still be divided, they can still be disappointed that "he played with bad taste".[157] Like looking at an abstract drawing or painting, the audience can be puzzled even though it is well drawn or painted. Even though the same audio is transmitted to the audience, the individual spectator is subject to the analysis, which creates an ambiguity in semiology from stimulus. Joanna Bailie proposes that gestalt in cinema is "a model of human perception where the different senses process their input in parallel, and perceived items are organised into coherent patterns to form whole percepts, which are somehow different to (or more than) the sum of their parts."[158] Gestalt can be attributed to most things in film, for example, narrative and character form and articulation. For Bailie, the mind "desires clarity over the complexity that constitutes the real world, and [the spectator] attempts to order the stimuli perceived in as simple and regular form as possible."[159] Audissino provides an example of narrative/character gestalt from *Gone with the Wind* (Fleming, 1939), with Tara's Theme being situated in the major key at the opening to establish the bustling O'Hara plantation, then later in the film the theme is reprised in the minor key when Scarlett returns home to a decrepit estate to communicate Scarlett's unhappiness; a narrative/character gestalt organised with major/minor musical modulation.[160] The mind has interlocked music and image to make sense of the narrative world. This way of thinking about film music as a 'whole' relates to the semiotic ambiguity created from the interlocking between music and image subject to the spectator. The 'whole' functions within a cinematic temporality. One thing that has constantly been mentioned in film music scholarship, ever since Prendergast and Gorbman, is the relation between time and movement in cinema. Both music and cinema only exist within a certain temporality, parallel to movement. This movement within the music-image also forms its basis, that is, the automaton that positions the spectator. The three movements that Deleuze described in his spiritual automaton only exist within time, like movement itself. For that transmission from image to spectator in the first movement of the automaton to exist requires a movement cast in duration.[161]

---

[155] The idea of the 'gestalt' of film music in this thesis, though not applied widely, comes from Emilio Audissino's gestalt/Neoformalism framework, and Joanna Bailie's theorisation of the "bad Gestalt', in Audissino, *Film/Music Analysis: A Film Studies Approach*., and Bailie, "Film Theory and Synchronization.", respectively.

[156] Audissino, *Film/Music Analysis: A Film Studies Approach*, 100.

[157] Kurt Koffka, *Principles of Gestalt Psychology* (Oxford: Routledge, 1999), 352.

[158] Bailie, "Film Theory and Synchronization," 70.

Joanna Bailie calls the "complexity that constitutes the real world," in this case the music-image stimuli that forms the basis of the first movement of perception, a 'good-gestalt' when it is clarified.[162] Bailie uses the screening of the 'Duelling Cavalier' from *Singin' in the Rain* (Kelly and Donen, 1952) as an example of bad synchronisation that results in the mind's inability to make 'good gestalt'.[163] The audience in the film find the desynchronization hilarious, and this allegorises that the spectator is only conscious of synchronisation when there is desynchronisation. The same can be said when music seems to mismatch the visual image.

---

[159] Ibid.

[160] Audissino, *Film/Music Analysis: A Film Studies Approach*, 109.

[161] I want to stress here that Gestalt Psychology is not a building block for music-image analysis. Gestalt, rather, is an overarching idea that contextualises the sense of a 'whole' image and will not be referenced frequently.

[162] Bailie, "Film Theory and Synchronization," 70.

[163] Ibid., 71.

If the music is sourced within the narrative space, it affects the inhabitants of the screened world as well as the experience of the spectator, and if the music is sourced outside the narrative space of the screened world, it effects only the spectator's experience. The diegetic music (music coming from the source of the narrative world), and underscore (music outside the narrative world), create completely different experiences for the spectator. This process requires separate analyses to locate their difference, and similarities, in definition and function. According to Claudia Gorbman, 'diegesis', relating to film music, constitutes the space-time universe and its inhabitants that are referred to by the principle of filmic narration.[164] Diegesis only exists in relation to screen and spectator, and in turn only exists within time. The fact that the spectator can identify the source of the music is related to the initial movement of perception in the automaton that is fed back into the image for music-image analysis through the second movement. Diegesis holds the music-image together and is thus integral to its semiotic ambiguity.

The concept of diegesis will structure the music-image analysis in this thesis, with separate chapters on the 'nondiegetic' (underscore) music-image, the 'diegetic', and their apparent difference (but potential similarity) in definition and function. I conclude this analysis of the music-image within the diegesis through a reading of what Deleuze identifies as the 'crisis' of the image, or what I refer to as the breakdown of the diegesis.

In the following set of analyses, I suggest throughout that the music-image is present in all interlockings of music and visual materials in cinema, and that such phenomena engage the spectator initially on the basis of intrigue.[165] In more general terms, a common ground between musicology and film theory is the study of movement and time. Their function within a music-image can be analysed to determine what film music is doing within a certain film. This analysis is based on the intrigue of an individual automaton to exemplify semiotic ambiguity and to highlight the music-image as a way of thinking rather than a theory that requires certain films or musical theory to affirm it.

---

[164] Gorbman, "Narrative Film Music," 195.

One final point to emphasise is the radically subjective results of the music-image experience within each spectatorial encounter. As I have suggested up to this point, and will develop further in the following chapters, I argue that there is a semiotic ambiguity holding music and image together because there is no fixed meaning, theory or communication attached to a discrete analysis of a music-image. Bridging the separate and often conflicted film and musicological methodologies through the framework of the music-image encourages the analysis of a single unified *audio-visual image*.

---

[165] The analysis of film music through the music-image model is similar to Kristin Thompson's Neoformalism approach to film analysis, as referenced by Emilio Audissino in, Audissino, *Film/Music Analysis: A Film Studies Approach*, 67-94. See, Thompson, *Breaking the Glass Armor: Neoformalist Film Analysis*. Thompson declares that the spectator analyses based on assuming what films consist of, how people watch film, and how the film relates to the world and society. When the spectator examines their assumptions, there is a systematic approach to analysis. As an aesthetic approach then, as Thompson argues, the spectator refers to a set of assumptions about "traits shared by different artworks [films]" and employs a set of procedures to analyse. See, ibid., 3. Grand Theories have emphasised that films are in part intriguing, but Neoformalism assumes that the analysis of film is based on the notion that films are generally intriguing. Thompson argues that one approach to film analysis cannot anticipate every possibility, and that analysis at face-value can have infinite possible readings. The analysis involves the initial intrigue, a careful (audio-visual) viewing of a film while thinking of film through the automaton framework, and in turn make sense of the film's own music-image. The fact that Thompson explains that part of the analysis process is to pass on results of this close viewing to others only strengthens the position of semiotic ambiguity due to the individual spectator's automaton.

Chapter 3: The Nondiegetic Music-Image

> "Image, sound effects, dialogue, and music-track are virtually inseparable during the viewing experience; they form a *combinatoire* of expression."[166]

Although the idea of 'combination' considers the separate nature of film entities, aural and visual components of the image are interlocked to generate one image of experience (for the audience) as well as expression (for the filmmaker) through the music-image. In this and subsequent chapters, I examine cinematic diegesis through the lens of the music-image, and I begin in this chapter with the nondiegetic music-image.[167] The nondiegetic music-image concerns the construction of an image combining music experienced outside the narrative space-time of the film. In Film Music Studies, the scholarly field that researches film music, nondiegetic music is typically called 'underscore', but this terminology is problematic when considering diegesis. Nondiegetic film music is not separate from the visual image (implied in the term 'underscore') but is rather intrinsic to audio-visual narrative form. In the first part of this chapter, I will attempt to redefine the concept of nondiegetic film music through the lens of the music-image; in the second part of the chapter, I examine the audio-visual components of the nondiegetic music-image, notably audio-visual harmony, time, tone, and space in several film analyses.

---

[166] Gorbman, *Unheard Melodies: Narrative Film Music*, 16.

[167] I use the term 'nondiegetic' in this chapter for consistency, as scholars have either used 'nondiegetic' or 'non-diegetic' for referencing the same term. For instance, Claudia Gorbman uses 'nondiegetic' in "Narrative Film Music.", and Ben Winters uses 'non-diegetic' in Ben Winters, "The Non-Diegetic Fallacy: Film, Music, and Narrative Space," *Music & Letters* 91, no. 2 (2010).

What is underscore and how is it sourced?

The term ‘underscore’ is hard to define in Film Music Studies as scholars in the shared fields of Musicology and Film Studies variously agree and disagree on its usage in film music analysis. For example, in their book *On the Track: A Guide to Contemporary Film Scoring*, Fred Karlin and Rayburn Wright define underscore as “supporting music, which the audience hears but the actors do not.”[168] This definition, which is widely accepted by scholars who use the term in their own work, conveys a spatiotemporal quality relating to film ‘diegesis’. As explained in Chapter 2, ‘diegesis’, according to Claudia Gorbman, is the space-time universe and its inhabitants that are referred to by the principle of filmic narration.[169] Applying this notion to film music, there are two basic principles we can employ: diegetic space-time and its externality, the nondiegesis, with a further metaphoric relation, metadiegesis.[170] Michel Chion bridges these two terms ‘underscore’ and ‘nondiegetic’ in his book *Audio-Vision*: "...to designate sound whose supposed source is not only absent from the image but is also external to the story world.” He concludes: “I shall use the term *nondiegetic*. This is the widespread case of voiceover commentary and narration and, of course, musical underscoring."[171]

The use of ‘underscore’ is specific to film music while the use of nondiegetic music can relate to the entire spatiotemporal universe, as Gorbman suggests. However, the congruency between nondiegetic and underscore as theoretical terms unravels in analysis. In *The Oxford Handbook of Film Music Studies*, James Buhler, and David Neumeyer state that:

---

[168] Fred Karlin, and Rayburn Wright, *On the Track: A Guide to Contemporary Film Scoring* (New York; London: Routledge, 2004), 39.

[169] Gorbman, "Narrative Film Music," 195.

[170] ‘Diegetic’ music occurs within the world of the story and can be heard by the characters within the space. See, OED Online, "diegesis, n.," Oxford University Press, https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/52402?redirectedFrom=diegetic.. ‘Nondiegetic’, or as some scholars note, ‘Non-diegetic’ music is not diegetic in the sense that its source is external to the world of the story, and not heard by the characters. See, "non-diegetic, adj.," Oxford University Press, https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/69156318.. ‘Metadiegetic’ film music is an extension of these two distinctions, as it is heard by the narrator, creating a psychological source within the diegetic world.

[171] Chion, *Audio-Vision: Sound on Screen*, 73.

> "the fundamental function of music in the later silent era was to *underscore* the underlying narrative structure of the film by establishing a musical unit of structure, the musical cue, that extended across individual shots, binding them together into a larger unit of narration, the sequence….[172]"

Film music would continue to 'underscore' in this way when the diegetic space-time of sound/music in the late 1920s is established. However, the nondiegetic realm of film music is scrutinised in Film Music Studies by giving it the sole purpose of underscoring, as seen in the example, "underscoring is in effect the place where the concept of nondiegetic music *was forged....*"[173] The problem in this explanation is that the nondiegetic space-time of film music has been positioned as a separate entity which 'underscores' a universe *already* existing rather than existing as a single universe. This distinction between music and image in Film Music Studies has led to scholars such as Gregg Redner, whom I reference in Chapter 1, attempting to simultaneously analyse aural and visual material as one image, and thereby conceptualising film music underscoring as a function of story rather than the entirety of visual form and narration. Favouring internal story analysis over audio-visual analysis can unintentionally turn the analysis in on the music itself instead of combining it with the spatiotemporal universe of the film. When mentioning the score to *Hamlet* (Kozintsev, 1964), Redner notes that Dmitri Shostakovich's score concerns itself with the internal narrative rather than the visual aspects of the mise-en-scene. [174] Redner also notes that the score links, because of the introspection, to particular characters, creating sub-plots of family groups. That is, Shostakovich allocates various aspects of the orchestra to principal characters. What is peculiar about this case is that Redner gives no reason as to why he has dismissed analysing visual style over story. The only argument Redner puts forward, which is quite vague, is that the score is not “visually-external.”[175] The music still, however, plays with the visual form of the image, and thus ‘underscore’ as a term can become convoluted in analysis. A case in point is when Johnny Wingstedt, Sture Brandstrom, and Jan Berg declare, “what (we think) we see is to a large degree determined by what we hear [and vice versa]", and yet with, “for the sake of analysis however, emphasis will in the following primarily be put on music’s contribution to how meaning is established in the multimodal interplay of the filmic narrative."[176] Ben Winters also disagrees with the dismissal of audio-visual analysis for the sake of referencing film music as underscore. Winters uses Chion's definition of nondiegetic sound, referenced earlier, to highlight the problem with the term ‘nondiegetic’. Winters uses the example of the theme for Indiana Jones in the *Indiana Jones* series, quickly noting that the music and the character Indiana Jones only exist together, not in separation.[177] The music may not be visually present, but the image of the character Indiana Jones would not exist without his theme music, *Raiders March*, which is intrinsic to the diegesis of the film.[178] This chapter argues that the term underscore, which has derived *from* the 'nondiegetic' (rather than leading to it as James Buhler and David Neumeyer state in *The Oxford Handbook of Film Music Studies*), exists within the film image just as film music sourced within the space exists within the film image. Nondiegetic music may not be produced physically within the space, but it functions and coheres within the *image of the space*. Ben Winters argues that the function of music extends beyond narration, a point with which I concur.[179] For the *Raiders March* to be associated with Indiana Jones, it needs to be, and is, part of the image. As soon as the first note of the first bar is played, the music and image cannot exist without each other in the film's space-time. The music-image is thus always a synchrony, and its elements are equal.

---

[172] Buhler, "Music and the Ontology of the Sound Film: The Classical Hollywood System," 20.
[173] Ibid., 32.
[174] Redner, *Deleuze and Film Music: Building a Methodological Bridge Between Film Theory and Music*, 81.

---

[175] Ibid.

[176] Johnny Wingstedt, Sture Brändström, and Jan Berg, "Narrative Music, Visuals and Meaning in Film," *Visual Communication* 9, no. 2 (2010): 194.

[177] Winters, "The Non-Diegetic Fallacy: Film, Music, and Narrative Space," 224.

[178] Ibid.

[179] Ibid., 225.

Audio-visual harmony and dissonance

Up until this point in my thesis, I have addressed the music-image as a polyphonic super structure. Music cannot be a visual image, and the visual image cannot be music. If music and visual image are different ontological forms, can they contain the same effect? Music and visual stimuli create an image, indeed, but its structure as a single entity is based on opposing forces that emit semiotic cues separately. When these opposing forces interlock, played on one channel, they interlock in harmony or disharmony. Claudia Gorbman argues that music either 'resembles' or 'contradicts' the action or mood of what happens on screen, but this assumes that music is separate from and subordinate to the visual image.[180] Although Gorbman uses the word 'accompany', it is agreed that any music is sufficient to *interlock* with the visual image in film.[181]

Harmony (as well as disharmony) exists when there are two channels coexisting, like a piano fugue consisting of two melodies in the left and right hands played together, but their coexistence creates a third layer of meaning, as these melodies are not played separately. The same can be said for the music and visual within the music-image. It is acknowledged that these two entities can mean different things and serve different semiotic functions when separated, but together on one channel, like the piano fugue, they create a harmonious meaning. When the separate semiotic outputs of audio and visual are in close harmony with each other, the interlocked ambiguity is less obvious. For example, when Luther's henchman,

---

[180] Gorbman, *Unheard Melodies: Narrative Film Music*, 15.
[181] Ibid.

Otis, in the film *Superman* (Donner, 1978), is running downhill in a distant wide shot about to show what an incompetent henchman he is, a single French Horn expresses Otis' brainless, mischievous nature as he waddles clumsily. This interlocking of semiotically similar outputs results in audio-visual consonance generating a comedic undertone. They are so similar that there is almost no doubt in the interlocked meaning generated.

Is there such a thing as disharmony in film music? To recall from Chapter 2, Joanna Bailie called the bad synchronisation in *Singin' in the Rain* for 'The Dancing Cavalier' a "bad Gestalt" intended for comedic effect, as seen through the narrative audience. However, does the music always have to harmonise to interlock? What does music-image 'disharmony' look like? A dissonant music-image exists as does a consonant music-image. However, in a dissonant music-image, the aural and visual stimulus signify opposite semiotic meaning, generating a more ambiguous interlocked semiotic system in the music-image. An example of this would be an image of a boy joyfully reconciling with his parents with crying strings in a minor key, aurally suggesting sadness on its own. This interlocking of music and image results in audio-visual dissonance. Audio-visual dissonance operates similarly to musical dissonance, but instead of occurring expressly through the musical form (dissonance between musical notes, for example), it is created between the music and visual forms. This audio-visual dissonance is not necessarily a diminution of a music-image that could have been, but rather the character of the semiotic output the visual emits. The reason the spectator can deduce a dissonant music-image shows a unity in semiotic disunity to generate new meaning.

A deep dive into close textual analysis of film music may not always match the perception of film. David Ireland writes, "incongruities that are identified by close textual analysis, for example, may not always be consciously recognised by the perceiver in the moment of engaging with a film text."[182] This is not to say that incongruencies do not exist, they clearly do, but only in the separation of aural and visual forms. This incongruence is erased by the yielding of one image. In the bombardment of music and visual image, separate pure music and visual meanings, whether incongruent or not, refer to each other to create a new meaning.

In the film, *The Neon Demon* (Nicolas Winding Refn, 2016), for instance, the soundtrack is entirely 'vaporwave' music. Vaporwave is a music genre built through computer technology, consisting of piracy and recirculation of audio-visual fragments, and indicates a film with an obsession with a particular production technology from roughly 1980 to the mid-1990s.[183] Already, this seems like an unconventional choice due to the lack of musical instrumentation. However, the image that is created is representative of both the visual narrative and the sounds of vaporwave music. Vapor cannot be harnessed, it cannot be pinned down, it lingers then eventually fades away, like waves, old and new, as they pass by, transferring energy from one place to another, from one time to another. Vaporwave is built on this principle of loose waves of vapor that consume time and space. The spaces that vaporwave generates are non-places in the imagination.[184] The genre therefore cannot exist without modern technology and the computer technology that produces it. Grafton Tanner suggests that vaporwave is the "musical product of a culture plagued by trauma and regression in late capitalism…It awakens the experiencer to the cultural maladies that stunt political discourse and shun human empathy."[185] Ross Cole, who cites Grafton Tanner, argues that vaporwave is basically the articulation of "the alienating nowhere of capitalist technocracy" while staging the "critique of somebody else's utopia."[186] The sounds of vaporwave are distortions of a once ubiquitous soundscape – technologies that were once new reappear as outmoded objects stained with time.[187] If *The Neon Demon* were to be analysed in terms of the music-image, it would begin with its one channel output. The film revolves around a young Jesse, who aspires to be in the Los Angeles fashion industry and is consumed by the jealousy and narcissism of the other fashion models. If you were to look at a poster or a promotional still from the film, the images are filled with neon lighting, complementing the film's title.

---

[182] David Ireland, *Identifying and Interpreting Incongruent Film Music* (London: Palgrave Macmillan, 2018), 11.

[183] Ross Cole, "Vaporwave Aesthetics: Internet Nostalgia and the Utopian Impulse," *ASAP/Journal* 5, no. 2 (2020): 300.

[184] Ibid., 306.

[185] Ibid., 303-4.

[186] Ibid., 302.

In the 'runway' scene, when Jesse 'becomes' the 'neon demon', she is consumed by a narcissistic regard for her natural beauty brought on by the industry. This consumption echoes that 'critique of somebody else's utopia' (Jesse's) and articulation of the 'alienating nowhere of capitalist technocracy' (the runway and fashion industry). Jesse's innocence is captured by the move from a medium profile shot to a series of close-ups of her face in harsh blue lighting as she carousels the automatic runway. She is not actively participating in showcasing her beauty; instead, the non-space reels her in (see Figures 12-13). The distortions in the soundscape are heard, unharnessed, moving in and out of audible reach in a strobic pulse. These sounds crudely loop in echoes and shifts in speed. There is a fascination with the hollowness of this empty space that Jesse moves through; it is a simulation of ubiquity. For these once new technologies to reappear or stand out, everything else must be paused or muted. In the scene, the diegetic sounds of camera flashes are drowned out by the nondiegetic vaporwave soundtrack littered with hi-fi nostalgia of neon lights and objects. The nostalgic image floods Jesse's innocent face with that innocent, nostalgic shade of blue. The pulsating light show on the runway surrounded by a sea of darkness spotlights her innocence corrupted by narcissism as the blues become reds and one close-up becomes multiple. She kisses her image that surrounds her in a prism of mirrors while the echoes of digital distortion envelope her image (see Figure 14-15).

---

[187] Ibid., 316.

Figures 12-13: Jesse in blue light on the runway in *The Neon Demon*.

Figures 14-15: Jesse in red light, becoming the ‘neon demon’ in *The Neon Demon*.

Injected with the formal colours and patterning of narcissism, Jesse has become the neon demon. She is consumed by the decayed capitalist utopia, and she loves it as much as herself. The wave of the vapor has washed over Jesse and controls her. The distorted energy of narcissism has transferred from the other models to Jesse, dissolving her innocent utopia. The image of vaporwave with Jesse on the runway is constructed through this interlocking; the separate incongruencies have become congruent, a harmony in presupposed disharmony through one image. The oddly specific choice to use vaporwave music in *The Neon Demon*, and especially in the runway scene, harmonises within the incongruencies that these two aural and visual images present separately. Instead of two images of music and visual source exhibited side by side, these sources collide to generate a new audio-visual language and its potential signification.

In the opening of *The Social Network* (Fincher, 2010), two inharmonious audio and visual sources create audio-visual harmony when interlocked. After the simple yet pivotal opening sequence in which Erica breaks up with Mark, Mark runs back to his dormitory room at Harvard University to create Facebook, prompting the first audio-visual cue in the film. If the proceeding analysis were to separate the visual language of Mark running back to his dormitory room from the musical codes of the polyrhythmic music and vice versa, there would be two different narratives trajectories for the film. Congruent meaning is born from interlocking the incongruent audio and visual language. A piano and synthesiser sounding like a tremolo on guitar hesitate on three G major chords, creating a pedal point in the depths of the image. The music begins as soon as Erica calls Mark an 'asshole' and storms out of the pub and Mark starts out for college. The fact that the pedal point cues in as the conversation ends is evidence that the music relates to the aftermath of that conversation - Mark being rejected by Erica. However, this pedal point continues for the entire cue. There is an underlying rage simmering beneath the surface of the wide shot covering Mark’s journey. The pedal point harnesses the visual language in a state of ambiguous meaning for Mark. Erica’s rejection of Mark is for Mark’s good, but also for the bad. Mark moves within the towering shots of the university campus, a figure in a sparse landscape, piercing through the space in this unresolved tension (see Figures 16-17). The pedal point in the synth hangs on the constant strumming of a low D, while the piano plays G major chords that do not resolve, and the music never returns to G tonic, remaining on the 5th, D major, holding tension. Something will happen, there will be resolution between Erica’s description of him as both a probable 'successful computer person' and an 'asshole', and the wide shots emphasise that the result might be significant. The minimal number of musical notes used also suggest clear focus on this tension and the clear, minimalist sound of code. This rejection by Erica has spawned the founding of Facebook through hope and inner rage on a large but intimate scale. If these camera shots were analysed solely, the lack of close-up would perhaps suggest that Mark is lonely and emotionally confused after the breakup, but this audio-visual harmony has synthesised a new image, a new meaning out of incongruence. Critically, the sound composition synthesises with a visual and narrative formalism to create an integrated semiotic field.

Figures 16-17. Wide shots of the university campus as Mark runs back to his dormitory in *The Social Network*.

There are many examples of films that create audio-visual harmony from a recipe of incongruent ingredients. In *Picnic at Hanging Rock* (Weir, 1975), the incorporation of eerie pan-flute and organ, which sound mystical and ethereal, coupled with the vibrant but strange and threatening Australian landscape forbodes a sense of the supernatural.[188] The opening of *Persona* (Bergman, 1966) blends string atonality with the clumsy construction of celluloid imagery, creating a meditation on the nature of film itself as an object of experiment. More famously, Erich Korngold's sweeping symphonic opera with full orchestra for an intimate tale of people growing up in a mid-west American town in *King's Row* (Sam Wood, 1942) seems overblown. However, arguably the most famous and noticeable examples of incongruence come from period films with contemporary scores, whether original modern orchestration or pre-existing popular songs.[189] Whatever the 'incongruencies' or disharmonies that come from the textual analysis, music and visual images harmonise upon interlocking. To echo Stanley Cavell, who writes "the particular hurt or crudity or selfishness or needfulness or hatred or longing that separates us must be given leave to declare our separateness, hence the possibility of our connection," film music seems incongruent when separated from the visual, but that very incongruence shows the pivot of a bond that creates new possibility of meaning.[190]

---

[188] D. Bruno Starrs, "Film as Text [1931-2007]: "Picnic at Hanging Rock" and the Puzzle of the Art Film," *Screen Education* 48 (2008): 141.
[189] Both *Marie Antoinette* (Coppola, 2006) and *The Great Gatsby* (Luhrmann, 2013) use contemporary pop soundtracks for period setting and narrative, for example.
[190] Cavell, *The World Viewed: Reflections on the Ontology of Film*, 128.

Audio-visual time: duration, rhythm, and tempo

Film music does not exist without time, as cinema does not exist without time.[191] The music-image follows the same principle, that is, that the movement of percept and concept between image (from which the percept is received) and spectator (from whom the concept is created) operates through time. In referring to film as an 'object', this coexistence of film music and cinema within time plays on the one phenomenological level. Within this construction of the film, the shot is subject to time. As Gilles Deleuze writes, the shot is a potential montage and time depends on the movement within montage and belongs to this movement.[192] Since the shot is not a free-flowing body, like that of nature, as movement and duration exist in the construction of images, typically 24 per second, movement is subordinate to time. However, the cinematic spectator does not experience film as a physical object even though time orchestrates the fabrics of the image. The spectator does not watch a display referencing time, the spectator is involved in an experience of orchestrated time. Thus, for example, the ontological distinction of film diegesis becomes more subjective based on experience – a form that does not exist without the construction of image.[193] Experiencing film music as part of the image, especially in narrative cinema, there is a multi-faceted reference of time. The existence of a film's duration, that is, the construction of images within montage, is distinct from the experience of narrative time. Narrative time can jump outside the 3-dimensional space-time of the film form. In this sense, narrative movement and time constitute an autonomous world from the principle that the spectator cannot control the perception of movement within the image. Audio-visual time is a different experience to a film's runtime.

---

[191] Much has been already said about cinematic time and the philosophy of time, in general. So much so that this section will not tackle the philosophical tentacles of this enquiry, but will acknowledge that the music-image, and cinema in general, does not exist without, and exists within, time.

[192] Deleuze, *Cinema 2: The Time Image*, 35.

[193] I go into more detail on the ontology of film diegesis based on the location of film music and experience in Chapters 6 and 7.

As Andre Bazin argues, cinema is of its essence "a dramaturgy of Nature."[194] Even though film music is bound by the duration of the film, it cannot exist without this fundamental parameter, its function lies in the diegetic time construct of narrative. In this space-time universe that only exists in the experience of the spectator, as stated by Claudia Gorbman, the music-image assists in the construction of narrative time; it would not be sourced within the narrative space. If music were sourced within the narrative space-time, the temporality of the music-image is bound by the experience of narrative constructs, as music is part of that experienced space-time. If the music were not sourced within the narrative space-time, the music-image is not bound by these narrative constructs, and instead, the narrative constructs are bound by the music-image, as music is part of the film image, and subjectively experienced by the spectator. This demarcation marks the distinction between diegetic and nondiegetic music-images. The nondiegetic music-image exists from within the construction of montage for the purposes of narration. The spectator experiences the interlocking of aural and visual image in time, allowing for the experience of audio-visual time, which can consider time-based parameters such as duration, rhythm, and tempo in analysis.

The first audio-visual time constraint is that of the duration of a shot or a sequence of shots. I want to suggest here that the narrative duration of shots and sequences is independent of the construction of shot and sequence duration, which David Bordwell distinguishes as ‘story time’ and ‘viewing time’.[195] David Butler argues that “time and music are intimately related and the consequences of that relationship can have a significant impact on our understanding and perception of the passage of film time: both within the world of the film and our own experience of the film unfolding before us.”[196] The music-image can manipulate the sense of narrative duration, regardless of 'viewing time', making the spectator process a dramaturgy of nature, rather than nature itself.

---

[194] André Bazin, *What is Cinema? Volume 1*, trans. Hugh Gray (Berkeley; Los Angeles: University of California Press, 1967), 110.

[195] David Butler, "The Days Do Not End: Film Music, Time and Bernard Herrmann," *Film Studies*, no. 9 (2006): 51.

[196] Ibid.

For a film seemingly shot in one take like *1917* (Mendes, 2019) to hold and release tension and create a sense of urgency to its movement of imagery, film music fills its visual void. By comparison, a film like *The Turin Horse* (Tarr, 2011), with its many long sequence shots without music, feels the weight of duration, demanding more concentration from the spectator and expenditure of their time. Little wonder then that Bela Tarr's films are regarded as 'slow cinema'. When music is part of a notably 'long take', the sense of duration is altered. As Bordwell states, extended duration of shots does not necessarily imply the extended duration of musical reference.[197]

For the scene of Enrique's funeral in *I am Cuba* (Kalatozov, 1964), it is famously achieved in a single take. The camera remains within the funeral procession without cuts, beginning in a mid-shot of Enrique's close but brief friend, then departs and remains distant from human faces. However, the movement of the camera keeps the camera and the spectator with the crowd. The marching 4/4-time signature rhythm organises the harmonies of low register strings and timpani in a repeated loop, presenting the image of the crowd moving in an orderly procession. Diegetic sound is mute in the image, letting the music and camera movement complement each other. When the camera begins its ascent up the side of the building as Enrique's coffin enters stage left and moves through the space, trumpets blast a counterpoint melody in a higher register to comment on this resistance to cutting the image. The spectator can see people watching the procession from their apartments as the music pushes the camera up the wall. There is a slight unharnessing nature to the camera, freeing the camera from any artificiality that could detract from a sense of spectatorship of the camera itself. The spectator feels the weight of the apparatus being pulled and dragged through the open space of the narrow street. The camera seems to hover mid-air above the crowd at multiple points in the sequence, but never remains stationary. The music situates the camera in the narrative dramaturgy of space. The wonderment of the camera freely moving through the space is sustained through the musical urgency of the march-like tempo. Everything and every space in the area is connected narratively, including the people on the street, the residents watching from their apartments, and even the workers packaging cigars. As Lida Oukaderova argues, it is a subjective experience that began from a point of view to many different viewpoints, indivisible from an objectively surveying gaze.[198] The funeral procession viewed as a continuum between specific viewpoints is harnessed by the differential duration and continuum of marching counter melodies of strings and brass. Enrique is mourned and grieved as a Cuban revolutionary, and the long duration of this collage of viewpoints is manipulated, and the stretched temporal void filled, by the marching music-image, a march that requires movement in order to alter the pace of existing movement. This effect can be said about other films that champion longer than usual shot durations, that film music exercises a certain manipulation of mathematical duration for narrative duration.

---

[197] Here, Bordwell is referring to Classical Hollywood cinema, but this is true to film outside this period. See David Bordwell, "Time in the Classical Film," in *The Classical Hollywood Cinema: Film Style & Mode of Production to 1960*, ed. David Bordwell, Janet Staiger and Kristin Thompson (London: Routledge, 1985), 47.

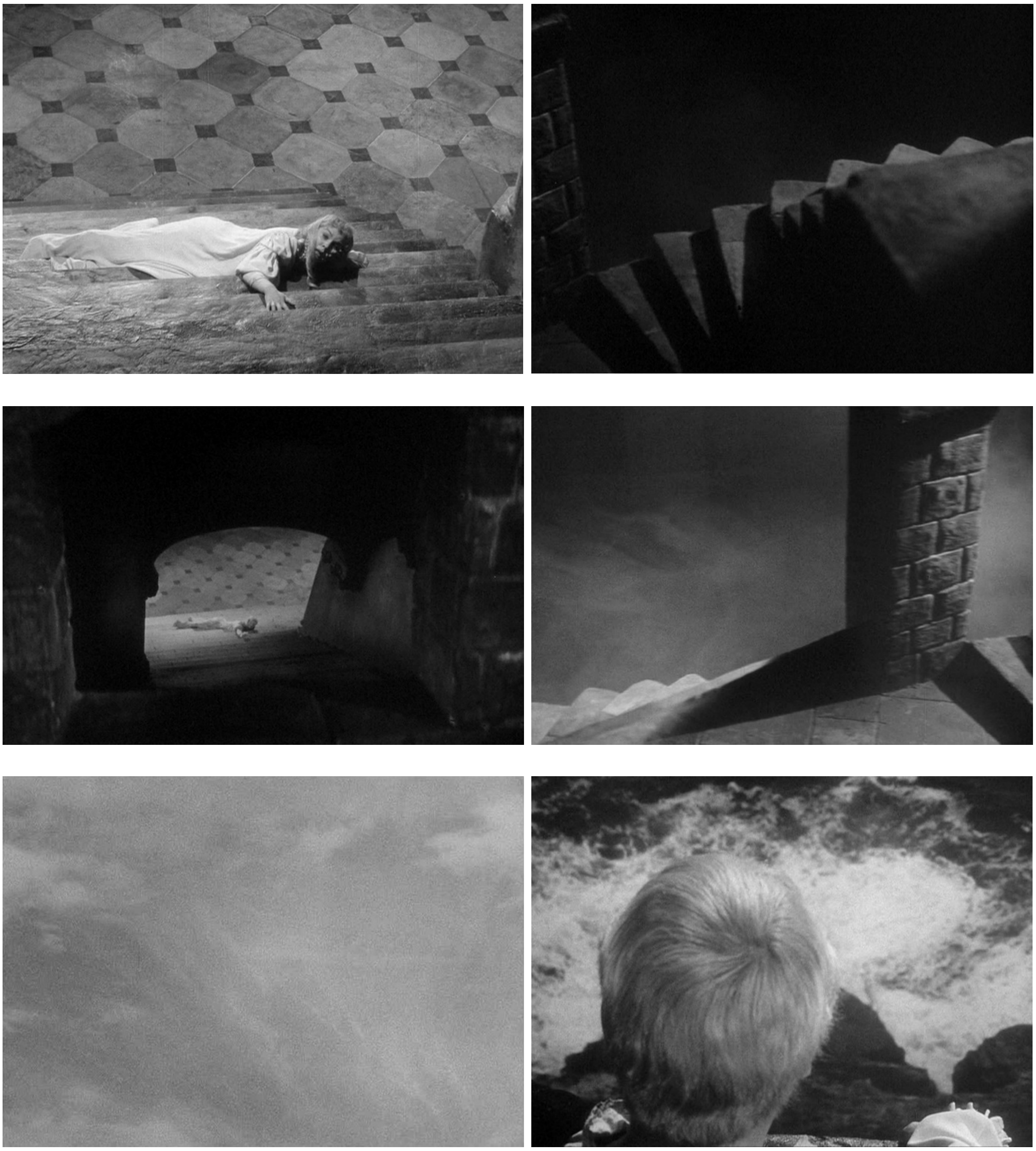
Contrary to long shot durations, films exhibit music-images for shots of shorter durations to stretch story time. A great example of this technique is the preceding shots of the famous ‘To be or not to be’ soliloquy from Lawrence Olivier’s *Hamlet* (Olivier, 1948). After Hamlet throws Ophelia onto the steps on learning that Claudius and Gertrude are using Ophelia to spy on him, the camera falls behind Hamlet as he climbs the spiral staircase to deliver his soliloquy to a turbulent sea. Instead of following Hamlet in one fluid motion and shot, there are many shots stitched together cleverly by the chiaroscuro contrast of light on the steps and the shadows of corners. This motion of transitioning shots seems to speed up narrative time, due to the scurry of strings and brass rising in pitch and volume reaching a crescendo through unstable chromaticism. The music seems to push the camera up the staircase faster than the physical apparatus could be carried. Hamlet races up the stairs after feeling betrayed by his parents, leading to the pinnacle moment of “To be or not to be.” The music-image enables the illusion of faster physical movement for a dramatic and unstable scurry up the stairs to match Hamlet’s feeling of betrayal. As Ben Winters states, music unifies the montage of images to narrate an external perspective to the events the spectator is witnessing, passing what would seem a long journey up the stairs as seen in *Hamlet* in a more compressed time frame.[199] [See Figures 18-23]

---

[198] Lida Oukaderova, "I Am Cuba and the Space of Revolution," *Film & History* 44, no. 2 (2014): 11.

---

[199] Winters, "The Non-Diegetic Fallacy: Film, Music, and Narrative Space," 236.



Figures 18-23. A sample of the many shots stitched together to reach Hamlet at the top of the staircase in a hurry in *Hamlet*.

Parallel to the duration of time in shots and sequences, tempo (the speed of the music) can govern the ‘tempo’ of the music-image. Where film music, in general, affects the sense of narrative duration (‘story time’), the specific tempo, the speed of the music, can manipulate that story time regardless of shot or sequence duration. Vsevolod Pudovkin on the synchronous nature of the introduction of film sound to the existing visual image states: "Always there exist[s] two rhythms: the rhythmic course of the objective world and the tempo and rhythm with which man observes this world.”[200] It is easy to forget this principle, that the objective truth of the duration of the image exits with or without the experience of the spectator with the world. This is how the spectator differentiates a dramaturgy of nature from nature itself, and why Kathryn Kalinak argues that film music tempo provokes a physiological response to the existing nature of images.[201] For instance, in the case of the film *In the Mood for Love* (Wong Kar-wai, 2000), the visual image slows down to complement the slow tempo of Shigeru Umebayashi’s famous “Yumeji's Theme”. However, this explanation of the nature of the image downplays the tempo of spectator experience. Slowing down the image exaggerates bodily movement, as seen in the film experiment *Olympia* (Leni Riefenstahl, 1938). Exaggerating bodily movement, that is, giving ordinary, natural activity of the body extended or *elongated* screen time, not natural to that activity, dramatizes the image, giving it fuller attention. In the early scenes of Mr Chow and Mrs Chan’s early encounters passing by each other, the melancholy waltz (a tempo of ¾ time signature) of Yumeji’s Theme captures their bodily movement, but a waltz is incomplete without a dancing partner. This is why Yumeji's Theme only plays during the slow-motion montage of the two characters’ movements juxtaposed to keep them together. Yumi Braester points out that Yumeji’s Theme repeats through this interaction to bring attention to itself, as the exaggeratedness of bodily movement hijacks the spectator's attention.[202] The image choreographs audio-visual dance routine through montage, creating a shared gaze of Mr Chow and Mrs Chan, a dance partnership held together by images rather than bodies. Interestingly, this slow-motion montage of Mr Chow and Mrs Chan connected through the splicing of images and spaces is never projected in total silence. Audio-visual tempo is achieved through this physical waltzing of images. What this example elicits is that tempo created by the visual form (the slow-motion cinematography) and the music (the melancholy waltz), forces the spectator to configure audio-visual meaning, that is, a shared tempo created through the ambiguity of two separate outputs.

---

[200] V. I. Pudovkin, "Asynchronism as a Principle of Sound Film," in *Film Sound: Theory and Practice*, ed. Elisabeth Weis and John Belton (New York: Columbia University Press, 1985), 87.
[201] Kalinak, *Settling the Score: Music and the Classical Hollywood Film*, 9.

On the topic of montage in *In the Mood for Love*, the music-image can create a sense of nondiegetic rhythm, creating meaning through faster or slower rates of story time from outside (metric) or within (rhythmic) the narrative space. I argued that Eisenstein's Vertical Montage theory of audio-visual rhythm using 12 shots from his film *Alexander Nevsky* is problematic because it does not consider the complexity of musical form when determining that 'music's vertically synchronous relationship with the horizontal lines of existing montage' creates polyphonic rhythm. Many scholars have dismissed Eisenstein's highly original practice and theory with composer Sergei Prokofiev, but I also argued that Eisenstein's attempt to acknowledge that there is indeed an audio-visual rhythm brings attention to the music-image that requires the spectator for it to exist. Lee Jacobs argues that film rhythm involves how duration is structured.[203] I also suggested earlier that film duration is defined through the relationship of music and visual image in narration; and thus, even though Eisenstein's ambitious theory falls short, his initial idea rings true. Graig Uhlin argues that film rhythm "*describes* how the moving image embodies time, gives it material form, by providing a patterning or structuring of time relatively independent from narrative plotting."[204] On this definition, nondiegetic music operates inside and outside the narrative space, as part of the image. However, Uhlin does not acknowledge the 'narrative patterning' involved, that is, the dramaturgy of the nature of film as experienced by the spectator. This simple acknowledgement (that Uhlin does not make) considers that film can only be analysed if there is a spectator who experiences the world separate from the nature of the film form.

---

[202] Yomi Braester, "Cinephiliac Engagement and the Disengaged Gaze in In the Mood for Love," in *A Companion to Wong Kar-wai*, ed. Martha P. Nochimson (West Sussex, UK: Wiley Blackwell, 2016), 212.

[203] Lea Jacobs, *Film Rhythm after Sound: Technology, Music, and Performance* (Oakland, California: University of California Press, 2015), 51.

'Metric' audio-visual montage concerns the temporal rate of the images in a sequence in accordance with the timing of the nondiegetic music. Through experience, it seems the aural component, rather than the visual component, is in control of the image. Michel Chion agrees that "music can directly express its participation in the feeling of the scene, by taking on the scene's rhythm…."[205] The Soviet filmmakers of the silent film era capitalised on and theorised this phenomenological experience of the image, especially Sergei Eisenstein in *Battleship Potemkin* (Eisenstein, 1925). Royal S. Brown argues that the film calls attention to the montage process through its dialectical approach, and that music has no choice but to cooperate with the metric patterning of Eisenstein's process.[206] After all, Eisenstein has said himself that music is part of the montage process.[207] This montage process builds a language through the creation of shots, rather than through the sum of its juxtaposed parts, with music joining the creation process.[208] With montage as a language in this 'thought cinema', music as part of its very construction of language seems justified. This is not simply true of the silent Soviet era, as film has evolved into a medium of communication through this very creation. Rather, metric audio-visual montage, that is, the creation of shots and music that build rhythm, is present throughout the history of cinema. A notable example is the ‘training montage’ in sports films, used frequently in the *Rocky* franchise, in which a main character prepares for a major narrative event through intensive yet mundane exercise routines. In these montages, it is not so much the activities performed by the character that is important (the images), it is the construction of rhythm built through cutting the action to music, as if the image itself is in training.

---

[204] Graig Uhlin, "Film Rhythm and the Aporetics of Temporality," *Quarterly Review of Film and Video* 36, no. 2 (2019): 110.
[205] Chion, *Audio-Vision: Sound on Screen*, 8.
[206] Royal S. Brown, "How not to think film music," *Music and the Moving Image* 1, no. 1 (2008): 5.
[207] This is a citation of Sergei Eisenstein by Claudia Gorbman, in Gorbman, *Unheard Melodies: Narrative Film Music*, 15.
[208] Eisenstein, *The Film Sense*, 7-8.

These audio-visual montages arise also when a character is compelled to complete a task. In the ‘peach montage’ from *Parasite* (Bong Joon-ho, 2019), the Kim family plot to replace the Park family’s housemaid, and the sequence of shots within the montage is orchestrated and bound by the rhythm of the music rather than the narrative task within the space. Similarly, in the ‘Running One’ sequence in *Run Lola Run* (Tykwer, 1998) in which Lola hangs up the phone call with her boyfriend Manni for the first time and runs to the bank for the first of three times in the film, music works with diegetic sound to decide *when* to splice imagery, making Lola arrive at the bank within the duration of the music rather than the actual time it would take her to do so. In a wide shot of Lola running a straight line from behind, a series of jump cuts are made for no other reason than to accelerate her movement to the rhythm of the music, as if the cuts are the bars of musical phrasing.

On the other hand, ‘rhythmic’ audio-visual montage concerns movement within the image in accordance with the movement of the music, making the diegetic principles control the rhythm of the image. The most famous example of this is the ‘Odessa Steps’ sequence from *Battleship Potemkin*, where the temporal rate of images is in rhythm with the “syncopation of the soldiers’ marching descent against the metrical cutting.”[209] Since it is a silent film in which a soundtrack is not bound to the celluloid, many iterations have been made for the film, but all seem to apply the holistic intensity of the total movements in this scene.[210] For instance, the Edmund Meisel original score “parallels the implacable descent of the Tsarist troops with a relentless percussive pounding and chromatic phrases in the strings” parodying military tattoos, a grotesque theme that is experienced in all versions of the scene.[211] Instead of the pounding aural rhythm orchestrating the image, the diegetic movement choreographs the music and, in turn, the holistic image. However, to reiterate a point, this ‘rhythm’ does not exist without the subject’s (aural or visual) counterpart. For such an audio-visual rhythm to exist, there needs to be the existence of movement, a fundamental mode of time, by which film music, and film itself, live.

---

[209] Uhlin, "Film Rhythm and the Aporetics of Temporality," 112.

Audio-visual tone and texture

Since the introduction of film sound, film music has enjoyed the success of staples that define its use in the medium. Suggesting a scene such as a hero crashing through the window of the villain's hideout to rescue the heroine, immediately a major fanfare from horns comes to mind announcing the hero's entrance and setting the tone for his heroic rescue. Then a climactic cadence in the strings following a triumphant major 7th (a V7) resolves the moment the hero gets to kiss the heroine. Frank Lehman argues that this awareness of tonality is due to the idea that it "sounds like film music".[212] What does Lehman mean by this? If the music-image is considered, Lehman suggests that film music tonality exists from the harmony created in audio-visual interlocking. "If film is an apparatus designed to transport the filmgoer, it is hard to imagine that apparatus functioning smoothly without the aid of tonal design."[213] The OED defines musical 'tone' as "a sound of definite pitch and character produced by regular vibration of a sounding body; a musical note."[214] Tonality then refers to the "relationships existing between tones or tonal spheres within the context of a particular style system."[215] Basically, if the chord of C minor is played, C is the root, minor is the modality (or quality) of the chord, and C Minor is the tonality.[216] Tonality in film, on a technical level, considers how light registers on the film.[217] This tonality is a visual register, with lighting belonging to the mise-en-scene, controlling contrast levels, exposures, and use of colour. In the narrative film, these tonal decisions are made to provoke a certain response or mood from the spectator, or to draw or avoid attention. With aural and visual images interlocked, audio-visual tonality refers to harmonizing of separate channels of music and visual tone (this can include congruence and dissonance) to create one channel of mood or thought. Audio-visual tonality can include tension, release, pitch and volume, through the assemblage of cultural codes and conventions.[218] For instance, at the beginning of *North by Northwest* (Hitchcock, 1959), at the end of the title credits, the overture ends in a repeat of shrieking dominant seventh chords. E major can be heard overlaid with B-flat in the horns and trombones creating a tritone. These tense dominant chords need releasing in a cadence to a tonic (first chord in a harmonic series), but they never do. The spectator hears this as Alfred Hitchcock makes his iconic onscreen cameo missing his bus, a narratological effect of tension arising and never releasing as we can all relate to the sudden panic and disappointment of missing a bus. Having said that, it is worth noting that there are different cultural understandings of tone. As Leonard Meyer points out, “different cultures as a rule have different style systems, different ways of organising musical space. In one system the normal repertory of tones may be five, in another seven, and in still another only three.”[219] This suggests that non-Western tonal cultures, for example, may see the tension in the example above in another narratological light due to different style.

---

[210] This only applies to the multiple versions of the film I have encountered.

[211] D. Travers Scott, "Intertextuality as ‘Resonance’: Masculinity and Anticapitalism in Pet Shop Boys’ Score for Battleship Potemkin," *Music, Sound, and the Moving Image* 7, no. 1 (2013): 61.

[212] Frank Lehman, "Tonal Practices," in *Hollywood Harmony: Musical Wonder and the Sound of Cinema* (New York: Oxford University Press, 2018), 15.

---

[213] Ibid.

[214] OED Online, "tone, n.," Oxford University Press, https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/203150?rskey=J3vEQI&result=1.

[215] Leonard B. Meyer, *Emotion and Meaning in Music* (Chicago; London: The University of Chicago Press, 1956), 214.

[216] C minor can also refer to the key, where C is the tonic and minor is the modality.

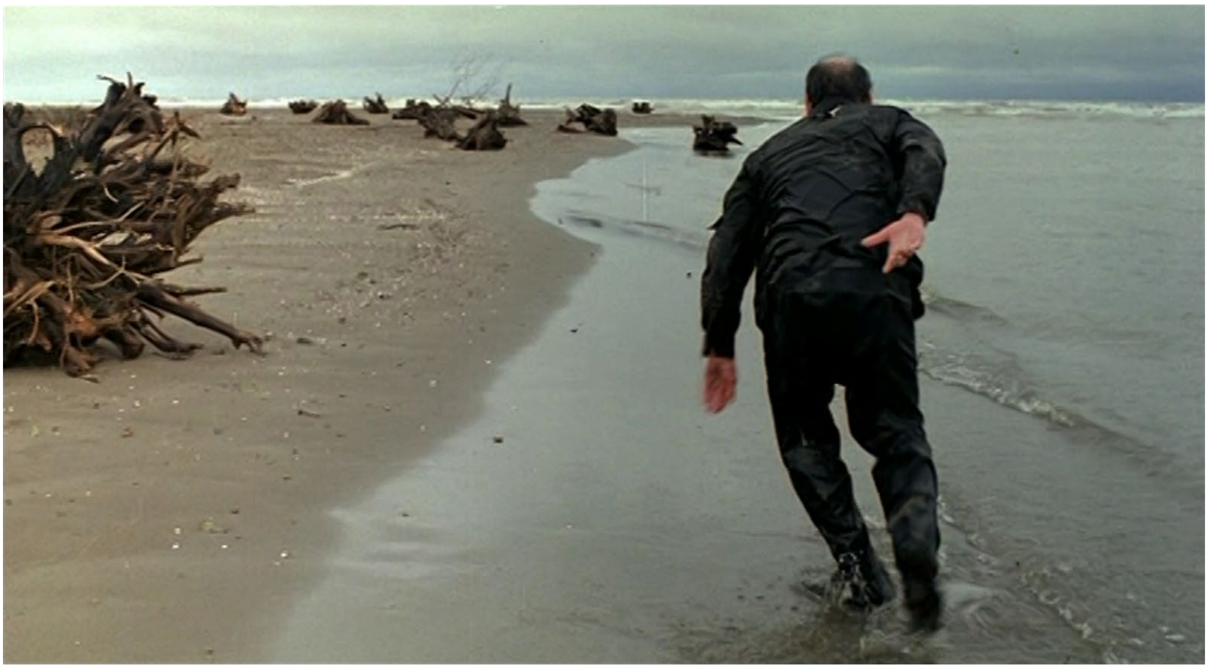
[217] David Bordwell, Kristin Thompson, and Jeff Smith, "The Shot: Cinematography," in *Film Art: An Introduction* (New York: McGraw-Hill Eduction, 2019), 159.

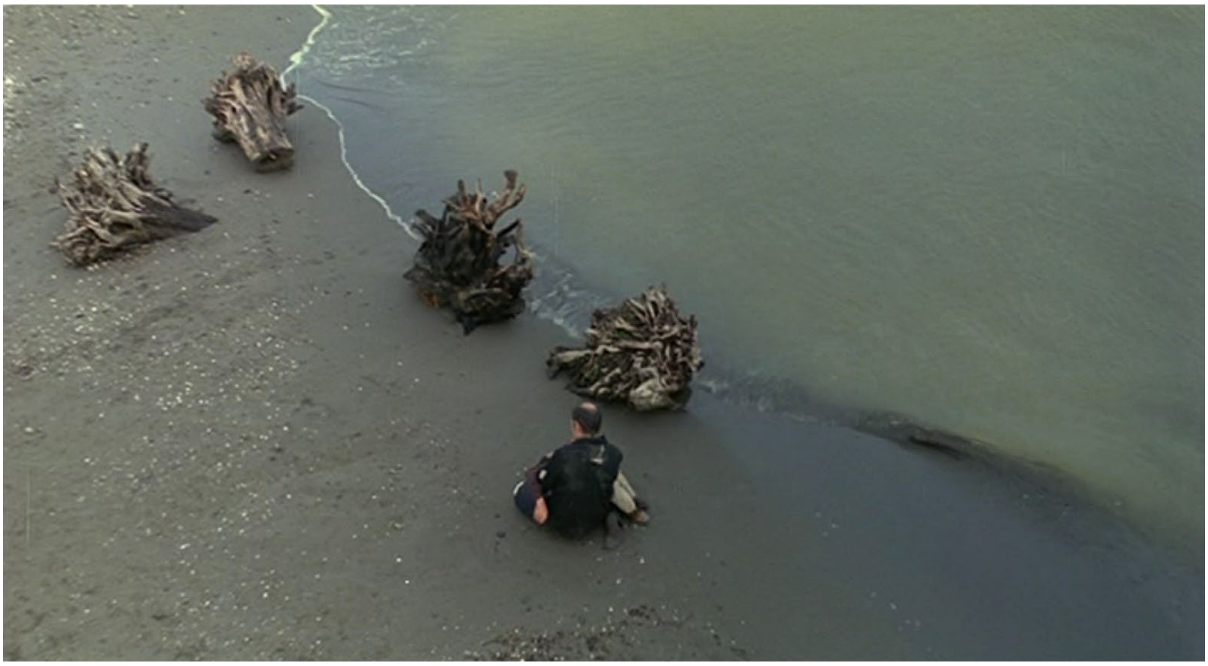
[218] Notice that these examples can be attributed to either music or film; it is when they are interlocked that a new layer of meaning is made, based on existing attributes.

Regardless of culture, audio-visual tone can be experienced in almost any film with music for purposes of creating a certain narratological mood. In the Iranian film *The Color of Paradise* (Majidi, 1999), music is used sparingly to comment on the relationship between Hashem and his blind son, Mohammad. Throughout the film, Hashem neglects his son due to the shame and burden of Mohammad’s blindness. Mohammad wants to experience ‘God’, which to him is the visual beauty of the world, a beauty that he cannot experience. At the end of a film full of glorious and melancholic moments, Mohammad is washed away down a rushing river on the trip back to Tehran. Hashem attempts to swim after him only to wake up later on a beach. Here music is cued for one of the only times in the film, as there is a panning shot of Hashem waking up to seagulls and a deserted beach. The strings swell as he sees Mohammad’s body. Are we experiencing Mohammad’s tragic death or sympathy for Hashem’s loss? It is unclear, but the image is filled with the minor mode of piano and strings to greet the bleak, grey colours of the shoreline with melancholy. The camera begins up high on a crane and tracks into Mohammad’s hand as Hashem holds his body. The innocent, angelic voices of a children’s choir pierce the melancholy in a hymn to this loss. However, as Mohammad’s hand begins to move and lights up gloriously as if touched by God, the image and film close in a tone of ambiguous hope. The hymn is sorrowful, but its tone suggests hope. That final crane shot now has double meaning: it pinpoints the isolation of grief, but perhaps it is the point of view of the hand of God.[220] [See Figures 24-26]

---

[219] Meyer, *Emotion and Meaning in Music*, 132.





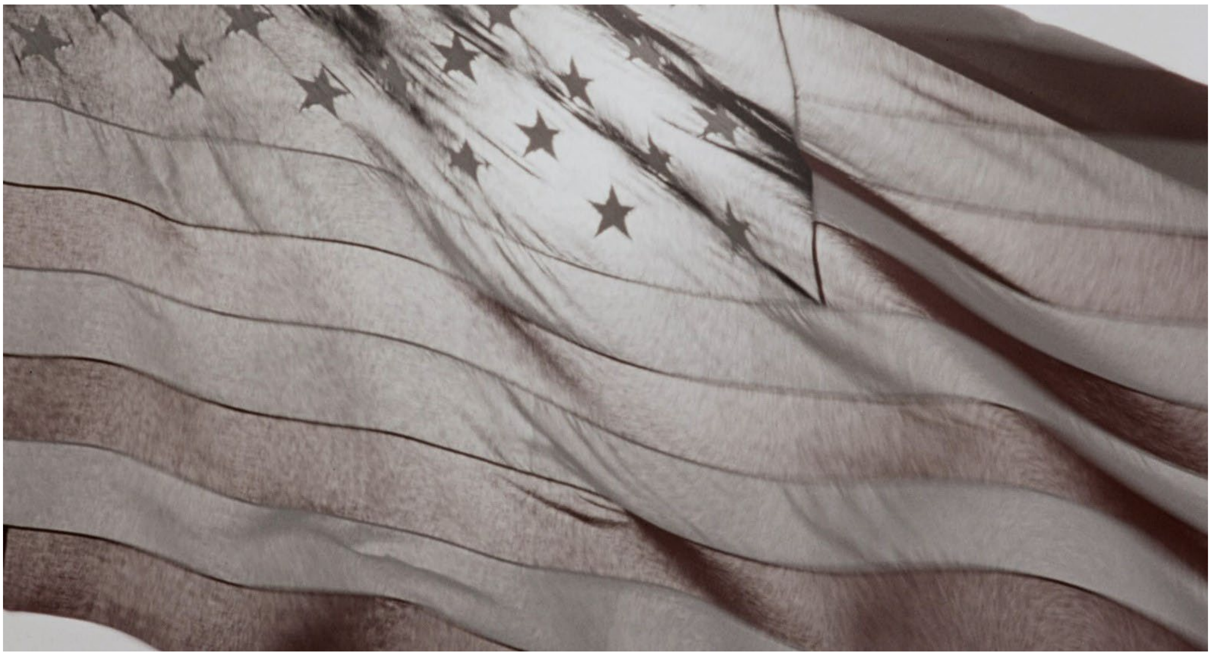
---

[220] This soundtrack was in the version of the film I watched online on a streaming service. Interestingly, a different stripped back soundtrack is used on the DVD version of the film with only strings cueing as Mohammad's hand begins to move transitioning into the end credits. After experiencing both versions, I had completely different experiences of Hashem's response to Mohammad's death.

Figures 24-26. Hashem notices his son, Mohammad washed up on the shoreline, and goes to hold Mohammad's body as a crane shot tracks into Mohammad's hand which lights up gloriously in *The Color of Paradise*.

This style of audio-visual tone from *The Color of Paradise* is on display everywhere in film and for varying degrees of mood to be experienced. In the opening of *Midnight in Paris* (Allen, 2011), the diegetic instruments of the Parisian streets are inserted into the nondiegetic image, constructing a romantic, jazz depiction of Paris, which helps set up the film's nostalgic narrative. Cleverly, in the opening of *Citizen Kane* (Welles, 1941), two depictions of Charles Foster Kane's death are portrayed with completely different moods. The first is the poetic connotation of the bare landscape of Xanadu palace through the montage of wide angles zoning in on Kane's deathbed. The loneliness of Kane's last moments ending with him muttering "Rosebud" feels sombre through held minor key brass notes. The following 'News on the March' sequence is a prosaic depiction of the news of Kane's death with the use of triumphant major key strings ending in 'funeral march'-style trumpets. The newsreel-style narration completely lacks mood, leaving news reporter Rawlston to wonder who Kane really is, triggering the film's search for meaning. Yet the prior sequence has already teased the spectator with "Rosebud."

Saving Private Ryan (Spielberg, 1998) and *JFK* (Stone, 1991) offer a variation on this theme of poetic and prosaic contrast by opening with the limiting harmonies of marching drums and military bugles to signal American justice. *Saving Private Ryan* constructs a connoted image of bittersweet patriotism with a transparent flag blowing in the wind while *JFK* constructs a denotated image of that bittersweet patriotism through the archival footage of President Eisenhower's farewell address to the US nation, in which he declares an imbalance of liberty and the might of military influence. [see Figures 27-28]



Figures 27-28. The connoted, poetic opening image of *Saving Private Ryan* (top), and the denotated, prosaic opening image of *JFK* (bottom).

On the opposite side of the spectrum, aural and visual stimuli harmonise and interlock so closely that the tone is beyond doubt. This can range from the excitement and awe of witnessing the grandiose wide-angle image of Isla Nublar from the air for the first time with high register trumpet fanfare in the major key in *Jurassic Park* (Spielberg, 1993) to the heartache of Emma's death in *Terms of Endearment* (Brooks, 1983). In the intimate, bleak space of a hospital room, Emma and her mother, Aurora connect one final time in close-up, shot reverse shot with a flute, harp and strings calling and responding to each other in minor key harmony, matching their final farewell but creating one layer of meaning for interpretation.

In another layer to this audio-visual tone, musical instruments can be played in various ways, and different instruments can play the same note, chord, phrase, or melody in different ways, which distinguishes the quality of the tone or the texture of an image. In studies of music, this is called timbre.[221]

Using stringed instruments for extended note duration is a highly expressive practice.[222] In the final scene of the Australian film *Kokoda* (Grierson, 2006), Japanese soldiers attack Australian-held positions on the Kokoda track at night. This final scene is constructed in four parts: the tension and release of the surprise attack by the Japanese, the 39th Battalion's colonel gives a sentimental, patriotic speech to the surviving soldiers the next morning, a return to the attack of the night before with the colonel's speech narrated over the top, and the withdrawal of the soldiers. The opening attack holds a 7th chord in the strings and releases when the first gunshots are fired, with the low visibility of night-time and awkward close-ups of the soldiers' rifles making for a tense surprise ambush. However, as soon as the colonel delivers his rousing speech (reminiscent of the ‘band of brothers’ speech from Henry V) to the men in broad daylight the next day (a few shots later), the image from the night before is affected. The image returns to that night attack, this time lowering the volume of the diegetic gun sounds and inserting the densely layered stringed orchestra playing sustained notes in the minor key to saturate the image with sentiment, matching the heartfelt words of the colonel: “gallantry…courage…fortitude…finest soldiers I’ve ever seen…history will remember you.” The same diegetic image of night attack has turned from tension and horror to acts of gallantry, courage, and fortitude through the sentimental choice of adding layers of strings.[223] This device and effect will bleed into the final section through a cross dissolve edit, keeping the music intact as the soldiers exit the battlefield through the sad and sombre atmosphere of rain in a forest. A similar effect is achieved in the opening battle of *Gladiator* (Scott, 2000), in which the diegetic world slows to a distorted slow-motion and is drained of battle noise as the sentimental music of a stringed orchestra marks the turning point of victory. Mood is summoned by the audio-visual tone in play with the narrative outcome of the scene.

---

[221] OED Online, "timbre, n.3," Oxford University Press, https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/202089?rskey=Kpaswo&result=3.
[222] Meyer, *Emotion and Meaning in Music*, 66.

In contrast to this cloudiness of emotion, texture can also be thinned out in an image by using solo instruments only. In the film *Predator* (McTiernan, 1987), a single echoing tribal drum playing patterns of triplets is heard during the opening jungle scenes, creating a hollow tension of the lurking presence of the Predator prior to its first big reveal. Every time this occurs, there are random shots of trees and bushes, usually shown through zoom lenses. There is no pivot to these shots as they cut back to the soldiers wandering, as the Predator has not yet been revealed, so the images are ominous echoes in themselves due to the lack of resolution. As soon as the Predator is revealed, the drum pattern disappears and the jungle loses its tightness, releasing this echo.

---

[223] Brian Macfarlane, in an article written on the film, suggests that this is done ‘contrapuntally’, though is concerned that the scene is left ambiguous as to how it is meant to be read. However, with the music overlayed, it is clear that the image is clinging onto the colonel’s words, to honour the fighting efforts of the Australian soldiers in battle. See Brian McFarlane, "‘Kokoda’: Lost and Found on the Trail: To Australians of a Certain Age, and Even to Some Younger, the ‘Kokoda Trail’ Has Acquired the Sheen of Legendary Significance," *Australian Screen Education*, no. 42 (2006): 12.

A similar situation occurs in the theatrical cut of *Troy* (Peterson, 2004), in which a solo tribal drum is used for the fight between Achilles and Hector in front of the Trojan gate. It is extremely tense as there is no structure or phrasing to the drumming, nor known outcome to the fight, in a mixture of close-ups and wide shots of the fight and of the Trojans watching from the gate. Their emotion is ambiguous as they hold their breath to this tightening of the image. Interestingly, this tension is completely lost in the director's cut of the film, in which a string and wind orchestra replaces the drum for the exact same shot sequence; this textual choice has completely changed the tone of the image due to the music alone. Similarly, if stringed instruments were substituted for the mechanical pounding percussion and electronically synthesised score in the shots of the Terminator's first arrival in the year 1984 from the future in *The Terminator* (Cameron, 1984), the image would lose its futuristic, robotically macho energy. The choice of musical instruments with certain visual elements or shots can completely alter the narratological expression understood by the spectator.

Audio-visual space

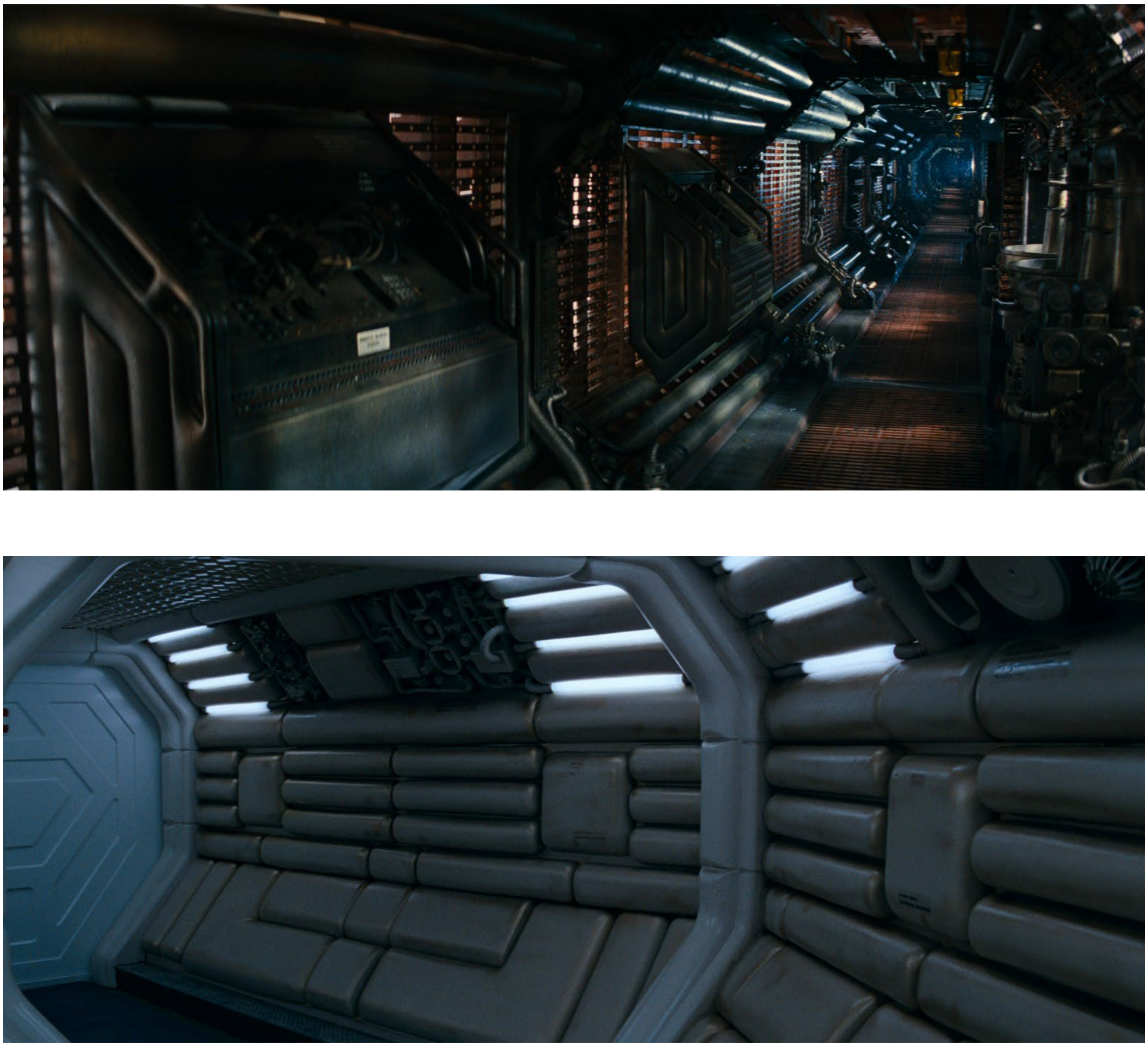
Since the definition of nondiegetic film music refers to the 'outside [of] the narrative space,' it is hard to imagine film music situated outside the narrative space influencing that space. The construction of the narrative space is in the film itself, as much as the space. I want to end this chapter on the notion that began it, that it is difficult to distinguish between the binary notion of diegetic and nondiegetic music-images when all relate to the narrative space-time in some way. This is why diegesis is distinguished by the experience of the film.

Narrative space does not exist without the construction of the film image which includes music when present, making it an audio-visual space.

If the music cannot be understood to be *within* the narrative space, that is, the narrative subjects are able to interact with the music in some way or the music is seen to be performed within the space, then it governs the space they populate instead. Since the experience of this governing interlocks with the visual representation of this space, audio-visual space is defined by how it is presented. This means that nondiegetic music does indeed interact with the diegetic space, but it *constructs* rather than inhabits that space. However, when cinema has the effect of "…flattening the real on a two-dimensional plane…" from the lack of speech and ambient sound, as in the silent era, music supplied the psychological effect of three-dimensional reality to the film world.[224] When there is a lack of soundscape, artistically or technologically, music can appear to inhabit the space it constructs. For instance, for the opening sequence inside the Nostromo spaceship in the original *Alien* (Scott, 1979) with its lack of narrative input due to the unpopulated, eerie hallways as the crew sleeps, the music has freedom to fill in the physical space, as well as the narrative space, making the discordant, echoes of strings seem to bounce off the walls. The lack of melody in the repeated pairing of A-flat – F quavers played by a flute are sequenced in a decrescendo as the camera moves through the hallway, as if the A-flat is blown and hits one wall and the F is blown and hits the other, moving further and further away into the void of the empty spaceship, emphasising that echoing emptiness (see Figures 29-30).

---

[224] Claudia Gorbman cites scholars who argue that the silence of actors has a psychological effect of flattening the fictional world into a "two-dimensional plane." Music compensated for this silence and brought back the fictional world's three-dimensional reality. See, Gorbman, *Unheard Melodies: Narrative Film Music*, 37.



Figures 29-30. The empty hallways of the Nostromo spaceship at the beginning of *Alien*.

As seen in the example from *Alien*, David Yacavone cites Ben Winters arguing that audio-visual space is hard to define in the binary distinction of diegetic and nondiegetic as, through spectatorial experience of film, they are both part of the story's world, the presented reality.[225] However, Yacavone further argues that nondiegetic music is not part of the narrated, represented, or fictional reality, with which I agree.[226] The music does not literally bounce off the walls in *Alien*, nor is it being played within the space; it is more a psychological effect as the spectator turns a two-dimensional plane into a three-dimensional space, or a flat image into an audio-visual depth of field.

---

[225] Daniel Yacavone, "Spaces, Gaps, and Levels: From the Diegetic to the Aesthetic in Film Theory," *Music, Sound, and the Moving Image* 6, no. 1 (2012): 23.

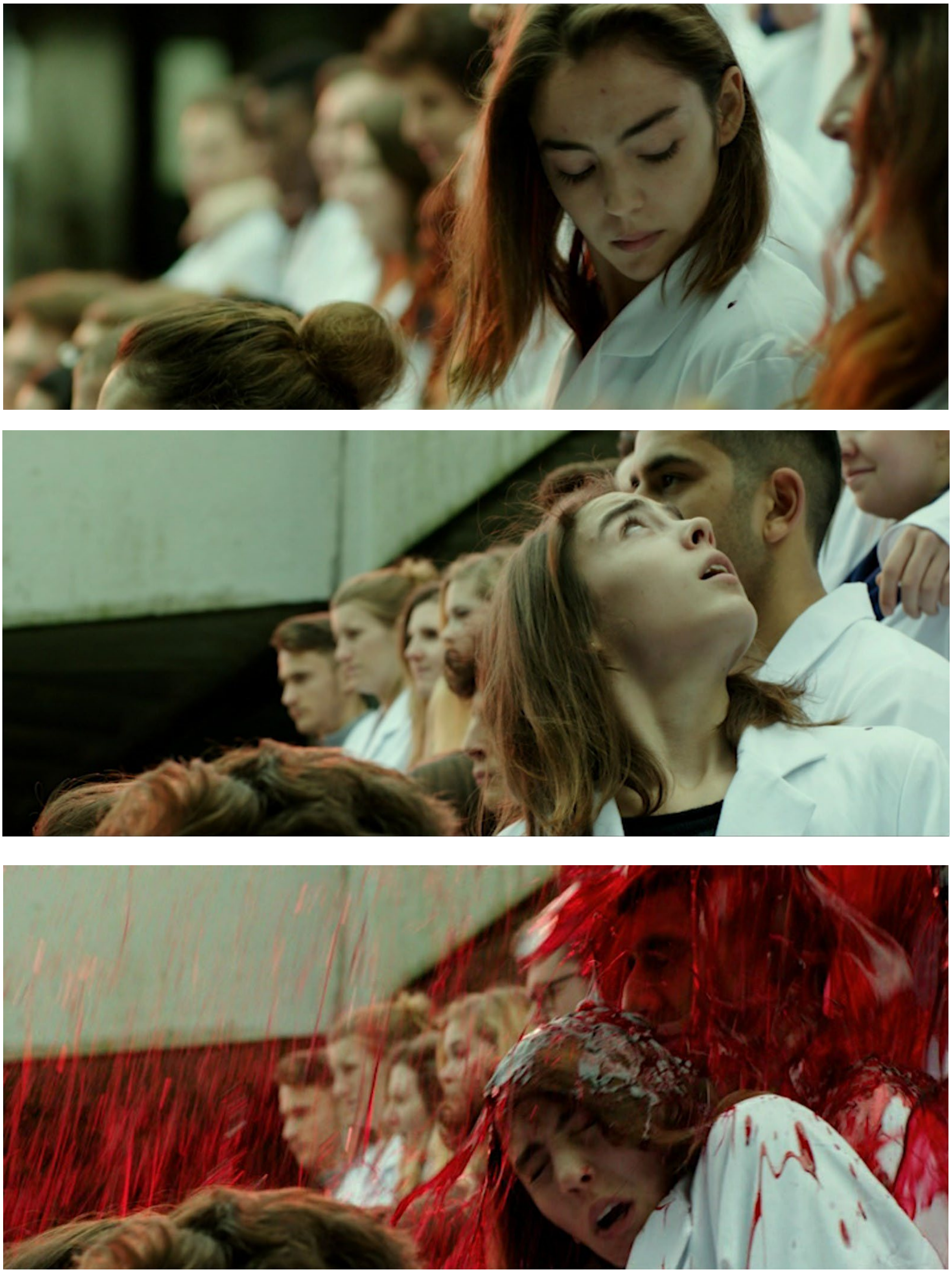
Another example of this three-dimensionality is in the film *A Hidden Life* (Malick, 2019). At first glance the opening images of the film appear to be a simple tonal harmony of strings and romantic Austrian vistas, but on further inspection the wide-angle of the fisheye lens being used wraps around the space to envelope the love and lifestyle of Franz and Fani in the swelling romance of strings. However, it is far-fetched to believe the music is located physically within that space due to this audio-visual paradox. I will expand upon this notion of construction to explore how the nondiegetic music-image constructs an audio-visual space functioning in relation to the spectator experience of the film.

The nondiegetic interacts, even synchronises, with diegetic narration to construct spatial awareness for the spectator and narrate the world. In the film *Raw* (Julia Ducournau, 2017), when Justine and her first-year classmates at the veterinary school are welcomed by existing students, including her sister, Alexia, they are splattered with blood from above, referencing De Palma's *Carrie* (1977) (see Figures 31-33). Diegetic sounds except for the sounds of blood splattering on the students' heads evaporate into thin air, creating a hollow vacuum. The sounds provide room for distorted guitars in multiple pitches (and thus, layers) to audibly screech and distort the space. The guitar score seems to come from the space like the blood, as all that is heard is blood and guitar in a sound vacuum; but for the spectator, it constructs the narrative image of meaning. Justine is shown in close-up during this sequence; it is a pivotal moment in the film as she is overcome with desire for meat despite being a vegan.

---

[226] Ibid.

The music-image creates a space of moral and ethical disgust through Justine, a space of distorted guitar-fuelled blood spatter. Sometimes specific instruments that resonate environmental properties are used to characterise a certain landscape. Although the landscape is diegetic, the characters inhabit and experience the space, and this characteristic is reserved for the spectator who experiences the narrative landscape.



Figures 31-33. Justine and her classmates are splattered with blood from above in *Raw*.

In the example of *Hostiles* (Cooper, 2018), a very unfamiliar instrument, the Yaybahar (there is only one in the world), with its highly unusual, unearthly sound, is used to create an "abstract, alien or other" landscape.[227] *Hostiles* is a revisionist Western that intends to subvert the standard format of the Western by depicting a harsh, unrelenting 'west' of the American frontier in the 19th Century. The Yaybahar, like the guitar in *Raw*, conveys this sentiment with its inhuman narration of a space conquered by human lawlessness and violence. It creates a resonating chamber of this harsh landscape, giving the ground, the trees, the sky, as well as the violence, the heroes and the criminals that inhabit the landscape an abstract, alien, foreign spatial chamber.

With this three-dimensionality that nondiegetic film music can create, it is hard not to think that the music is located within the space. Instead, the fact that nondiegetic music *creates* the ontology of this three-dimensional space makes the nondiegetic music-image a construction of that very three-dimensional audio-visual space. When Andre Bazin argued that images preserving memories in "…the image helps us to remember the subject and to preserve him from a second spiritual death," that preservation is constructed within the constraints that make the image possible, that is, by the camera taking the photograph or recording the space.[228] In film, the construction of audio-visual 'space' relies on the spectator's experience to make sense of its construction. As Stanley Cavell asserts on spectatorial film experience, the spectator puts their scepticism aside to make sense of the narrative reality in relation to spectator reality. [229] As difficult as it is to define the binary nature of diegesis based on existing scholarship, nondiegetic music-images are defined by the way in which the spectator experiences a certain film, providing vast potential for analytical inquiry.

---

[227] Mark Kermode, "Kermode Uncut: Unknown Instrument," YouTube video, https://www.youtube.com/watch?v=p1t4EUXjVa4.
[228] Bazin, *What is Cinema? Volume 1*, 11.
[229] Cavell, *The World Viewed: Reflections on the Ontology of Film*, 188.

Conclusion

In this chapter, I have provided an overview of film music that is intrinsic to the construction of the audio-visual image. This is an image that, according to Gorbman, does not exist within the film diegesis yet constructs a certain audio-visual narrative in each film. In the analysis of audio-visual space, however, I have demonstrated that the binary definition of film music's location outside and inside the diegesis becomes unstable. A reconsideration of nondiegetic film music has demonstrated that there is a certain audio-visual harmony, time, tone, and space in film experience that can completely change the coordinates and potentiality of film analysis. I will explore these issues further in the next chapter with an analysis of *Jaws* (Spielberg, 1975).

Chapter 4: A Music-Image Analysis of *Jaws* (Spielberg, 1975)

In this chapter, I provide an extended analysis of a nondiegetic music-image in *Jaws* (Spielberg, 1975). I will demonstrate that aural and visual stimuli cannot exist separately; instead, they exist within the same phenomenological plane, and are interlocked as a singular, hermetic image form.[230] With perhaps the most iconic film score in cinematic history, *Jaws* presents an obvious and rich object of analysis. The film enjoyed critical acclaim and popular success, smashing box office records and earning three Academy Awards, including one for John Williams's score. Another remarkable point to consider is that, for a film about a killer shark terrorising a small American summer-vacation town on an island, the shark does not appear for a good hour into the film. Yet the shark has an overwhelming presence and effect on the spectator.[231] For the roughly 43 music cues in the film, with most occurring once the shark has been revealed in the second half of the film, there are only a handful that do not associate with the shark. This is not to say that the score predominantly revolves around, or 'represents' the shark, as Giorgio Biancorosso claims, but rather that the film score and in turn the film's music-image participate in world and perception building, with the shark occupying its space-time.[232]

The *Jaws* title sequence

Before the film opens, all the spectator is aware of narratively is that the film is probably about a shark, assuming they have noticed the film title and/or poster.

---

[230] See Chapter 2 for a breakdown of how aural and visual stimuli are interlocked in this singular, hermetic image form.

[231] This is not to disregard the sound of the image, which is important here, as sound also constitutes and contributes to narrative content and meaning. Sound is not ignored in this film analysis, but this is a film music analysis, through my model of the music-image.

[232] Giorgio Biancorosso, "The Shark in the Music," *Music Analysis* 29, no. 1-3 (2010): 306.

After the Universal Studio logo comes and goes, the film fades to black, but listening closely, there are muffled sounds of orcas (killer whales) which give this blank image an ambience of the natural world beneath the sea. The image of the black abyss lasts for a few seconds, gradually increasing in volume, until a not-so-natural sound of strings in unison play a low E and F. There is no clear indication if the E that begins the music in *Jaws* is tonal as film music can commonly begin atonally. However, if the opening E-F were to be considered tonal, in a tonal chord progression, the E and F are situated only a semitone apart. Since the E is the first note in the melodic progression played by the strings, the note is assumed to be part of the tonic chord. The note E is the 5th note in an A minor triad, which could make the tonic chord an A minor chord. The E moves to the F, the flat-sixth of A minor, which makes this a melodic progression from the 5th to the flat-sixth in the key of A minor.[233] Only the notes E and F are used here, not chord triads, but in order for there to be a harmonic progression from E-F, I have situated the E-F interval in the key of A minor if the interval is indeed tonal. There is nothing uncommon or generally unsettling about the move from the 5th to the flat-sixth chord, but in this instance in the film it does not sound congruent, and E-F is a half-step (or a minor-second) containing dissonance, as it does not resolve.[234] Whether tonal or not, this is an unnatural sound to begin the film. This unnatural sound pierces through the ambient sound of the orcas and their voices retreat into complete silence. However, the sound of rippling water is still present. This musical dissonance exists within the ocean, where the orcas *were*. From here, we get the first title cards on the black image. Once the orcas have retreated, the progression from E to F repeats, with the E accented a little louder, increasing that emphasis on dissonance. It is a very unsettling sound, as the spectator does not know what it is trying to say, nor can they grasp any sense of melody. It may be said that this is a pure musical code because all there is to experience is the musical dissonance, but the fact that the spectator is forced to look at a blank image makes their experience even more unsettling.

---

[233] The movement from E-F is a semitone and could be atonal, but the interval can also be tonally situated in the key of A minor through the 5th to flat-sixth.

[234] Musical dissonance, opposite to consonance, is generally an interval or chord that requires resolving as it contains a nontonal/nonharmonic note in a tonal series. For example, the tritone does not belong in a harmonic series. Dissonance on a beat is unsettling to the ear, but dissonance off-beat usually go unnoticed as they are used as non-chord tones that shift the melodic line. For example, playing a G# in the key of C is dissonant as it does not belong to the scale of C, but it can be used to progress to a consonant, like Am in a stepping motion.

The blank image supported by title cards then quickly transitions to an image of underwater where you can see the seabed (see Figures 34-35). The first sense of physical space is established, and it is not a pleasant image. The water is murky, the colours from seaweed and sand are desaturated, and the camera moves forward unharnessed and mobile, suggesting a point of view. It is no ordinary point of view because the spectator knows that humans would not move within this space in this manner, increasing the mystery and unease of the gaze. As this image occurs, that dissonant E-F that remains is no longer isolated, it is now in pulsating rhythm of 4 crotchet beats per bar, with the use of quavers (half the value which doubles the speed), making it repeat in a loop four times per bar; the dissonance moves (see Figures 36-37). The rhythm created in E-F adds a D every 6th and 7th beat to create further dissonance. There is now a concrete musical cue flooding the image, and the unsettling, murky visual of the moving camera now moves to its pulsating dissonance that began with the singular E-F progression. Dissonance of aural and visual stimuli has established audio-visual congruence. This short rhythmic pattern is called an ostinato, and from the moment the spectator experiences this opening image, the suggested point of view shot is interlocked with this ostinato.[235] On creating this ostinato, especially for the unsettling and ambiguous opening, John Williams said: "I fiddled around with the idea of creating something that was very, very brainless... All instinct... meaning something [that] could be very repetitious, very visceral, and grab you in your gut, not in your brain."[236]

---

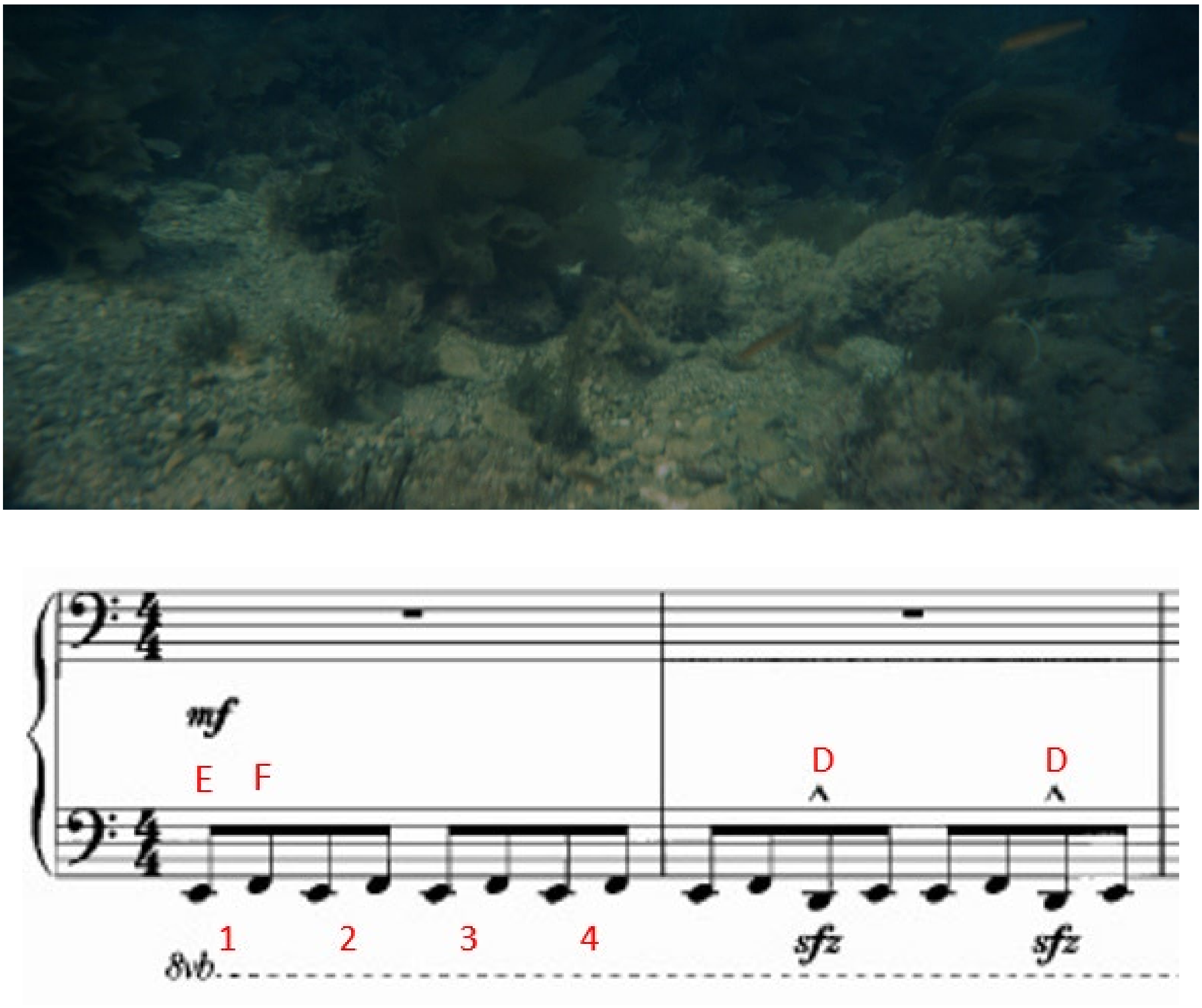
[235] For a survey of the 'historical remarks' on the function of the two-note ostinato in film music, see, Peter Moormann, "Composing with Types and Flexible Modules: John Williams' Two-Note Ostinato for Jaws and its Use in FIlm-Music History," *Journal of Film Music* 5, no. 1-2 (2012).

A
ZANUCK/BROWN
PRODUCTION

Figures 34-35. The transition from the blank title card image to an underwater shot.

---

[236] Emilio Audissino, "Jaws: Williams's Neoclassicism Floats Up to the Surface," in *John Williams's Film Music: Jaws, Star Wars, Raiders of the Lost Ark, and the Return of the Classical Hollywood Music Style*, Wisconsin Film Studies (Madison, Wisconsin: University of Wisconsin Press, 2014), 111.



Figures 36-37. Dissonance moves in an underwater point of view shot.

The title card ‘*Jaws*’ now appears centre screen, with the notes of a dominant 7th chord played in melodic progression like an arpeggio by a tuba in a higher register than the E-F in the strings (see Figures 38-39). This is no cheerful dominant 7th either. The notes E-flat – G – D-flat played as an arpeggio and held on the D-flat are independent of the key of A-minor, making it a non-chord tone. The arpeggio form also highlights each interval in the chord, one of which being the G – D-flat tritone interval, another dissonant, unsettling interval sounding like a chord that needs resolving. According to *The Oxford Dictionary of Music*, a ‘tritone’ is either an augmented fourth, or a diminished fifth, which is unresolved and uneasy on the ear.

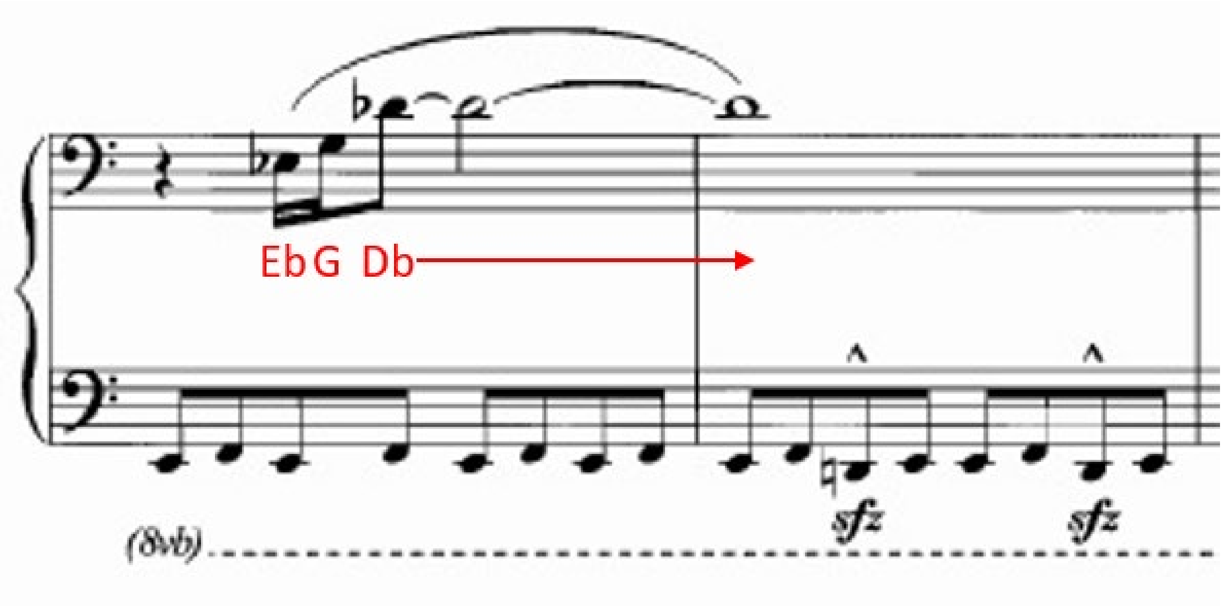
In medieval times, its use was prohibited and was called 'diabolus in musica', “the devil in music.”[237] The tritone in the non-chord dominant 7th makes this mobile non-chord tone movement through the water the point of view of something undomesticated or potentially devilish. The chord repeats a few times with the E-F ostinato remaining in the bass to create a pulsating pedal point. The spectator cannot identify the point of view, and combined with the sound of this musical dissonance, it is a deeply unsettling music-image. As the camera moves through seaweed, suggesting that it is part of the space if it can interact with it, the ostinato and chords increase in volume and texture with the strings adding instruments and now playing in octaves (the same note at different pitches) to add more intensity to the dissonance and growing sense of unease. Even before seeing the shark or the first attack, the spectator suspects that a dangerous physical object is moving through the water. Birger Langkjær, in an article analysing *Jaws* and on the topic of ‘hearing’ a film, argues that “music can be said to add to the visuals.”[238] He references Gorbman in suggesting that the two-note ostinato, the menacing shark theme heard even before the camera in the film reveals the monstrous shark closing in on unsuspecting swimmers, provides the spectator with advance knowledge of the narrative threat.[239] The dissonance created in the music with the unsettling underwater point of view shot provides the spectator with this advanced knowledge and fear, experienced as integrated stimuli constituting the audio-visual image.

---

[237] Michael Kennedy, and Joyce Bourne Kennedy. *Oxford Dictionary of Music.* 6th ed. s.v. "tritone." Oxford: Oxford University Press, 2013.

[238] Birger Langkjær, "Hearing things in music for films: Music, Fiction and Engagement," *SoundEffects* 3, no. 1-2 (2013): 70.

[239] Ibid.



Figures 38-39. A tuba plays a dominant 7th chord progression as the main title card cues.

The death of Chrissy

The opening music-image in *Jaws* is fully established within the narrative space-time with the killing of Chrissy. Up to this point the spectator has experienced an image governed by an immediate narrative function.[240] Now, a narrative-image is introduced as Chrissy goes swimming at a beach party. The audio-visual cue begins with an underwater image, this time looking directly above at the silhouette of Chrissy's body swimming from right to left, illuminated by the dim glow of the moon (see Figure 40). Since the camera is not moving, there is no urgency to the distant sound of violin in the bass softly playing the E-F interval in an irregular rhythmic pattern (E-F, then E-F-E-F, then E-F, etc), like a nervous heartbeat. Accompanying this are the sounds of harp and vibraphone, playing parallel ascending arpeggios, with the vibraphone playing shorter notes, indicating a faster scurry. From my own analysis of this image and its semiotic systems, it is apparent that the sharper sound of the harp represents Chrissy's swim in the moonlight, while the vibraphone, with its damper-sounding timbre, is completely soaked in water.[241] Neither instrument is natural sounding; they are commonly assigned to fantastical, abstract music 'looking for images', as K.J. Donnelly puts it. Nineteenth century composers of the late romantic period wanted their music to summon images, like Claude Debussy's *La Mer* and his *Images* suite.[242] But in this instance it creates an image of someone or something else swimming, suggesting another point of view.

---

[240] In the title sequence for *Jaws*, no sense of story on a larger scale has been established yet, and no characters have been introduced. The initial music-image is constructed in conjunction with the film's larger narrative if the film were to play for its full duration, indeed, but the sequence alone does not amount to anything narratively substantial.

[241] I would like to stress that this specific analysis is an original analysis and experience of this music-image, meaning the interpretation of this instance, but other instances in the film (and other films explored as examples) are not closed off analyses. I expect other analyses to differ based on subjective experiences and music-images. The music-image involves an open model of interpretation and film experience.

[242] K. J. Donnelly, "Soundtracks without Films," in *The Spectre of Sound: Music in Film and Television* (London: British Film Institute, 2005), 151.

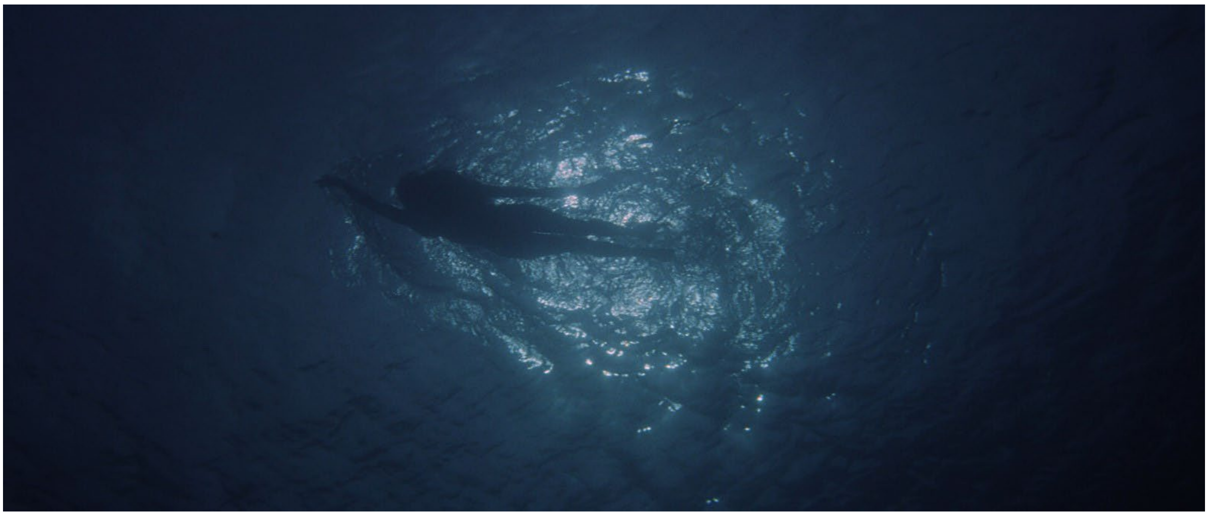


Figure 40. A point of view of Chrissy from the depths of the ocean.

After a quick cut to the sea surface showing Chrissy's cheerful face in close-up, the image cuts back to that underwater vantage point, as the camera moves slowly towards her. The distant, pulsating E-F in the strings has transformed into that same ostinato from the opening credits, but with a slower tempo matching the slower movement of the camera towards Chrissy. There is also the sound of a xylophone quickly descending in pitch through an arpeggio, with violins holding the E-F together in its minor-second chord.[243] The harp and vibraphone arpeggios have quickly phased out as the camera moves towards Chrissy with the significantly dissonant E-F pulse. The movement of the point of view creates tension once again, but this time it has a target. The ostinato now has a narrative function, to signal the impending terror of an attack. The image cuts to Chrissy in close-up above water again, but this time she is tugged from underneath. As Emilio Audissino puts it, the sound of a violent sforzando horn rips "with a kind of 'rrrruhah'!" invoking the shark's sudden and ferocious bite, but also conveying Chrissy’s sudden shock and pain.[244] The next 30 seconds of Chrissy’s violent death are complete audio-visual chaos. There is no tonality to the music in the strings, with the occasional shreds of a xylophone and chaotic poundings of timpani. These cues are paced out, occurring only when Chrissy is being tugged onscreen and ending only when she submerges completely. Ending the audio-visual cue at this point ensures that the image is associated with this underwater threat, from which the cue began. Giorgio Biancorosso agrees that this is when the ambiguity of the underwater point of view has been resolved.[245]

---

[243] According to the OED, an 'arpeggio' is "the employment of the notes of a chord in rapid succession instead of simultaneously; a chord thus played or sung." See, OED Online, "arpeggio, n.," Oxford University Press, https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/10945?rskey=9zE17s&result=1&isAdvanced=false.

The theme for this underwater image has become Pavlovian now that the killer has attacked its first victim, in the sense that the spectator now expects violent action and threat when hearing the cue. Noel Carroll provides a good example of the diegetic barrier in this sequence:

> “When the heroine is splashing about with abandon as, unbeknownst to her, a killer shark is zooming in for the kill, we feel concern for her. But that is not what she is feeling. She’s feeling delighted. That is, very often we have different and, in fact, more information about what is going on in a fiction than do the protagonists, and consequently, what we feel is very different from what the character may be thought to feel.”[246]

Carroll correctly and indirectly (as Carroll is not mentioning the music) observes that there is a diegetic barrier between characters and spectators. The spectator knows things the characters do not due to the music-image. Chrissy did not hear the threat approaching her from underwater, but the spectator did. This space-time is specifically for the spectator and not the inhabitants of the space. A limited interpretation assigns this diegetic dichotomy only with the music when analysing film music; a richer, further interpretation acknowledges the diegetic barrier within all film parameters, as inhabitants of the film's diegesis do not *see* what the spectator sees either. If the music is considered part of the image, perhaps the ongoing debate about the definition of diegesis would be less binary. For instance, in Audissino's analysis of Chrissy's death, he never attempts to reconcile the audio and visual stimulus he draws upon, and instead keeps them separate. Since diegesis is always concerned in what defines 'space-time', and narrative space-time for that matter, the phenomenological import of the 'image', that is, what is being fully experienced, is entangled in a crossing of diegetic frames and visual-aural materials. The audience experiences the shark's two-note motif as part of the image, whether it is sourced off or on-screen. What matters is that it is there to be experienced; the audio-visual image is always synchronised, always part of the whole.

---

[244] Audissino, "Jaws: Williams's Neoclassicism Floats Up to the Surface," 113.
[245] Biancorosso, "The Shark in the Music," 306.
[246] Carroll, *Philosophy as Art: A Contemporary Introduction*, 90.

The death of Alex Kintner

The next music-image cues the next time the camera is underwater, this time for the scene of Alex Kintner's death. Within the narrative space, from the attack on Chrissy until this point in the film, Mayor Vaughan deems Chrissy's death a boating accident, convincing Chief Martin Brody to keep the beaches open. This narrative point determines that the threat is still inherently nondiegetic, as only the spectator is aware of the underwater presence. The underwater point of view is given the autonomy to exploit the diegetic space and attack Alex Kintner unsuspected. The cue is prompted by mystery and anticipation. The shot preceding the attack is of a piece of wood floating on the surface of the water, not fetched by a beachgoer's dog that has been previously shown on the beach. Many thoughts enter the spectator's mind: "What happened to the dog? Did the dog meet Chrissy's fate?" Montage supplies the answer by cutting to the now familiar underwater image looking towards the surface.

Unlike the sequence of Chrissy's death beginning with a stationary camera, this sequence begins with movement which cues the E-F ostinato once again. The number of string instruments playing the ostinato has increased, as the spectator already fears what to expect with the similar music-image in Chrissy's death. The mystery of the gaze was resolved with Chrissy's death, allowing the texture of the image to increase into a more ferocious point of view. The camera scans the surface above dotted with bodies treading water, but finally lands on an inflatable raft that Alex Kintner was seen paddling into the water earlier. Now with greater urgency, the point of view rushes towards the raft. Upon making this vertical climb towards Alex, the devilish tritone within the E-flat dominant 7th arpeggio that is heard with the '*Jaws*' title card is used here played on a harp. The arpeggio created in the dominant 7th chord is repeated a few times, sounding like a realisation, as if the point of view *sees* Alex paddling.

As the underwater camera moves towards Alex and the ostinato reaches the surface for the first time in the film, there is a cut to a wide shot from the beach, the vantage point of the diegetic world and Police Chief Martin Brody, with the ostinato still playing. The music-image containing the E-F ostinato is no longer nondiegetic, no longer hidden by the film world, as a man on the beach points off-screen to the water saying, "Did you see that?" The nondiegetic image is no longer lurking in underwater mystery, no longer an impossible gaze; the event of the image only perceived by the spectator has been exposed to bystanders on the beach and other swimmers. Horns accompany the strings playing the ostinato in a lower, grittier register, as the spectator sees the disturbing close-up image of Alex struggling to survive the underwater gaze, with the dark red colour of his blood flooding the image. This image only lasts a couple of seconds until cutting to Brody looking past the camera at the scene unfolding off camera. In a call and response action, the cut to Brody has distorted the spectator's gaze on Brody, and subsequently Brody's gaze at the water, as Brody has now identified this impossible underwater gaze that the spectator has only been aware of up until this point in the film. In a direct quotation of Alfred Hitchcock's point of view push-pull camera action in *Vertigo* (Hitchcock, 1958), when Scottie has a feeling of vertigo while climbing, the camera tracks into Brody's face while zooming out simultaneously.[247] It provides an optical illusion crumbling the world in around him (see Figures 41-42). This shot is given many names in Film Studies, but I will call it the 'Trombone' shot as it mimics the movements of a trombone to both distort pitch and distort the world around Brody.[248] The trombone movement in the image is echoed in the music, as violins perform a glissando descending and then ascending in pitch, as if moving in and out. When the diegetic eyes of a main character have seen the impossible happen, and the underwater image reaches the diegetic surface and kills a diegetic inhabitant, the music-image shifts gear. The music-image has begun to emit chaos in the diegetic narrative space after the shark is noticed by characters for the first time. All the instrument groups - strings, brass, woodwind, and percussion - play separate atonal melodies, making the scurry of bodies fleeing the water in a string of wide shots occur in a blender of chaotic rhythm, before dying away when all bodies have left the space of the water. The chaos of the image is created due to the diegetic space above water being caught off-guard by the underwater threat.

---

[247] Warren Buckland, "“Duel With A Shark”: Jaws (1975)," in *Directed by Steven Spielberg: Poetics of the Contemporary Hollywood Blockbuster* (New York; London: Continuum, 2006), 99.

[248] I am not the first to call this shot the 'trombone shot', as this is both referenced by Todd McGowen in, Todd McGowen, "Steven Spielberg's Search for the Father," in *The Real Gaze: Film Theory after Lacan* (New York: State University of New York Press, 2007), 139., and by Nigel Morris in, Nigel Morris, "*Jaws*: Searching the Depths," in *The Cinema of Steven Spielberg: Empire of Light* (London; New York: Wallflower Press, 2007), 61., to name a couple of examples.

Figures 41-42. The trombone shot.

From this point in the film, there are no more underwater music-images like the ones depicting the deaths of Chrissy and Alex. The next time there is a similar visual image underwater that may prompt this music-image, that of the fake shark gaze on the 4th of July, music is absent. The shark's music-image, the impossible nondiegetic gaze (the film's title sequence, Chrissy's death, and Alex Kitner's death), has now become diegetic and so will no longer be experienced due to its exposure to inhabitants of the diegesis who are not its victims. The death of Alex Kintner is also the scene where Chrissy's (and now Alex's) death is linked to a shark, and not dismissed by authorities of Amity Island as a "boating accident," allowing the beaches to be open and the shark to kill Alex. The underwater gaze shrouded in mystery and subsequent terror has been diegetically identified; even though the shark killed Alex, his death killed the impossible gaze. The underwater gaze that is established through the interlocking of aural and visual points of view, as explored in this analysis so far, is now familiar within the diegetic space, and so the E-F ostinato image is only experienced with the shark in the camera's sight, as the shark too has become a diegetic object, acknowledged by other inhabitants within the space. The gaze has shifted from the nondiegetic to the diegetic space-time.

This music-image analysis complicates the existing debate surrounding the term diegesis as the "universe in which the story takes place," requiring a rigid demarcation between diegetic and nondiegetic worlds.[249] The famous E-F ostinato, which would conventionally be classified as a nondiegetic sound, clearly functions within and effects the narrative space. The image contained within the diegetic world is able to manipulate diegetic space-time, and yet is experienced only by the spectator. All other occurrences, that is, of the inhabitants hearing or seeing the source of the music-image, are diegetic.

Face to face with the shark

Two final music-images are central to my analysis of the music-image in *Jaws*, both taking place once the spectator and the gaze of the main characters, Brody, Quint, and Hooper have physically seen the shark. When Brody comes face to face with the shark for the first time, his shock is felt via the soundtrack in a visual-aural movement in what is called in Film Music Studies "mickey-mousing". Mickey-mousing is the tight synchronisation of sound and visual rhythm, most notably in music when it imitates the physical movement on screen, such as “the use of a glissando when a character slides down a rope.”[250] James Wierzbicki notes that mickey-mousing got its name after the Walt Disney film *Steamboat Willie* (Disney and Iwerks, 1928) is recognised as the first film to use a “tightly synchronised soundtrack,” featuring the character Mickey Mouse.[251] However, taking this origin story too literally only refers to *this particular* use of film music as synchronous, *whereas all* film music is synchronous to the very instances in which it occurs. In the shark reveal scene in *Jaws*, Brody’s shock, and in turn the spectator’s shock, is merely the beginning of an audio-visual cue that interlocks with the visual sequence. The sudden accented tuba, the same instrument that introduced the spectator to the shark with its dominant 7th tritone, has now been introduced to Brody. His shock awakening that distorts the image can be seen as the distortion of diegesis, the nondiegetic becoming diegetic. Stunned by what he has seen, Brody slowly walks backwards into the cabin to tell Quint what he has seen. The spectator knows that Brody understands the full horror of the shark for the first time, with the high-pitched shriek of the tuba and horns now descending in pitch through the tremolo. Tremolo is the fast reiteration of notes making a trembling sound and providing a sense of motion. The strings play this way through descending tritone intervals, with an echo of muted horns, its trembling dissonance tensing the image that has stunned Brody.

---

[249] Jeff Smith, "Bridging the Gap: Reconsidering the Border between Diegetic and Nondiegetic Music," *Music and the Moving Image* 2, no. 1 (2009): 2.

The shark has gone back underwater, but the threat has now been realised within the narrative space. When Quint comes on deck to see for himself what Brody saw, the use of tremolo in the strings continues to descend; the spectator suspects that Quint will see the shark too and realise the magnitude of the threat. In a simple reverse shot cut (sharing Quint and Brody’s gaze), the music-image of the shark is now within the diegetic space (see Figures 43-44). The ostinato is now used to show the shark from the characters' perspective, and not the perspective of the shark. The threat that once only the spectator knew is now known by the main characters within the diegesis of the film. The camera sits behind the shoulders of Quint and Brody, with the dorsal and tail fin of the shark coming towards them and the spectator. Hooper turns around and sees the shark offscreen too, and so now the ostinato builds in texture and returns to its ominous tritone with the tuba and muted trumpets playing the ostinato in reverse from F-E, as the shark is enclosing the diegetic space. The ostinato builds in volume as the shark draws closer to the boat and then evolves into harp glissandos and ascending trumpets announcing the shark's monstrous size as it swims past from a bird's eye view (see Figure 45). Trumpets enter the image as the shark swims by in a short phrase ending on a cadence, signalling the shark's awe-inspiring reveal to the diegetic space.[252] Hooper and Quint describe the shark's might with, "That's a 20-footer" and "25, three tons of him," respectively. This music-image rejects the dissonance of the ostinato to reflect on how extraordinary this creature is. It is now man versus beast. The spectator, rather than experiencing one or the other within the music-image, now experiences only these two opposing forces inhabiting the diegetic space.

---

[250] Jacobs, *Film Rhythm after Sound: Technology, Music, and Performance*, 96.
[251] James Wierzbicki, *Film Music: A History* (New York; London: Routledge, 2008), 104.

---

[252] Although a cadence is heard for the first time in the film here, I note later on that the first cadence related to the shark's ostinato is heard towards to the end of the film. A cadence in music is located at the close of a musical passage, typically the final two notes, ending on the tonic (the key). A perfect cadence is when the penultimate chord is the dominant 'V' and resolves on the tonic 'I', resulting in complete harmonic and melodic closure. See, OED Online, "cadence, n.," Oxford University Press, https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/25947?rskey=kISb2f&result=1.

Figures 43-44. The shark’s music-image enters the diegetic space.

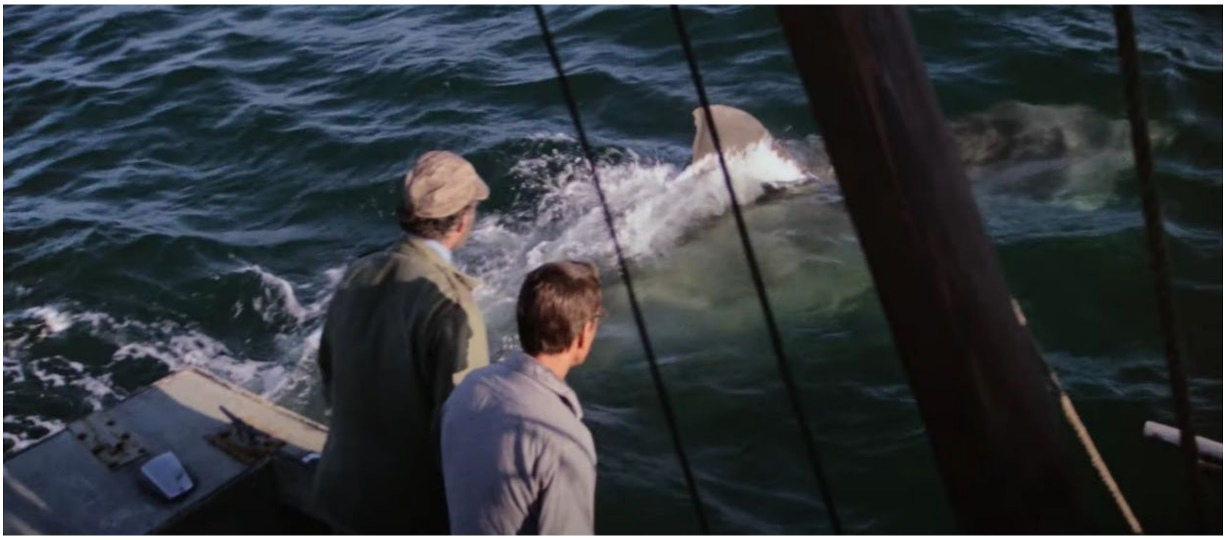


Figure 45. The shark's monstrous size is seen by the characters and the spectator for the first time.

In the audio-visual cue following Hooper spotting the shark returning ("Oh boys, I think he's come back for his noon feeding!"), the E-F ostinato, now firmly established as the diegetic presence of the shark, has become a pedal-point, as the shark returns one last time before the three main characters are split up. A pedal-point is a note or series of notes typically sustained in the bass while a harmony is played in another register. This is the basis of a fugue, where two melodies are in counterpoint, or as in this scene in *Jaws*, in combat. The image begins with a wide shot of Quint looking at the shark coming towards the boat and all the spectator hears is the ostinato, the threat, approaching (see Figure 46). The threat of this image is countered with that opposing melody in a higher register. As soon as Quint yells "Hook me up another barrel!" as the spectator sees Hooper jump onto the bow, the spectator hears the notes C, D and E-flat played in its own little ostinato in an attempt to rival the shark. These notes are close to each other on the musical scale, not leaving the image much room to move around freely. In terms of pitch, they register very closely to the shark ostinato. There is tension in the fight back against the shark, they match each other closely, however the shark's E-F underneath is the marker for a pulsating beat. With this, the temporal rate of images in the montage increases. There is an audio-visual intensity to the human comeback. With the visual image, since the camera is on board the boat, the rate of cutting is dictated by the upper melody. This montage heralds a new attack against the shark, who continues to signify through the dissonant beat beneath the water. The defence from the men provides a glimmer of hope, but the outcome is uncertain.

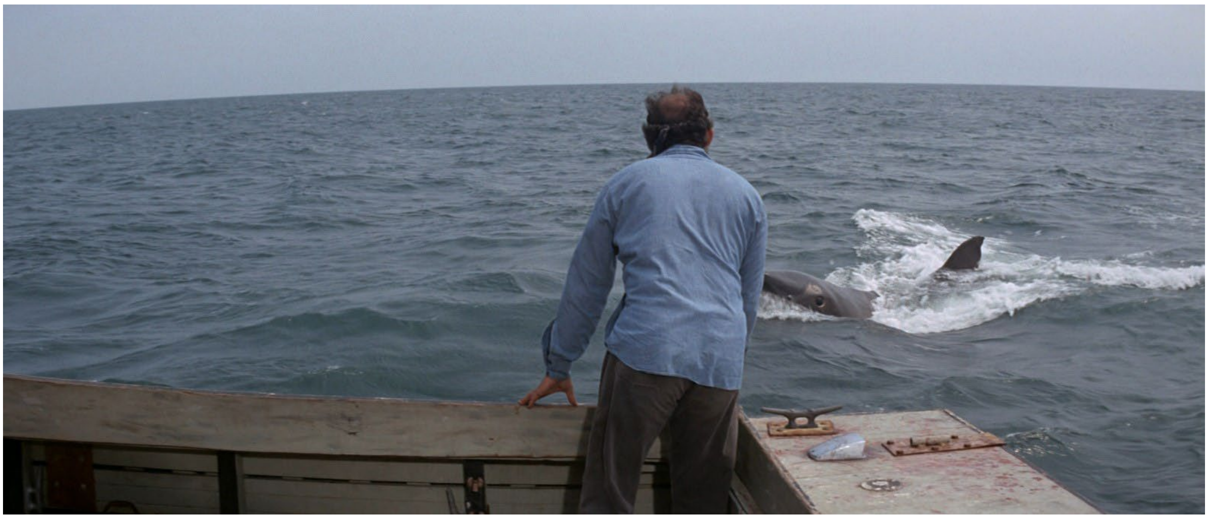


Figure 46. The shark approaches the boat with the ostinato.

When Quint shoots the shark with the harpoon and the barrel attached trails the shark, timpani (drums) modulate the key and disintegrate the existing fugue of ostinatos. While the timpani play, Brody looks out the window with a concerned expression as the barrel zooms over his face in the window. This modulation is a physical pounding matching Brody's concerned face: will this barrel keep afloat? The camera cuts to a wide shot of the boat chasing the barrel (the shark), with a very cheerful string melody playing. Authoritative brass instruments, romantic woodwinds, and a light-hearted beat from percussion can all be heard, sounding like an Erich Korngold pirate melody familiar from films such as *The Sea Hawk* (Curtiz, 1940) and *Captain Blood* (Curtiz, 1935); it is bright, optimistic, full of colour and triumph.[253] As the shot shows the barrel in the foreground and the boat behind, the music-image suggests the threat has reversed: the shark is threatened by the boat now progressing through the water to the bright pirate melody (see Figure 47). That tension in the 'Great Shark Chase' just earlier, which is held together by a fugue of closely bound ostinatos, one of which is the shark's, with the frantic visual cutting of preparations to shoot the shark with a harpoon, has been released. The spectator is prompted to believe that the three men have the upper hand. What is interesting is that even though the shark has entered the diegetic world around the boat in the second half of the film, it no longer dominates the space. The nondiegetic space had control over the diegetic space because the inhabitants could not see or hear the threat. Since the spectator is the only one to experience the nondiegetic music-image, the diegetic world is under the authority of nondiegetic space-time. Now that the nondiegetic threat of the shark, the impossible image, is now possible through its diegetic reveal, it must compete for dominance with the other inhabitants of the space. As soon as the barrels (and the shark) submerge, it becomes a threat again, and this is heard when the pirate melody modulates back to the minor key in a low register played by ominous brass and woodwinds. The shark has retreated down into the depths.

---

[253] Charlie Brigden specifically references the tone colours of Erich Korngold's scores for *The Sea Hawk* and *Captain Blood*, in Charlie Brigden, "Revisiting John Williams' Score for Jaws, 45 Years Later," RogerEbert.com, https://www.rogerebert.com/features/revisiting-john-williams-score-for-jaws-45-years-later. In an interview in 2012, John Williams makes note that the music when the boat begins chasing the shark "…becomes very Korngoldian." See, Jon Burlingame, "The music of Jaws: an interview with John Williams," Limelight, https://limelightmagazine.com.au/features/the-music-of-jaws-an-interview-with-john-williams/.

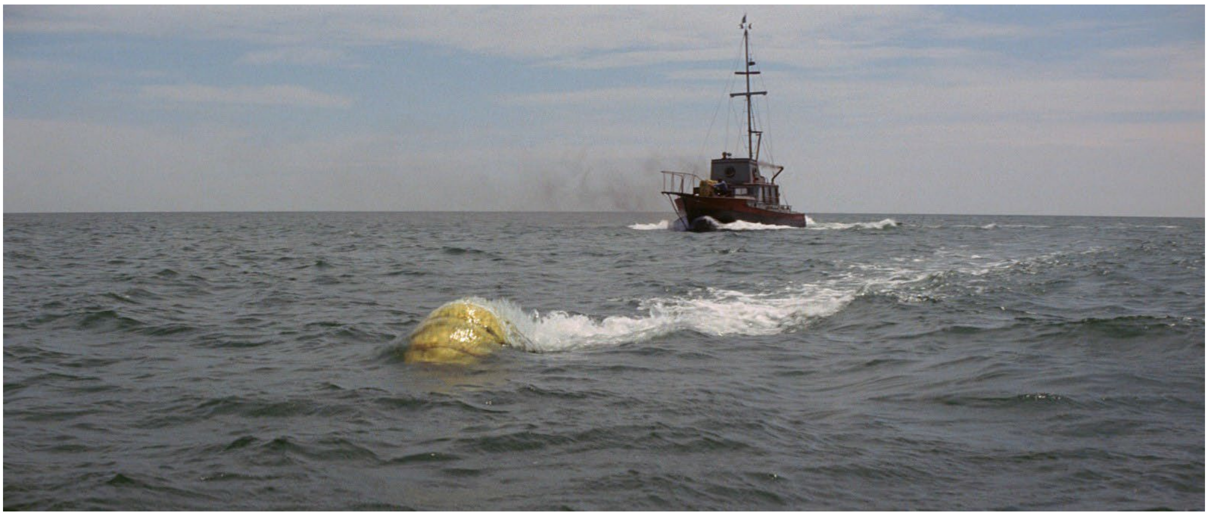
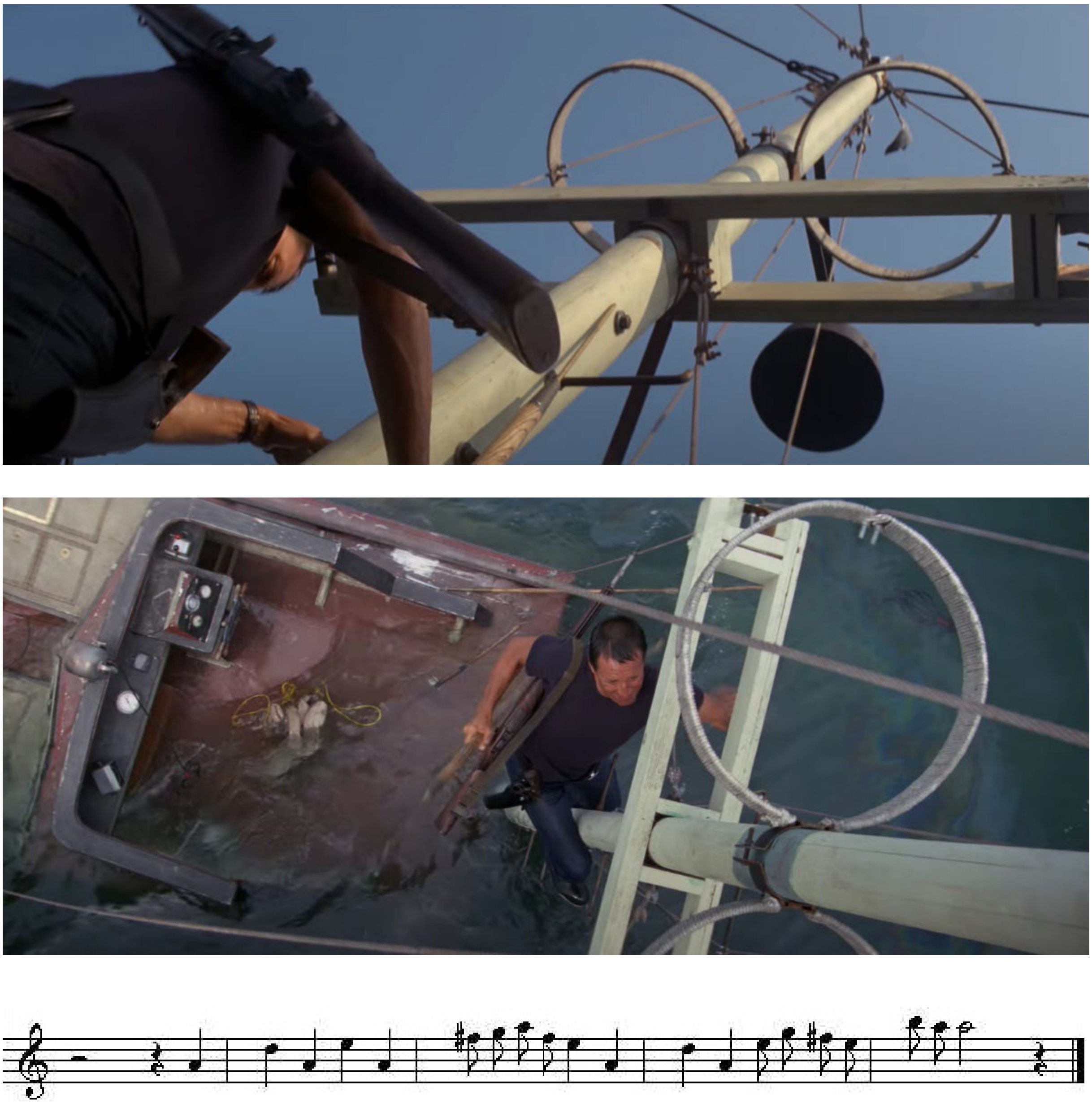


Figure 47. The shark is threatened by the boat in this reverse angle pirate melody.

The death of the shark

In the final scene, Quint is dead, Hooper is hidden underwater, and Brody alone remains active for a final showdown with the shark. The audio-visual cue begins once again with the wide shot image of the shark coming towards the sinking boat from a reverse shot following a shot of Brody scanning the horizon for the shark. Instead of the ostinato, only the E in the E-F is played as a pedal point beat. The shark is lurking in the shot but there is uncertainty as to its threat in these final moments. This uncertainty reveals a very unsettling string melody played over the top of the E in the same key to mask a revenge. As Brody equips a rifle and a harpoon that he has not held before, he looks nervous as he climbs the boat mast. The F returns to the E-F, tipping the scales back to the shark, because of Brody's nervousness. The shot cuts to underneath Brody in a tight angle as he climbs the mast, making him look massive in scale. There is hope in this shot from his towering figure; a nondiegetic theme which sounds familiar to the sea shanty Quint sang, "Spanish Ladies", now envelopes Brody in the image; I will call this 'Brody's theme'. The ascending melody in the major key played by horns, then again with a flute, matches his ascent of the mast but also the growing authority and confidence that Brody now has over the shark. But he still has an uphill battle as the next shot shows, with a contrasting high angle with the same music (see Figures 48-50). A few shots from different angles follow, each getting wider to accentuate his isolation in the water in what will surely be his final task one way or the other.



Figures 48-50. Brody's Theme with contrasting high and low angle shots of Brody.

The melody for Brody's theme dims into an aimless harp arpeggio when there is a cut to a random shot of water moving. The shark is within the space, but its location is not known. The shark suddenly rises out of the water in a close-up of its face, and the ostinato briefly recues. A series of shots are chaotically sequenced together as Brody stabs the shark. One largely textured, loud, low register minor-second interval is played in succession by the entire string family to mimic his stabbing, similar to Stravinsky's 'Glorification of the Chosen One' from his *The Rite of Spring* and the shrieking violins from The Shower Scene in *Psycho* (Hitchcock, 1960).[254] Instead of the high shrieks from the shower scene, the low stabs resonate with the shark's presence but also the strong, steady hero stabbing the villain instead of the villain surprising the heroine. As Brody stops stabbing the shark, it submerges once again and the music mimicking the stabbing also ends.

The low E note from the ostinato returns as Brody repositions in a wide shot, ready to shoot the scuba tank that he lodged in the shark's mouth (see Figure 51). A reprise of the music that accompanied the barrel montage I mentioned earlier enters the space, but this time with strings instead of brass. It is now the same instrument as the pulsating E for the shark, as Brody and the shark are now locking horns. The E-F ostinato returns when the spectator sees the shark underwater drawing near to camera, in one final showdown of man versus beast (see Figures 52-53). The opposing melody builds in texture and volume as the rate of images increases, similar to when Quint was preparing to shoot the shark but this time the opposing ostinato drifts apart from the low E-F, so far so that the melody has nowhere left to go but disintegrate. Brody keeps missing and missing with every shot he takes, all from different camera angles to heighten the extreme tension. The music halts completely as Brody says in close-up “Smile ya son of a…” before he shoots the tank and the shark is finally killed, along with its ostinato. When Brody realises that he has killed the shark and the nondiegetic threat that entered the diegetic world has been destroyed, horns have modulated to A-major by using an F# instead of an F as the image prepares the first ever cadence relating to the ostinato within this modulation and new key.[255] This is the ultimate satisfaction, the dissonant music-image has been resolved. That E which began the dissonant minor-second, which caused terror and dissonance by oscillating with the note F, has been raised a semitone to F#, modulating the key from A-minor to A-major. This key modulation brings the image in line with Brody’s Theme (see the vital F# in Brody’s Theme indicating the A-major in Figure 50) and subsequently, Brody. The modulation to A-major, the key of Brody’s Theme, paired with Brody’s screams of joy and the image of the shark nowhere to be seen after the explosion, signifies a return to safety from tension; a victory for Brody over the shark (see Figures 54-55).

---

[254] Emilio Audissino refers to the violent strings in *Psycho* while discussing the shark attacks in *Jaws*. See, Emilio Audissino, "John Williams and Contemporary Film Music," in *Contemporary Film Music: Investigating Cinema Narratives and Composition*, ed. Lindsay Colemen and Joakim Tillman (London, UK: Palgrave Macmillan, 2017), 226.

---

[255] I note earlier that the first cadence is heard in the diegetic reveal of the shark through trumpets. This is the first time that dissonance in the ostinato has resolved through a cadence. See, Online, "cadence, n."., for a definition of musical ‘cadence’.

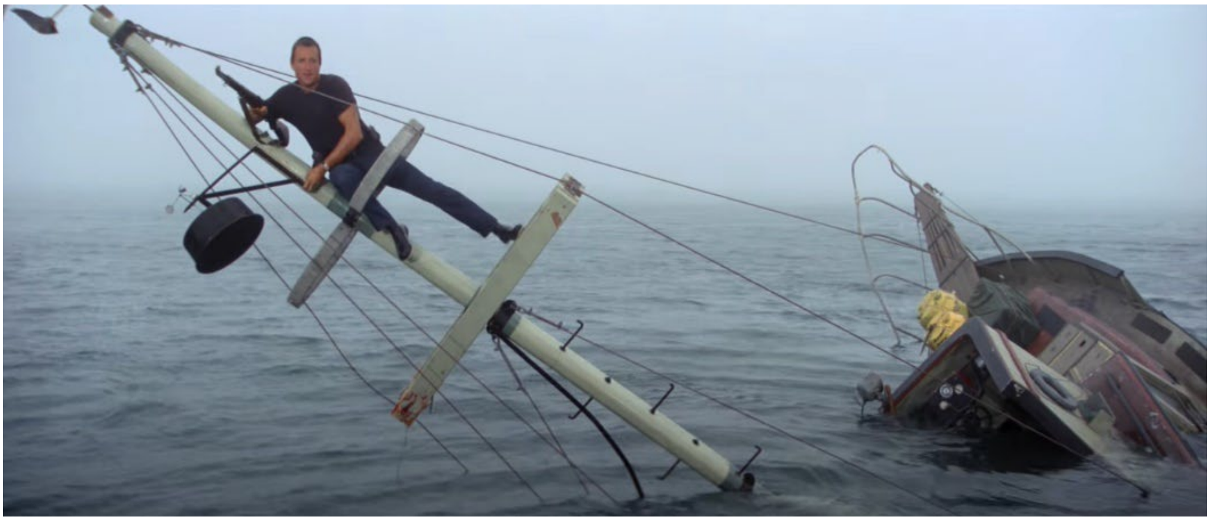
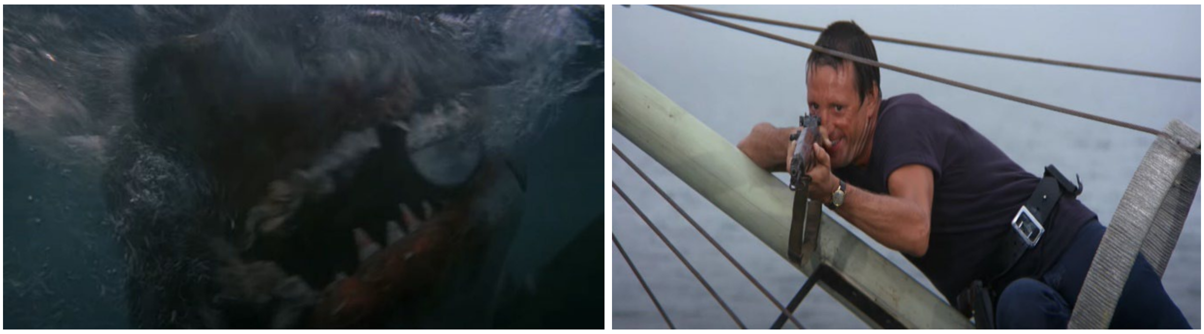


Figure 51. The low E returns with Brody in a wide shot.



Figures 52-53. The ostinato returns in a montage of man versus beast.

Figures 54-55. The ostinato dissipates with the joy of Brody and the destruction of the shark.

While the string section is in the process of modulating to the key of A major, the spectator is shown the carcass of the dead shark descending into the depths of the ocean, with descending arpeggios played on piano to interlock with the descent and the blood trickling (see Figure 56). The piano arpeggios use the vital C# and F# to indicate the modulation to A-major. The A-major piano arpeggios are harmonious instead of dissonant, which is a first for an image of the shark, because it no longer lives. As the strings and piano arpeggio resolve on the tonic of A-major, the camera cuts to Brody's face. The diegetic threat to Brody has been resolved. As Brody and Hooper paddle back to shore, this A major is now firmly in place, matched with the absence of dissonant minor-seconds and impossible underwater threats.

Figure 56. The descending shark carcass with descending piano arpeggios.

Conclusion

Before the shark's visual reveal, the shark has complete control of the spatiotemporal universe of the film, but once it rises to the water's surface and is seen by the main characters it enters the diegetic space where it locked horns with other inhabitants of diegesis. For the film to end happily, the devilish dissonance carried over from the nondiegetic space must also be resolved, and the film finds a way to do this with a music-image that is fully realised diegetically.

In this analysis of *Jaws*, I have attempted to demonstrate the way in which film music does not ‘underscore’ the film image, but rather interlocks in an audio-visual super structure. Within this music-image analysis, the shark transitions (I deliberately use a musicological term) from an ‘impossible’ nondiegetic image that threatens the narrative space and characters to a fully present entity expressed as an audio-visual diegetic monster. It is within this diegetic shift that the shark moves from a point of centrality and command within the diegesis to its death and aural-visual diminution.

Chapter 5: 'When' is the Music-Image?

In my analysis of *Jaws* in Chapter 4, it is clear that music is not present for the entirety of the film's duration. There are specific points in the film with film music and others without film music, indicating that there is a decision and recognition about *when* music-images present within the film. If film music is part of the construction of the audio-visual image, whenever there is music in a film, there is a music-image to experience. In this chapter, I explore the question of whether the nondiegetic music-image is truly 'unheard' (as Gorbman proposes), before analysing 'when' the music-image occurs in a given film, using *The Hours* (Daldry, 2002) as my example. As I suggested in my analysis of *Jaws*, the same film music can repeat at several points in a film, which occurs in the E-F string ostinato linked with point of view (POV) underwater shots or visual images of the attacking shark; musical and visual semiotic cues in repetition signal and formulate narrative cues. The repeated audio-visual semiotic phrases link thematically, which provides a framework for my discussion in the concluding section of this chapter: "When music-images repeat" and "When is the first music-image" in a given film.

Is nondiegetic film music ever really 'unheard'?

This research project began with encountering Roger Scruton's deep sense that something is 'missing' from film music when that music is taken out of a film and played in a concert-hall setting.[256] The same can be said for the music-less film that remains. Film music is not placed in a film not to be heard; it is there for a reason. For instance, for the first hour of *Jaws* the terrifying underwater threat of the shark, which had not been physically revealed to that point, would not have been as emphatic without the film music that attached to the visual and narrational fields of the film. The title of Claudia Gorbman's ground-breaking study on film music contains the word 'unheard', but what does Gorbman mean by this? In the introduction to an anthology of essays on film music in *Music and Cinema*, David Neumeyer, Caryl Flinn and James Buhler assert that "Music is 'unheard' because its narrative functions are folded into, or readily overcome by, others: "on the soundtrack, dialogue routinely takes precedence over music."[257] Similarly, Jeff Smith suggests that for Gorbman and others, film music "somehow escapes or eludes the film spectator's perceptual awareness during the viewing experience."[258] However, this 'unawareness' (which in itself is an undefined state of attention to the image) is only plausible if we adopt the traditional model that film music is subordinate to image and narrative form.[259] One of the main reasons for this, Gorbman argues, is that music is (generally) the only entity of film which is nondiegetic, while all other entities are part of the diegetic discourse: camera, mise-en-scene, editing and sound.[260] It is important to note that Classical Hollywood film music aims to be seamless in a film, not drawing attention to itself. This seamlessness is achieved because it meshes with the narrative in one harmonious rhythm. Gorbman further begs the question of why film music is even in a film if it lowers the threshold of belief and has the aim to be unheard in this way. Sean Cubitt quotes Christian Metz in his book *The Cinema Effect* that: "In some sense all cinema is a special effect," and if this were to be taken literally, with its proposed construction of production commodities, then music is part of that construction as much as the diegetic discourse that Gorbman asserts.[261] The stitching together of shots into a sequence is as much cinematic

---

[256] Scruton, *Music as an Art*, 179.

---

[257] David Neumeyer, James Buhler, and Caryl Flinn, "Introduction," in *Music and Cinema*, ed. James Buhler, Caryl Flinn, and David Neumeyer (Hannover; London: Wesleyan University Press, 2000), 1.

[258] Jeff Smith, "Unheard Melodies? A Critique of Psychoanalytic Theories of Film Music," in *Post-Theory : Reconstructing Film Studies*, ed. David Bordwell and Noel Carroll (Madison, Wisconsin: University of Wisconsin Press, 1996), 230.

[259] Gorbman, *Unheard Melodies: Narrative Film Music*, 2.

[260] Ibid., 3.

[261] Sean Cubitt, *The Cinema Effect* (Cambridge; London: The MIT Press, 2004), 1.

‘reality’ as the E-F ostinato played by the string section of an orchestra in *Jaws*, both of which do not exist without the construction of cinema.

Nonetheless, I have to admit experiencing *Spider-Man: Far from Home* (Jon Watts, 2019) in a cinema all to myself with excellent surround sound, I shared aspects of Gorbman’s thinking:

> “We opt to focus attention on the narrative and visual realities on the screen before us...we forsake contemplating that abstract arrangement and rearrangement of sound which is music, because it is nonrepresentational and nonnarrative and does not inhabit the perceptual foreground of the narrative film."[262]

During the screening of *Spider-Man: Far from Home*, I had to make myself consciously aware of the music in the sequence in order to experience its ‘pure’ musical codes. However, this is not to say that the music is not heard, that it is not there, or that it should not be heard.[263] Martin Marks, analysing the film music of *The Maltese Falcon* (Huston, 1940), a Classical Hollywood film, suggests that most people do not hear film music most of the time while experiencing a film, unless it is an essential part of the narrative or draws attention to itself. This is the case for Marks even though composer Adolph Deutsch composed music for nearly half of *The Maltese Falcon*’s runtime.[264] The ‘effectiveness’ of film music Gorbman alludes to is based on the conscious acknowledgement that it engages with the diegetic space-time. This is not to declare that all nondiegetic film music is in the ‘background’ or devoid of diegetic space-time, as it still exists and engages with the film as much as the diegesis itself. There is no question that music is often neglected in studies of film because of the subordinate position it is thought to have in the experience of a film, but a comprehensive analytical engagement with film form remains limited if film music is excluded from the image analysis.

---

[262] Gorbman, *Unheard Melodies: Narrative Film Music*, 12.

[263] As well as Gorbman, Annabel Cohen offers valuable insight into the topic of film music’s “inaudibility”. See, Annabel Cohen, "Film Music: Perspectives from Cognitive Psychology," in *Music and Cinema*, ed. James Buhler, Caryl Flinn, and David Neumeyer (Hanover, New Hampshire: University Press of New England, 2000).

[264] Martin Marks, "Music, Drama, Warner Brothers: The Cases of Casablanca and The Maltese Falcon," ibid. (Hannover; London: Wesleyan University Press), 161.

As the opening statement of this chapter suggests, a film stripped of its film music would have something missing; it would be a completely different film and experience for the spectator. Whether the film music is consciously experienced as lurking in the background or blaring in the foreground, it is and remains part of the image. That said, the music-image is not experienced by analysing the score on a microscopic level. Rather, I argue that an appropriate analysis of the music-image consists of the macro-musicological themes within the music-image form. These macro themes are associated with the film itself. Aural and visual materials are always interlocked when they both exist; the music-image is therefore always synchronous.

To draw out this principle that the music-image is always ‘heard’ in the sense that it exists when cued, I will analyse a random survey of film sequences that traditionally present as ‘unheard’. The first is from Ingmar Bergman’s *Wild Strawberries* (Bergman, 1957), a film not known for its use of music. There are far fewer audio-visual cues in the film than in *Jaws*, with not even half as many (roughly twenty). All cues in the film are related to Isak Borg’s dreams or memories. For example, in Isak’s first dream sequence with its abstract, Salvador Dalí-like imagery, he sees his own dead body grab him by the arm. As he approaches the hand raised out of the coffin, the hand is in the foreground to the right while Isak is positioned to the left towards the back of the shot. The hand is of interest to Isak, so he approaches, and of interest to the spectator as the hand is held geometrically to the right of the image while Isak is held to the left (see Figures 57-58). Through this action, a soft, tense string chord is played, registering Isak’s interest as a *concern*. The shot cuts to a close-up of the hand beginning to move from Isak’s point of view and an added string chord brushes on top of the existing chord, creating a very unpleasant, tense sound accompanied by an atonal harp. This visual image is horrific in composition, while the music reanimates the dead hand. The sequence begins cutting more frequently with this atonal tension while the hand grabs Isak's arm; Isak is not sure what this means but it is horrific, distorting the image. The head of the dead body rises out of the coffin, and it is Isak. The music intensifies, creating unease at this revelation through a close-up shot reverse-shot montage of both Isak faces. Isak is looking at his own dead body, staring death in the eye. A great and unexpected "gong" is heard as he cannot shake the hand off his arm. The camera closes in on both faces, with the lens almost making contact, while the music screeches. Staring at death can mean many different things, but this music-image suggests that Isak is frightened by death. The music comes to a complete halt when Isak awakens from the dream. Even though this music-image does not draw attention to itself, it creates a separate meaning of the sequence from a purely visual interpretation. The music affected the visual form insofar as the intimacy between Isak and his dead body is deeply unsettling.

Figures 57-58. Isak and the arm in shot-reverse-shot, keeping both of interest.

Where is this music coming from if it is in Isak's dream, a space created by his psyche? Can Isak hear this horror that the spectator experiences? It is unclear, but it can be deduced that the music is not within the space of the dream as Isak did not noticeably react to a music source within the space. The music is clearly to help the spectator understand how Isak feels when he is presented with a nightmare of death staring him in the face, drawing him closer to his coffin. It is Isak's dream, but it is nondiegetic in the sense that the music cannot be sourced within the space. In the film, Isak is about to travel to Lund in Sweden to attend a ceremony celebrating 50 years of his medical service, but this music-image makes the spectator suspect that Isak feels that his time is up. In the scenes that follow, this theme of questioning existence continues with acceptance of the life that he has lived and the world he lives in. If the spectator did not 'hear' this music, or 'experience' this music-image, Isak's fear of wasting life established through death calling him to the grave would not be so clearly 'horrific'.

It is possible that this signal of narrative 'concern' is not contained in a sequence, nor the film narrative that progresses from that moment, but lingers beyond the boundaries of the duration of the film through the music-image.[265] In the final scene of John Carpenter's *The Thing* (Carpenter, 1982), after MacReady throws dynamite at the alien creature hoping that it is finally destroyed, a question remains: is the alien dead? MacReady sits down panting, exhausted from fighting the alien creature. Childs re-joins him and they sit and chat about what to do next, with the rest of their crew dead and stuck in Antarctica. MacReady says to stay put for a while and hypothetically suggests they kill themselves to be certain the job is done as the alien can live inside a host undetected. In the not so far nondiegetic distance, sustained notes in the high register are played on violins creating an eerie atmosphere for their conversation. The mood created is ominous as either MacReady or Childs could be an alien, creating a disturbing uncertainty in the spectator's mind. MacReady hands Childs an opened beer to share and as soon as the bottle touches his mouth, the main theme cues. This main theme seems an obvious way to end the film and fade to the end credits, but the main theme is deliberately musically unresolved and non-symphonic with its unusual synthesising pulse. I purposefully muted this final scene to see if it resolved visually, and the visual cue of resting and sharing a beer after the final battle with the alien creature seems uplifting; they can finally relax with no more alien threat. I played the sequence again with music and the result is unsettling instead. On a typical 88-key piano, the lowest F is played here on a synthesiser in a pulsating rhythm, playing two quavers on the first beat, sounding like a heartbeat (see Figure 59). It is so low on the keyboard that it has no real earthly voice. The choice of using a synthesiser creates an 'unworldly', 'futuristic' sound like Vangelis's score for *Blade Runner* (Ridley Scott, 1982) which released the same year. However, in the final scene of *The Thing* the image certainly does not look futuristic as the clashing colours of red, yellow and white from the fire and snow in the destroyed base camp look like an ancient barren wasteland, with MacReady and Childs icy and bloody (see Figures 60-62). What, then, is Ennio Morricone's music doing to the image? This synth heartbeat does not seem to complement their reunion, but points instead to some 'thing' that may be there, not visible in the visual frame, but present within the larger narrative frame of the film. The spectator knows from seeing the film that the alien creature is parasitic and assimilates living bodies to hide, taking over control until its cover is blown, revealing itself. When Childs takes a sip from the opened beer bottle, the music cue of the low F suggests that either Childs or MacReady could be infected with the thing, as by sharing of the beer bottle, the thing believes it can hide. I mention earlier that this low F is the main theme because it has been used before in the film, namely at the beginning in which the Norwegian explorers attempt to shoot the possibly infected Huskey. This connection, but also the unworldly 'heartbeat' motif, triggers a trauma that ultimately ends the film in a state of ambiguity rather than resolution. Heather Addison points out that before this, when Childs appears out of thin air and MacReady asks him, “Where were you, Childs?” Childs does not give a satisfactory answer. Addison claims that

---

[265] In a similar effect of in which the music-image extrudes beyond the narrative and durational field of the film.

> “Childs has been out of Mac’s sight and may be a Thing; the same is true of Mac, both from Childs’s and the viewers’ perspectives (previously, some crew members were revealed as Things without their transformations having been presented on screen). Neither these characters nor the viewers know whether the [alien creature] has been eliminated…”[266]

However, this is not clear on its own, with the ending suggesting resolution after the final confrontation with the alien creature. The music moulds tension with uncertainty, with the notion of the physical interaction of drinking and an alien pulse creating a visceral fear for the spectator, suitably ending a film deemed a ‘body horror’.

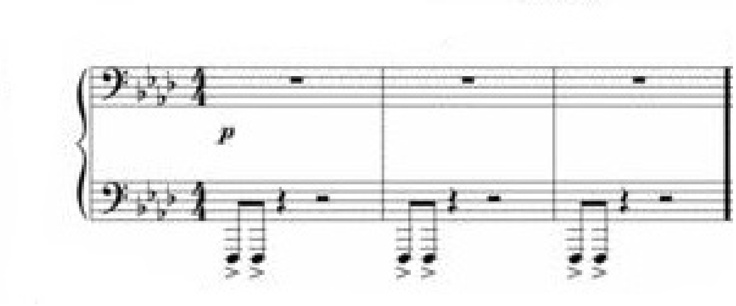


Figure 59. An extremely low F note played in a heartbeat rhythm.

---

[266] Heather Addison, "Cinema's Darkest Vision: Looking into the Void in John Carpenter's The Thing (1982)," *Journal of Popular Film and Television* 41, no. 3 (2013): 164.

Figures 60-62. The colour palette of a baron wasteland in *The Thing*.

While the series of low F notes creates this pulse, there are light shape-shifting chords played also by the synthesiser (see Figure 63).[267] This shapeshifting of chords echoes the shape-shifting persona of the alien creature, hosting on other creatures with a living ‘pulse’. Anne Billson concludes that the ending of the film “is not a commercial, crowd-pleasing finale, but the daring pays off. By denying the viewer[s] a sense of closure, Carpenter has ensured that the story will live on.”[268] The music-image has extended the ‘concern’ of the alien threat in *The Thing* beyond the final frame, into a realm of alien uncertainty.

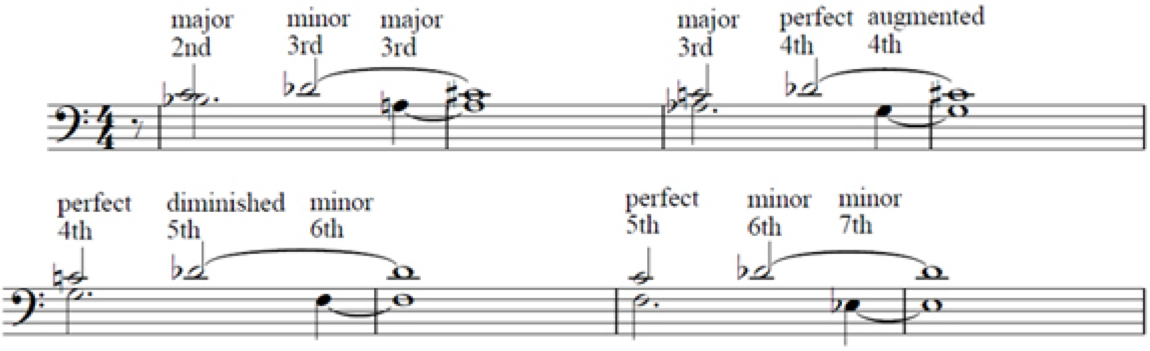

major 2nd
minor 3rd
major 3rd
major 3rd
perfect 4th
augmented 4th
perfect 4th
diminished 5th
minor 6th
perfect 5th
minor 6th
minor 7th

Figure 63. The shapeshifting chord progression echoing the shapeshifting persona of the alien.

Since film music is ‘heard’ whether it is ‘invisible’ within the visual image or not (think of the ‘seamlessness’ in Classical Hollywood film music), its lifespan depends on the audio-visual cue. A ‘cue’ in music is a “direction to enable a singer or player to come in at the right time after a long rest”; in cinematography, it is a “signal for action to begin or end.”[269] With the audio and visual image interlocked, the music-image exists only when it is cued, which rests its ontology on 'when' it takes place. However, film music generally does not occur for the entirety of a film; instead, it is interlocked at certain points. For instance, Sergei Eisenstein and Sergei Prokofiev matched the music with the vertical visual plane in the Battle on the Ice sequence in *Alexander Nevsky* to discuss a potential audio-visual 'vertical montage'.[270] Why is it that this music did not appear at another point in the film, but specifically for these visual vertices? If some film music composers can only place music at certain points, for example, during scene transitions, is that a fault of the composer who cannot find a way to create music-images or is it the case that the music cannot fit otherwise? To be clear, film music is not the end goal of a film production; it is commonly used sparingly and, in narrative film, for narrative effect. David Ireland claims that "[film music] plays a key role in providing information that influences how a perceiver might construct and interpret meaning in the sequence, and may impact their emotional response to the filmic events."[271] In his discussion of Kathryn Kalinak analysing the playing of the cheery song 'Stuck in the Middle with You' in a violent torture sequence in *Reservoir Dogs* (Tarantino, 1992), Ireland indirectly argues that this audio-visual interlocking is on purpose, even if the choice to play this song is obtrusive.[272] The music-image is cued in parts of the film when visual form alone cannot generate narrative effect. So, while music *can* play for the entirety of a film, its specific and limited phrasing heightens its use. Since there is cueing, the choice is for purely cinematic reasons. It is for the spectator to decipher and create meaning of their own. To analyse the music-image in any film, it is knowing 'when' the music-image occurs that contributes to its meaning. Recalling from Chapter 2, the spectator is part of the music-image automaton. The music-image is autonomous in the sense that the spectator enters into a relationship to the image, thereby forging the ontological form. When the music-image is cued, meaning is generated.

---

[267] Figure 63 retrieved from Hugo Fagandini, "Film Score Friday #3: "The Thing" (1982)," Son et Lumière, https://sonetlumiereblog.wordpress.com/2016/10/14/film-score-friday-the-thing-1982/.

[268] Addison, "Cinema's Darkest Vision: Looking into the Void in John Carpenter's The Thing (1982)," 164.

[269] OED Online, "cue, n.2," Oxford University Press, https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/45574?rskey=72N3xh&result=2.

---

[270] See Chapter 1 for an extended analysis of vertical montage in *Alexander Nevsky*.
[271] Ireland, *Identifying and Interpreting Incongruent Film Music*, 1.
[272] Ibid., 1-2.

‘When’ is the music-image in *The Hours* (Daldry, 2002)?

I have arbitrarily chosen *The Hours* (Daldry, 2002) with the only guiding criteria for analysis being the complex narrative structure of the film, which tracks the lives of three central characters across three different times and places. This narrational and situational complexity and richness results in an intriguing series of music-image cues.

The Hours follows the characters of Virginia Woolf in 1923, Laura Brown in 1951 and Clarissa Vaughan in 2001, all of whom intersect directly or figuratively through Virginia’s novel, *Mrs Dalloway*. After an introductory scene showing Virginia’s suicide in 1941, the film flashes back to Virginia in 1923 writing *Mrs Dalloway* while dealing with depression and nervous breakdown. Laura in 1953 is unsatisfied with life while reading Virginia’s book and identifies with the title character, Clarissa. Clarissa in 2001, as her name suggests, is clearly inspired by the character Clarissa in *Mrs Dalloway* as she prepares a party for her ex-lover Richard, who is dying of AIDS (and who is later discovered to be the son of Laura). She is obsessed with flowers, as is her namesake, but is also struggling with life. The film concludes by returning to the sequence at which it begins, Virginia’s suicide.

There are 18 audio-visual cues in the film that span over 56 minutes 57 seconds of screen time, leaving 57 minutes 43 seconds without music. This proportion indicates that music is chosen at specific points in the film. Out of the 18 cues, there are 10 different pieces of music scored, three of which repeat at different points in the film. Philip Glass, the composer who scored the film, is known for his minimalist music on and offscreen. Timothy Johnson suggests that,

> "…principal features of the minimalist technique include the five characteristics of the minimalist style...: an incessant formal structure, an even rhythmic texture, and bright tone, a simple harmonic palette, a lack of extended melodic line, and repetitive rhythmic patterns. Thus, the minimalist technique often produces long, harmonically static passages, characterized by consonance and built from repeated patterns and pulses."[273]

There are multiple music-images in the film that repeat melodic and rhythmic patterns, and passages that are harmonically static, bearing no endpoint and deliberately avoiding cadences. There is also minimal instrumentation, with only piano, strings, harp, and glockenspiel used. Roger Hillman and Deborah Crisp argue that there is a "nonlinearity, even circularity, of the music [that] matches the film's dramatic constellation, which moves like the crystals of a kaleidoscope…consistent with minimalism."[274] The music's minimalist nonlinearity matches the complex nonlinear narrative and editing structure of the film, avoiding textual reference, and rather basing itself purely on cinematic principles of camera and montage. From this, the three main characters, Virginia, Laura and Clarissa who occupy different spaces and times do not always retain their own scenes, with the subject shifting mid-scene. This creates a time continuum with all three characters sharing the same day in different narrative time periods through shared emotion and feelings.

The 18 music-images are psychologically motivated and avoid cueing on textual references. This makes the phrases relatively short in duration, but they repeat later in the film. The music-images do not distinguish time or place, or the three characters, instead interconnecting separate narrative threads into one whole. Cues occur mid scene rather than at points of dramatic change from one narrative section to another.[275] At arguably the most dramatic narrative turning point in the film, when Virginia's husband Leonard meets her at the train station after she runs away, there is no music until the final lines in their deliberation on Virginia's mental health. When Virginia states "this is my right" to live the life of the poet, even though she suffers from depression, the music finally cues in a horizontal fashion to ground the drama of the scene, leading nowhere melodically or harmonically, in the fashion of a cardiogram (see Figures 64-66). Virginia and Leonard walk into the crowd of people alighting the train on its arrival in an unobtrusive mid-shot as Virginia tells Leonard: "You cannot find peace by avoiding life, Leonard." The cardiogram music-image, now loud in volume and busy in appearance, gives the horizontal impression of life, or meaning passed on. The image does exactly that by passing the music onto Laura who reconciles with her son, Richard, after she decides not to commit suicide and abandon her son (see Figures 67-68). There are many other similar sequences in which dramatic volume is loud but musical volume relatively muted.

---

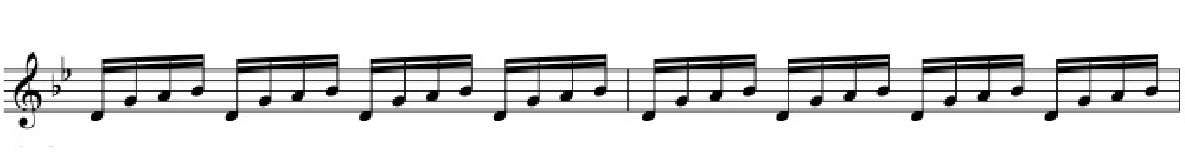
[273] Timothy A. Johnson, "Minimalism: Aesthetic, Style, or Technique?," *The Musical Quarterly* 78, no. 4 (1994): 751.

[274] Roger Hillman, and Deborah Crisp, "Chiming the Hours: A Philip Glass Soundtrack," in *Virginia Woolf and Music*, ed. Adriana Varga (Bloomington; Indianapolis Indiana University Press, 2014), 289-91.

---

[275] Ibid., 294.

This is my right



Figures 64-66. Music leads nowhere melodically as Virginia declares the life of the poet.

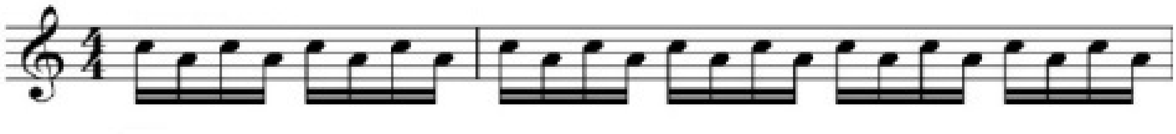
You cannot find peace by avoiding life, Leonard

Figures 67-68. The music-image continues with Laura.

As well as loud dramatic silence, there are many repeated musical cues that interlock with repeated character moods. Gorbman suggests that such repeated moments, musical motifs, since repeated, can carry representational meaning.[276] Johnny Wingstedt, Sture Brandstrom, and Jan Berg assert here that it is not just a repeating musical motif that occurs, but rather the repetition of visual and musical association as well.[277] However, in *The Hours*, the musical cues that do repeat do not necessarily interlock with parallel or complementary semiotic cues from the visual form. Each time a melody repeats, its repetitive narrative nondiegetic input lies in the ambiguity it creates as it interlocks with the visual image when cued. The musical piece that repeats the most has 5 cues out of the 18. The first time it occurs is when Laura's husband Dan leaves her and Richard alone, and Laura plans on making a cake. After Dan leaves for work, and Laura farewells him from the window in a mid-shot from outside the house, there is a cut to a wide angle of Laura looking back at Richard as the music cues in. The music contains a rhythmic pattern that oscillates between A and C in the A minor triad on strings, moving nowhere harmonically, only occasionally switching to the F major VI triad, which is a harmonic progression that is so subtle that it keeps a boundless horizontal rhythmic pulse. The wide image that is seemingly painless with the farewell of Dan on his birthday, has become consumed by the dampness of minor strings (see Figures 69-70). The room now seems soft and fragile, as with Dan gone Laura is left alone with Richard, a fragile relationship. Rather than categorise, the music creates an ambient mood. As soon as Laura declares that she and Richard will make a cake for Dan, the A minor chord turns into a one octave arpeggio form, making the aimless rhythmic pulse faster but also busier. The scene cuts to Clarissa in 2001, with the arpeggio remaining but rising in volume. The bustling pace of the dull arpeggio on piano increases the pace of editing in the film, creating a montage linked by this dampened rhythmic pulse. Clarissa is organising her evening party for Richard on the phone in a series of quick cuts with the spaces saturated in the colour blue. The music-image emits a cold, depressed mood lurking beneath the surface of spontaneity in the montage. The music dies as Clarissa enters the flower shop and speaks a few lines to the florist, with the infiltration of floral colours filling the image, drowning out the dampness. The joyous passion Clarissa has for life relieves the feeling of existential quandary.

---

[276] Gorbman, *Unheard Melodies: Narrative Film Music*, 26.
[277] Wingstedt, "Narrative Music, Visuals and Meaning in Film," 198.



Figures 69-70. Laura and Richard's fragile relationship face to face.

The same piece of music returns roughly 10 minutes later, after Clarissa speaks to Richard in his apartment about the party she is hosting for him. The A minor pulse of A and C returns when Richard holds Clarissa on her way out and they kiss briefly. This seems like a moment of romance in the brief close-up, but the blue colour saturation and the dampness of the rhythmic pulse heard earlier creates a narrative sense that grief lurks beneath the surface of her relationship with Richard.

Another 20 minutes go by in the film before this music occurs again, this time with Laura kissing her friend Kitty, after Kitty recounts her marriage struggles. Only this time the music cue *follows* the kiss, not during, when Kitty calls Laura "sweet." Very dimly in a low volume, the A minor string rhythm softly weeps. When Kitty leaves, Laura tosses the cake she and Richard made in the bin, after shouting at Richard, causing him to run to his room.

The following repeated cue comes minutes later, when Virginia and her sister Vanessa are talking about Virginia's mental health and are interrupted by Vanessa's children tending an injured bird. They discuss whether it is time for the bird to die and ask Virginia the meaning of death. Virginia gets teary and the music volume is raised when Virginia is left alone with the dead bird. She stares at the dead bird's eyes, then it cuts to Laura's eyes in a reverse shot through time and place. Laura is depressed and contemplates committing suicide. This is a parallel example of the shared, repeated emotion relating to existence, life and death. The final time this music is cued is when Virginia kisses Vanessa, asking whether she seems "better." Instead of a solid answer to comfort Virginia, the scene cuts to a solitary Clarissa in 2001 as her friend Louis leaves her apartment after a powerfully dramatic scene containing no music. These mid-scene cuts and audio-visual cues of repeated existential suffering are not based on any textual or visual reference, but the thoughts, feelings and emotions the three main characters share.

The longest cue in the film, that occurs once, is at the beginning when the spectator is introduced to the three characters, Virginia, Laura, and Clarissa, waking up and preparing for the day ahead. The nondiegetic minimalist rhythmic pulse of narrow, horizontal harmony, like the repeated patterns of the above cues, is motivated by the dread of preparing for the day instead of the motivation that drives existence. It fills the void of their existence and the diegetic space with repeated feelings of life's circularity, like the rhythmic pattern of the music itself. With three characters each controlling a first-person perspective in different times and places, the scenes they seemingly control tend to shift mid-scene, evident in this opening introductory passage. The music, like their lives onscreen, is going nowhere but forward into the day of turmoil. Roger Hillman and Deborah Crisp argue that the "minimalist patterning of melody, rhythm, and harmony seems not to demand closure, therefore lending itself either to fading out midstream or to the superimposition – or subtraction – of another melodic layer. Both tactics are frequently used, and "cadence" in the usual sense is avoided...."[278] This is apparent in all 18 audio-visual cues in the film, including the most linear, the introduction of the three lead characters.

Music-images are heard in the sense that they exist, but they are used and evolve throughout the film in deliberate cues, making the film experience dependent on when they occur. The selection of *The Hours* as a sample analysis for examining 'when' music-images occur may seem less obvious than a film like *Jaws*, but its example of minimalist harmonies interlocking with mentally dampened narrative-images demonstrate that nondiegetic music-images are used not purely for narrative and textual understanding. Cuing the musical track creates a music-image form that the visual field alone cannot emit.

When music-images repeat

The Hours' cues repeat at several points in the film. Jack M. Stein suggests that when the spectator experiences a musical cue that has repeated, a thought is triggered of the previous moment the music was cued, affecting the experience of the current musical cue.[279] This repetitive musical cue is what I call the 'thematic music-image'. The thematic music-image is based on Gorbman's definition of 'theme music', which is "…any music - melody, melody-fragment, or distinctive harmonic progression - heard more than once during the course of a film."[280] Music and visual image are interlocked within any given music-image, supplying an audio-visual association. When the same music repeats, closely resembling another cue, there is still an audio-visual association, but there is now also a memory association with the preceding music-image. The thematic music-image contains the cuing derived through memory association.

---

[278] Hillman, "Chiming the Hours: A Philip Glass Soundtrack," 295.

[279] Jack M. Stein, *Richard Wagner and the Synthesis of the Arts* (Detroit: Wayne State University Press, 1960), 75.

[280] Gorbman, *Unheard Melodies: Narrative Film Music*, 26.

I would argue that the thematic music-image is most likely (of all music-images) to be separated and listened to in isolation due to its repetition in film. Thematic music-images contain music that is most heard, in that the more a musical cue repeats, the more likely it will be as pure musical codes. When film music repeats, it evokes a memory, which Gorbman calls the "memory-motif," after Richard Wagner's "motifs of reminiscence," and this memory lasts beyond the duration of the film.[281] The memory that outlasts the duration of the film is subject to replay. However, this replay now exists in isolation from the film, stripping film music from its part in the music-image. The film music that is replayed in isolation still bears a strong association with the film image due to memory, and film narrative as the spectator will remember not just the music from its repetition, but the *film* image from which it derived. However, once the film music exits the music-image, repeated in isolation, it exists in isolation, and perhaps can be heard without having experienced the music-image or the film. This is problematic if applying merit to film music, as film music that is remembered or heard more than supposedly 'unheard' film music is undoubtedly favoured. When theme music is listened to in isolation, it is based on pure musical codes that never change. However, when the audio-visual makeup of the music-image is completely replicated in its repetition, thematic music-images both repeat and differ in experience and analysis when compared.

---

[281] Ibid., 28.

The thematic music-image is based on the idea of the ‘leitmotif’, popularised by the romantic composer Richard Wagner. Although Wagner never used the term ‘leitmotif’ in his written works, his music, particularly his operatic music of the late 19th century, is built on this idea. A 'leitmotif' is, "in the musical drama of Wagner and his imitators, a theme associated throughout the work with a particular person, situation, or sentiment."[282] On this basis, Adorno and Eisler define the leitmotif as musical trademarks such that persons, emotions, and symbols can be instantly identifiable in films, for audiences to understand what is going on.[283] However, they quickly assert that leitmotifs have no place in film, particularly Hollywood cinema, as cinema “seeks to depict reality” and only “leads to extreme poverty of composition”; in other words, the music of all film is bereft of pure aesthetic value.

To assess the merits of this harsh criticism of the leitmotif used in film, it is necessary to understand why leitmotifs are used in the first place and whether they have a place in film. Wagner was a great influence on film music, particularly Hollywood film music through key film composers such as Erich Korngold, Max Steiner, Miklos Rózsa and John Williams, some of whom migrated from Europe, borrowing and transporting Wagner’s ‘romantic’ style. The Romantic style is:

> "Designating, relating to, or characteristic of a movement or style during the late 18th and 19th centuries in Europe marked by an emphasis on feeling, individuality, and passion rather than classical form and order, and typically preferring grandeur, picturesqueness, or naturalness to finish and proportion."[284]

---

[282] OED Online, "leitmotiv, n.," Oxford University Press, https://www.oed.com/view/Entry/107180?redirectedFrom=leitmotiv#eid.

[283] Eisler, *Composing for the Films*, 2.

[284] OED Online, "romantic, adj., 7," Oxford University Press, https://www.oed.com/view/Entry/167122?rskey=m9T4MX&result=1.

Film composers have introduced this musical idea as film borrows a similar connection between music and image from the Wagnerian opera. This connection in opera existed before it did in film, and so the Wagnerian leitmotif has affected the ontology of film music's foundation.

Wagner references separate vocal pieces in opera as 'unconnected' in the sense that "it was only each single tone-piece that possessed in itself a connected form, derived from purely musical considerations…and imposed upon the poet as a compulsory yoke."[285] Here Wagner suggests that the reference to the same piece of music creates a connection with the action, or in the case of film, with the image. He further outlines that "the connection in these forms consisted in a theme…repeated capriciously on musical grounds."[286] Wagner then acknowledges that the dramatic action, which can be paralleled in film, becomes melodic from its association with musical friction:

> "The principle motives of dramatic action, *become melodic moments*, accurately distinguishable, and completely realising their purport, are fashioned, in their significant repetition, always well-defined – similar to rhyme – into a one and indivisible artistic form, spreading not only over the more limited portions of the drama, but over the entire drama itself, as a uniting and connecting series, in which not only these melodic points appear as mutely explaining each other, and therefore one and indivisible, but, also, the motives of feelings or phenomena incorporated in them, as containing the strongest and the weakest ones of the action, as mutually presupposing each other, and one and indivisible according to the essential attributes of the kind - manifest themselves to the feelings."[287]

---

[285] Richard Wagner, "Opera and Drama," *The Musical World* 34, no. 9 (1856): 132.
[286] Ibid.
[287] Ibid.

Dramatic action only becomes melodic if the two entities are one; one entity, musical idea or action, cannot be realised without the other.

This oneness, or unity, or what Wagner terms an 'alloy' in his book *Opera and Drama,* can "foul the comprehensibility" of the artistic purity of each entity.[288] Gorbman, as I noted in Chapter 1, claims that film music is not designed to be heard in its pure musical codes, and then to be played separately from the film; hence, Wagner's claims about the absence of musical purity in a musical piece attached to a film image. The leitmotif in film, however, operates and exists within the connection between musical and image form. If the repetition of film theme music outside of a film is based on the purity of the separated music as 'pure musical codes', that value rests on its musical codes in separation from the film. A theme repeats in a film based on its relationship to the image, which means that experience of a theme inside the duration of a film is based on theme-image, or in other words, the thematic music-image.

For Wagner to argue that the blending of drama with music "would need apology but could by no means gain acquittal" is therefore devoid of music-image experience.[289] It is easy to see where Adorno and Eisler's dismissal of the film leitmotif comes from, as Wagner himself believed that this blending of music and drama could only come at a cost. And yet this 'oneness' that Wagner repeatedly refers to in his works, referring to music, strengthens the music-image model of audio-visual interlocking. The oneness is a little different to the interlocking of music and image per se, but it parallels the moulding of drama and music. Marcia Citron argues that this problem that Wagner, Adorno and Eisler encounter is that the adjective 'operatic' is capable of many meanings, and these meanings can lead to different interpretations: "operatic can refer to a range of elements, such as genre, structure, expression, style, music, or tone…the possibilities are many."[290] Hans Keller even goes as far as to suggest that the leitmotif in the Wagnerian sense is hard to define in film as there is a discordance between Romanticism and film realism. However, this distinction Keller makes is precisely why the leitmotif is used sparingly in film, as not all films function in the Wagnerian operatic sense.[291]

---

[288] Richard Wagner, *Opera and Drama* (Lincoln: University of Nebraska Press, 1995), 120.
[289] Ibid., 121.

The leitmotif legacy that Wagner left behind has created a new language for film. Film music is interlocked with the film image to create a meaning differentiated from their separate semiotic channels, and if the interlocking repeats to create narrative connections, that repeated connection is the basis of the semiotic ambiguity that lies between music and image, or what Gorbman calls the “memory-motif.”[292]

When is the first music-image?

If film music can repeat and evoke memory, when is the first occurrence of the music-image? To understand where this ‘memory’ originates and its first point of connection, it is important to understand that a film generally begins before the first scene. For many films, the title sequence is when a music-image is first experienced. Title music is the least cinematic of all music-images to analyse, as these images rely on music for narration, which is explicitly the case in film overtures. However, film music is rarely separate from the visual image in these opening sequences in the audio-visual sense, as they are usually interlocked. For the title sequence theme to work, it needs to bear connective narrative tissue with the rest of the film.

---

[290] Marcia J. Citron, "Operatic Style in Coppola's Godfather trilogy," in *When Opera Meets Film* (Cambridge: Cambridge University Press, 2010), 20.
[291] Hans Keller, "Hans Keller: Essays on Film Music," *Film Studies*, no. 5 (2004): 115.
[292] Gorbman, *Unheard Melodies: Narrative Film Music*, 28.

Georg Stanitzek calls the title sequence a ‘complex intermediary zone’ in that it offers a transition before the film begins, but indeed the film has already begun.[293] What Stanitzek means by this is audiences may not give their full attention to the title sequence as they ‘transition’ into the film, but this is not to say that the title sequence is not an important part of the film. The French literary theorist Gerard Genette, although mainly referring to literature, describes the title of a text, which can include a film, the ‘paratext’.[294] The paratext is the title of a book or the title credits of a film that lies outside the boundary of the film, and is separate from the narrative body of the film. The point of declaring it as ‘outside’ of the film’s boundaries is to distinguish any paratextual message it contains.[295] This can be spatial, temporal, substantial, pragmatic, and functional characteristics. Pragmatic title sequences, for example, would concentrate on the individuals who are listed in the opening credits, but as title credits and sequences exist within the enveloped duration of the film (regardless of narrative structure), they contain the semiotic functioning of the film’s contents.[296] It can be said that this is the point in a theatrical production at which the curtain has not opened, but the duration of the production of the performance has begun. In many films, the title sequence does not open the diegetic world, acting as that curtain, whereas many title sequences exist within the diegetic world already in sequence. The title sequence is then the connective tissue of the cinematic narrative body; it is the bridge and acknowledgement of the real-world origin of a film within the story world.[297] In the realm of narrative (fiction) films, the titles become more semiotic than pragmatic as they transition to the world of make-believe rather than a source of intertextuality in non-fiction.[298] If title sequences are paratextual, they offer a guide or a blueprint of the film to come, which potentially includes the main theme or a series of themes that will be explored in the film.[299] The title sequence is almost never an actual 'scene' in the film; it instead attempts to introduce the scenes that will follow.

---

[293] Georg Stanitzek, "Reading the Title Sequence," *Cinema Journal* 4, no. 48 (2009): 44.
[294] Gérard Genette, *Paratexts: Thresholds of Interpretation*, trans. Jane E. Lewin (Cambridge: Cambridge University Press, 1997), 4.
[295] Ibid.
[296] Alessandro Cecchi, "Creative Titles: Audiovisual Experimentation and Self-Reflexivity in Italian Industrial Films of the Economic Miracle and After," *Music, Sound, and the Moving Image* 8, no. 2 (2014): 181.
[297] Phil Powrie, and Guido Heldt, "Introduction: Trailers, Titles, and End Credits," ibid.: 111.
[298] Cecchi, "Creative Titles: Audiovisual Experimentation and Self-Reflexivity in Italian Industrial Films of the Economic Miracle and After," 181.

James Buhler and David Neumeyer suggest that film music in title sequences are the most “predictable” categories for music placement in a film.[300] The spectator has not yet necessarily entered the diegetic world, and so music serves as an introduction to the construction of the diegetic world and a way to frame the feature film that will follow.[301] The point of this predictability is so that the narratological understanding of the drama that follows is firmly rooted in memory, and responds to the anticipation. In the following analysis, I examine four approaches to the composition of the filmic title sequence and the ways in which film can introduce the thematic music-image.

The Overture

The first way to introduce a thematic idea is the film ‘overture’. An overture, in the musical sense, is “an orchestral piece of varying form and dimensions, forming the opening or introduction to an opera, oratorio, or other extended composition, and often containing themes from the body of the work or otherwise indicating the character of it.”[302] Wagner even calls the overture one of the only ways to ‘connect’ multiple themes in a production

---

[299] I say ‘potentially’ as not all title sequence have music, and not all films have a title sequence.
[300] James Buhler, Neumeyer, David, and Rob Deemer, "Music in Main-Title and End-Credit Sequences," in *Hearing the Movies: Music and Sound in Film History* (New York; Oxford: Oxford University Press, 2010), 165.
[301] Ibid., 166.
[302] OED Online, "overture, n., 7a," Oxford University Press, https://www.oed.com/view/Entry/135281?rskey=X3ZfSK&result=1.

before creating the oneness of music to drama.[303] The overture, along with music over end credits, is possibly the least cinematic of music-images, as it relies heavily on pure musical codes (the notes, phrases and melodies) to narrate ideas. Such sequences also film segments in which film music is most explicitly ‘heard’. It is also the most musical form of the ‘intermediary’ zone of paratextual meaning, as Georg Stanitzek says, as it introduces the film almost entirely through musical theme. However, the overture is also almost always presented within a visual canvas, and usually in a still or set of stills, or an animated image. It acts like the front cover of a book that has not been opened yet, or the curtain in musical theatre before it is drawn; yet there is something to see and hear in the interim. The overture supplies an overview of what the spectator will experience behind the curtain or front cover.

Overtures:

> “were central to exhibition practices of the silent-film era, when live concert-style or song performances were normally also part of the program and, in theatres that could afford orchestras, overtures were often played before the main feature…these overtures frequently were taken from 19th-century operas."[304]

Speaking of 19th century operas and the overtures so often borrowed by early films, Steven Vande Moortele states that the typical overture was in sonata form, although the sonata was only one of several possible variants. Vande Moortele therefore defines the overture as a subcategory of a ‘symphonic first movement’.[305] Although film overtures sit comfortably in the sonata form, as the sonata is commonly structured around thematic material and ideas, I will specifically address the role of the overture role in the opening of a film.[306]

---

[303] Wagner, "Opera and Drama," 132.

[304] Buhler, "Music in Main-Title and End-Credit Sequences," 165.

[305] Steven Vande Moortele, "Introduction: Formenlehre, Genre, and the Romantic Turn," in *The Romantic Overture and Musical Form from Rossini to Wagner* (Cambridge: Cambridge University Press, 2017), 5.

The evolution of increasingly grand spectacles and larger screens through the classical Hollywood era called for operatic and epic formal compositions, which in turn prompted the rise of the overture as a form of epic narrativising.[307] The sword and sandal epic, *Ben-Hur* (Wyler, 1959), was filmed in 70mm and ran for a marathon 217 minutes, including a lengthy overture.[308] As Stephen Meyer suggests, "*Ben Hur's* leitmotifs… function as a means of appropriating the magnificence of grand opera."[309] The *Ben Hur* overture both explores the narrative themes of the film and compositionally frames the grandeur of the tale. Ralph Erkelenz argues that *Ben-Hur*'s overture, or any film overture for that matter, did not have to synchronise with any action on screen and so could take liberties with its various tempos.[310] However, Erkelenz overlooks the critical space of the visual screen, that is, the visual image, along and through which the spectator pairs the codes of the musical piece. The visual image intrinsic to the *Ben Hur* overture is a picture of Michelangelo's *Creation of Adam* (1508-1512) blown up to show only Adam's (man) and God's (divinity) hands (see Figure 71). This iconic image of great religious and aesthetic significance is overlayed by 'Overture' in Trajan font, a font type that shares its name with an early Roman Emperor. The interlocking of visual and musical form generates a third layer of semiotic meaning. The melody-driven overture contains five repeating themes in the film: Anno Domini, Judea, Esther, Miriam, and Friendship.[311] Other films from this Hollywood period that contain overtures interlocked with a visual image are *How the West Was Won* (Marshall, Hathaway, Ford, 1962), *Spellbound* (Hitchcock, 1945), and *It's a Mad, Mad, Mad, Mad World* (Kramer, 1963). None of these examples contain a blank screen, each uses a different image. This is to say that each overture is imbued with the compositional form of its visual association.

---

[306] According to the OED, the 'sonata form' is "a structure used in musical composition, which may be analysed as consisting of three sections (exposition, development, and recapitulation) in which the musical possibilities of the thematic material (typically two contrasting themes or subjects) are explored." See, OED Online, "sonata form, n.," Oxford University Press, https://www.oed.com/view/Entry/184563?redirectedFrom=sonata+form#eid1212072090.

[307] Although the overture is part of film history, there are very few examples and the majority belong to American films released in the Hollywood period of 1950-1970. It is no accident that most occurred in this era in Hollywood, as this was a time when major film studios needed strategies to increase audience numbers due to the rising popularity of television. One strategy was to increase the size of the film both narratively and in the making and presentation of film, but most importantly the size of the cinema screen, with the famous introduction of 70mm and Cinerama. The larger screen called for larger narratives and high drama, and with that high drama came an operatic quality which called for an overture.

[308] It was the third longest Hollywood feature film at the time, after *Gone with the Wind* (Fleming, 1939) (238 minutes) and *The Ten Commandments* (DeMille, 1956) (220 minutes). See Roger Hickman, *Miklós Rózsa's Ben-Hur: A Film Score Guide* (Lanham, Mayland; Toronto; Plymoth, UK: The Scarecrow Press, Inc., 2011), 65.

[309] Stephen C. Meyer, ""Leitmotif": On the Application of a Word to Film Music," *Journal of Film Music* 5, no. 1-2 (2012): 107.

[310] Ralph Erkelenz, "Ben-Hur: A Tale of the Score," *Pro Musica Sana* 5, no. 1 (2005): 6.

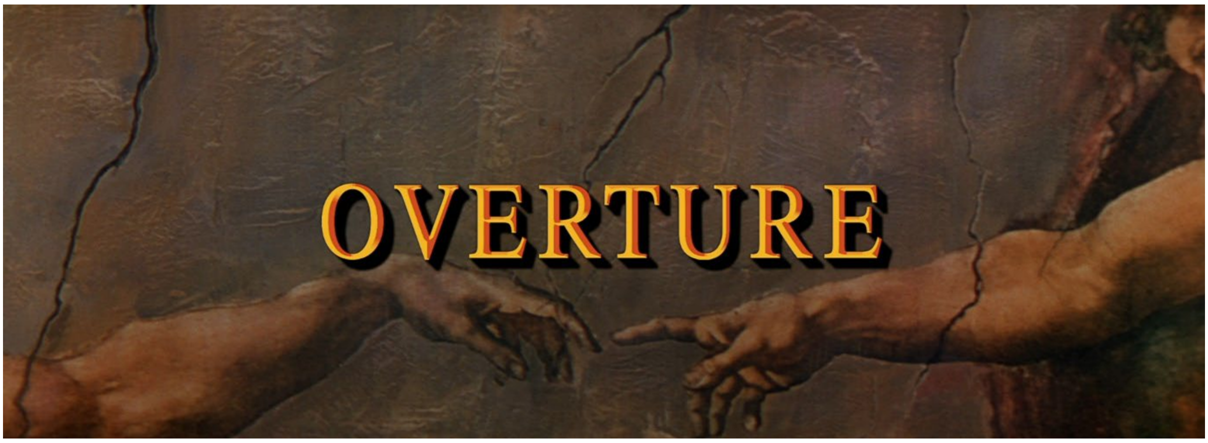


Figure 71. A blown up image of Michelangelo's *Creation of Adam* in *Ben-Hur*'s overture.

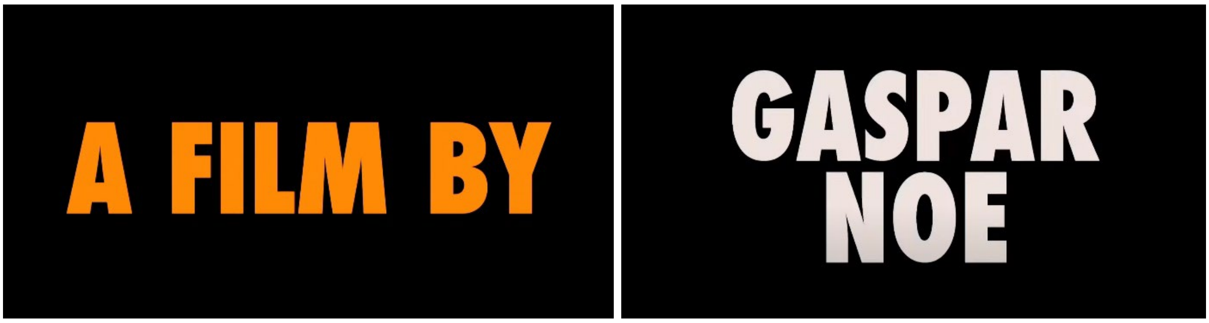
The Title Sequence

After an overture (unless there is no overture), a film's title sequence usually introduces the construction of the diegetic world. As the overture would lay out various musical themes accompanied with an image, the title sequence introduces a film's audio-visual themes, the first synchronisation of musical and visual semiotic systems (though still paratextual within the film narrative). There are four main forms of audio-visual title sequences in film: paratextual title cards, paratextual animation/visualisation, diegetic title sequences, and the delayed title sequence.

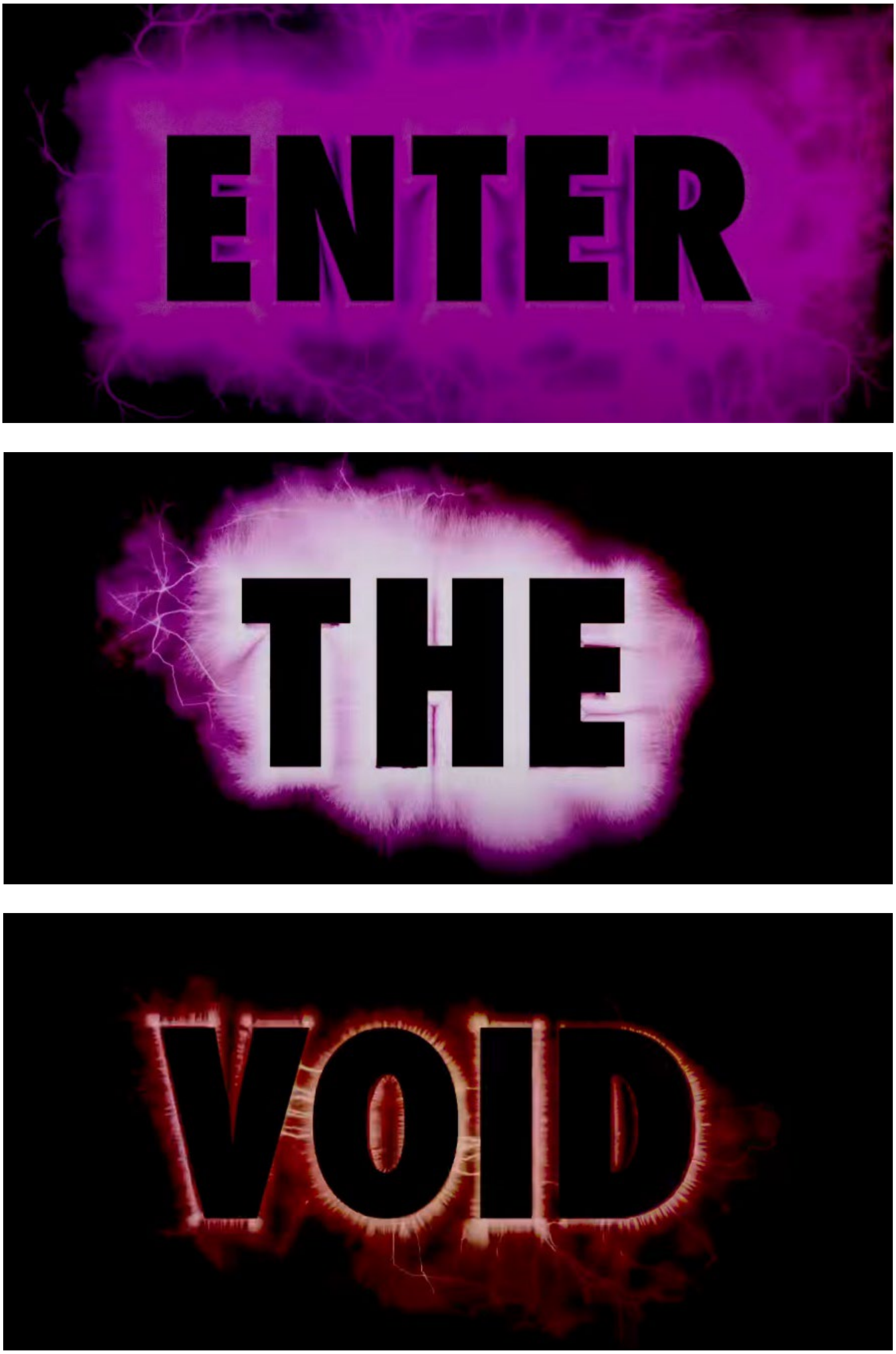
---

[311] See Chapter 5 of Hickman, *Miklós Rózsa's Ben-Hur: A Film Score Guide.*, for a comprehensive analysis of the score.

Paratextual title sequences (title cards and animation/visualisation) contain visual material never seen again in the film (unless there is a reappearance in the end credits) that establish the narrative that will follow. In the title sequence for *Enter the Void* (Noé, 2009), the spectator is assaulted with images of drug-induced epilepsy. The names of the film crew flash in strobe lighting to the sounds of dampened clicking, like a film projector spitting out the names of film crew or, more extremely, a machine gun firing live rounds at the spectator (see Figures 72-73). This imagery transitions to a distorted electronic fanfare that ignites the fiery neon lights of the film's title card (see Figures 74-76). In this metric nondiegetic montage, the title cards flash and cut in a syncopated rhythm to the music. The film here is introduced as a psychedelic pulsation of drug addiction conveyed by fusing the symbols of the underground nightclub drug scene with strobe lights and menacing electronic music. Constructing audio-visual imagery of this degree with the titles of the film anticipates a certain or similar narrative that will follow in the diegesis.



Figures 72-73. The opening of *Enter the Void*'s title sequence fires the names of the film crew in strobe lighting.


ENTER
THE
VOID

Figures 74-76. The title card for *Enter the Void* is ignited in fiery neon lights.

Similarly, in the title sequence of *Vertigo* (Hitchcock, 1958) the unstable oscillating arpeggios of strings and xylophone mould to the dizziness of spiralling vortexes as the title credits flash by (see Figure 77). There are two seventh chords played here in arpeggiating contrary motion. In contrary motion, a note from each chord is isolated and played together (as opposed to the two chords being played simultaneously as chords), moving in opposite directions like a mirror image (see Figures 78-79). On two occasions in this harmonic progression, minor seventh intervals are heard played together (F and E-flat; C and B). The seventh intervals played in isolation from the rest of the chords generates an unsettling sound with the hypnotic visual spirals to almost emulate an audio-visual grammar for the feeling of 'vertigo' to come in the film. Kathryn Kalinak notes that the arpeggios, one in a treble voice, and one in a bass voice, "move in opposite directions or in contrary motion."[312] Every few bars brass instruments blast an unsettling chord which attempts to stabilise the audio-visual vertigo, but there is tension that is too entangled to untie.

---

[312] Kalinak, *Settling the Score: Music and the Classical Hollywood Film*, 5-6.

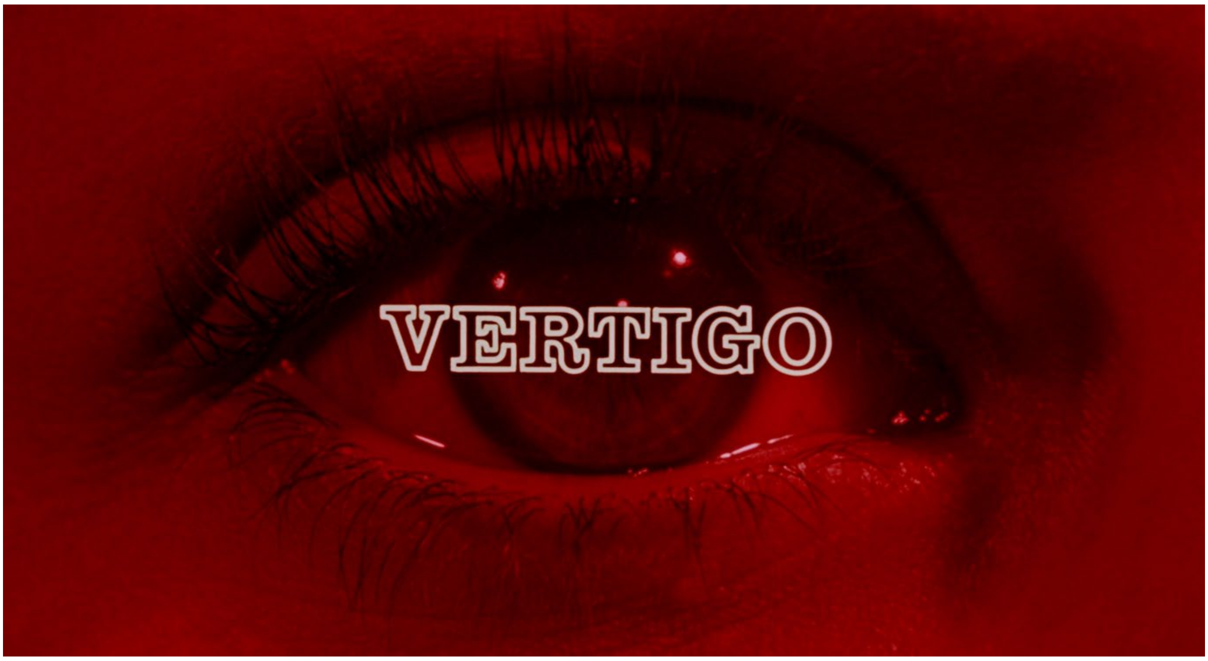
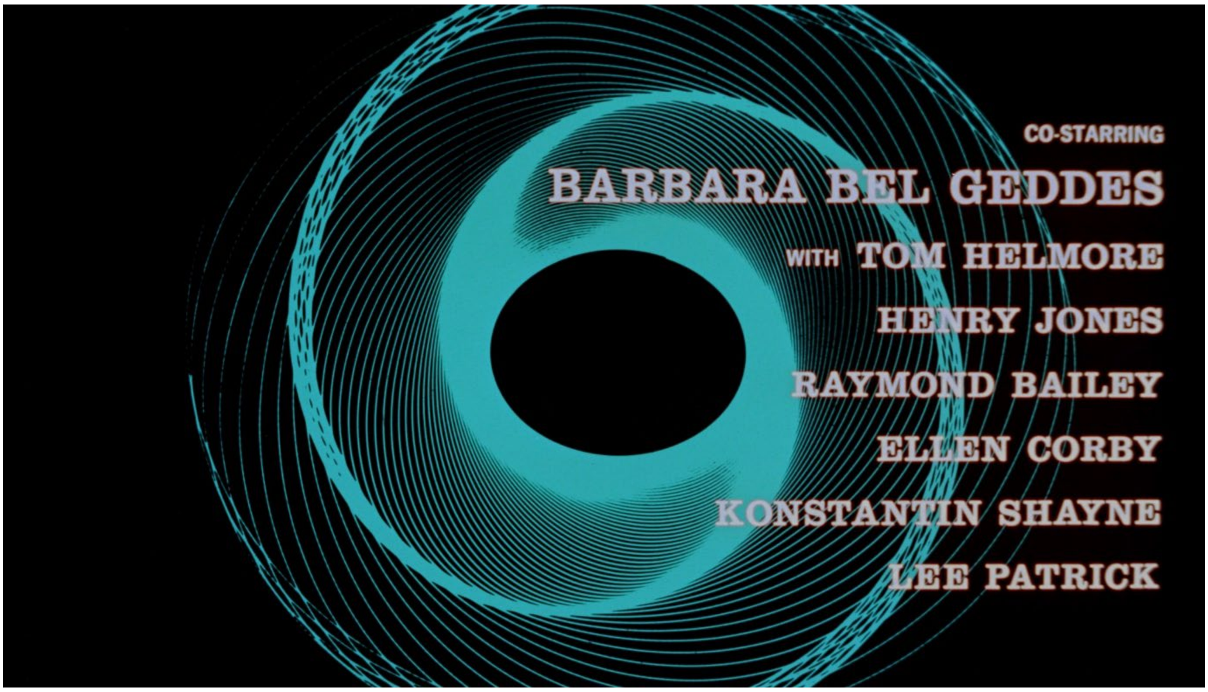
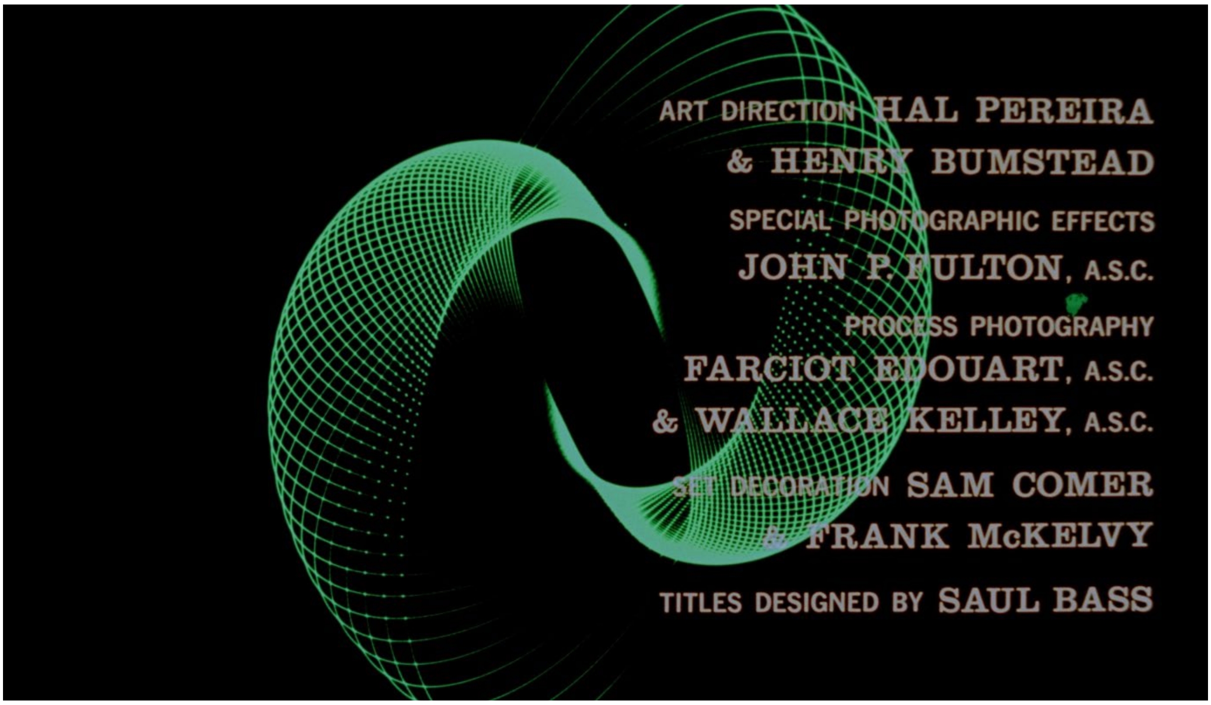


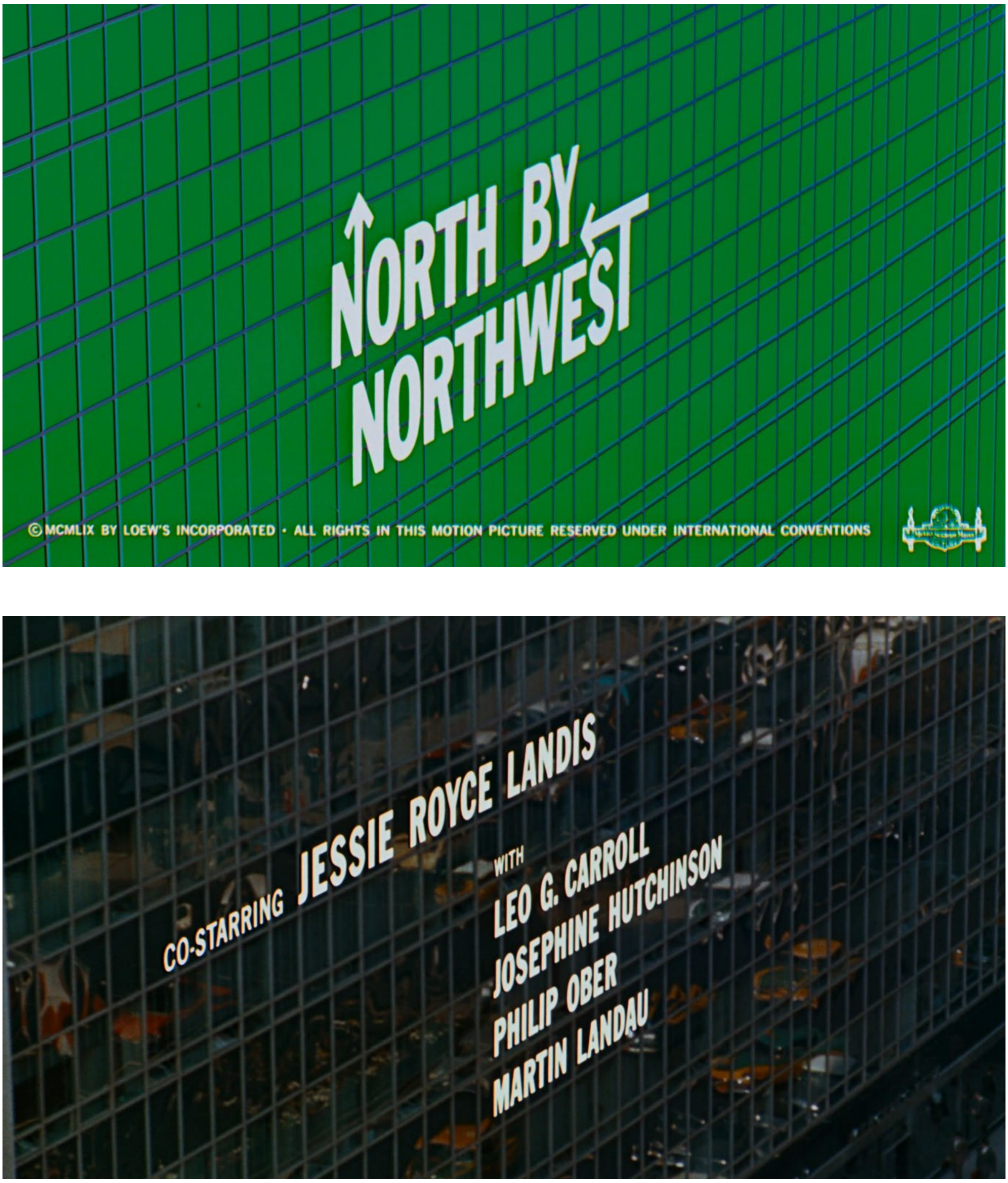
Figure 77. *Vertigo*’s title card.


CO-STARRING
BARBARA BEL GEDDES
WITH TOM HELMORE
HENRY JONES
RAYMOND BAILEY
ELLEN CORBY
KONSTANTIN SHAYNE
LEE PATRICK


ART DIRECTION HAL PEREIRA
& HENRY BUMSTEAD
SPECIAL PHOTOGRAPHIC EFFECTS
JOHN P. FULTON, A.S.C.
PROCESS PHOTOGRAPHY
FARCIOT EDOUART, A.S.C.
& WALLACE KELLEY, A.S.C.
SET DECORATION SAM COMER
& FRANK McKELVY
TITLES DESIGNED BY SAUL BASS

Figures 78-79. Examples of the spiralling imagery in *Vertigo*'s title sequence.

A more abstract Alfred Hitchcock example of narrative establishment is the title sequence for *North by Northwest* (Hitchcock, 1959). While introducing the title for the film and the main film crew credits, the animated imagery transitioning to a shot of the side of a New York building, which reflects the image of the bustling daytime street life below, shares meaning with a complex dramatic melody that shuffles between minor and major mode, and varying time signatures (see Figures 80-81). The music is not precise or easy to follow, sometimes seeming atonal in its complexity. However, the slanted window frames of the building provide symmetry but also subtle stability to the frame and the title cards. The meaning generated from this interlocking of musical chaos and visual rigidness (though with a slight italic slant in the font) anticipates a narrative that will include stability inside confusion. It is no mistake that the title sequence ends with the shrieking chaos of dominant seventh chords that never resolve in a cadence of closure, paired with Hitchcock's iconic cameo of missing a bus (see Figure 82). The film has tension that may never release or stop to wait for you.


NORTH BY NORTHWEST
© MCMLIX BY LOEW'S INCORPORATED • ALL RIGHTS IN THIS MOTION PICTURE RESERVED UNDER INTERNATIONAL CONVENTIONS
CO-STARRING JESSIE ROYCE LANDIS
WITH
LEO G. CARROLL
JOSEPHINE HUTCHINSON
PHILIP OBER
MARTIN LANDAU

Figures 80-81. The imagery of *North by Northwest*'s title card transitions to a shot of the side of a New York building.

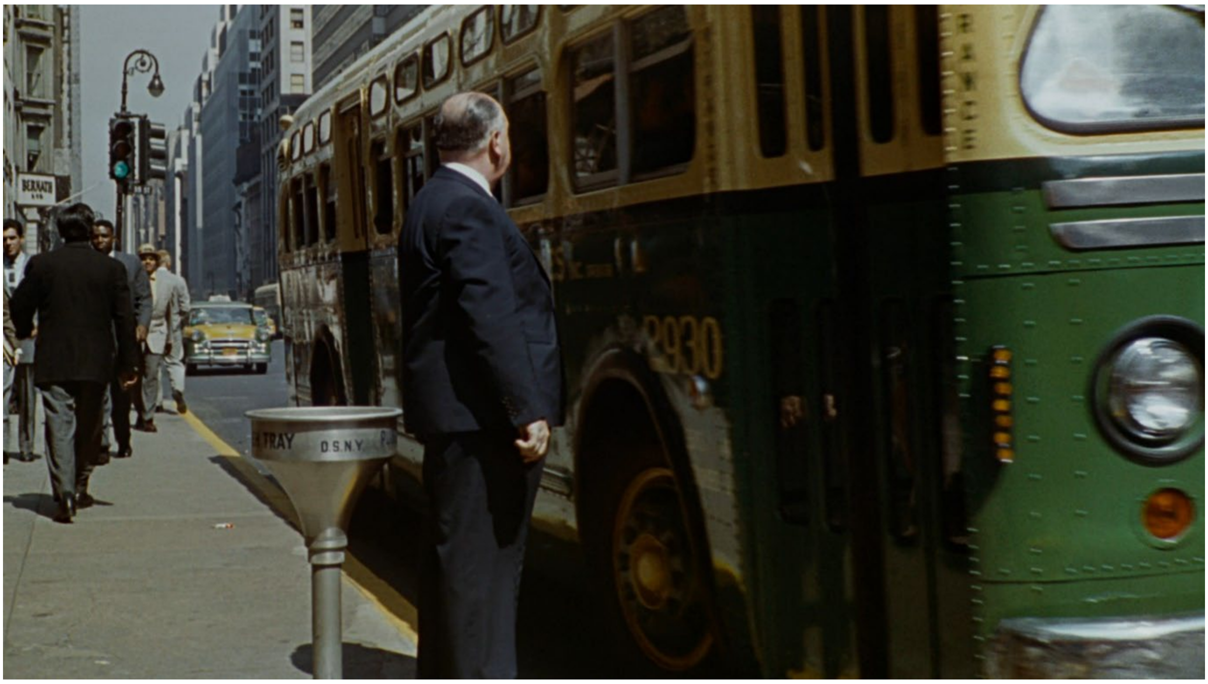
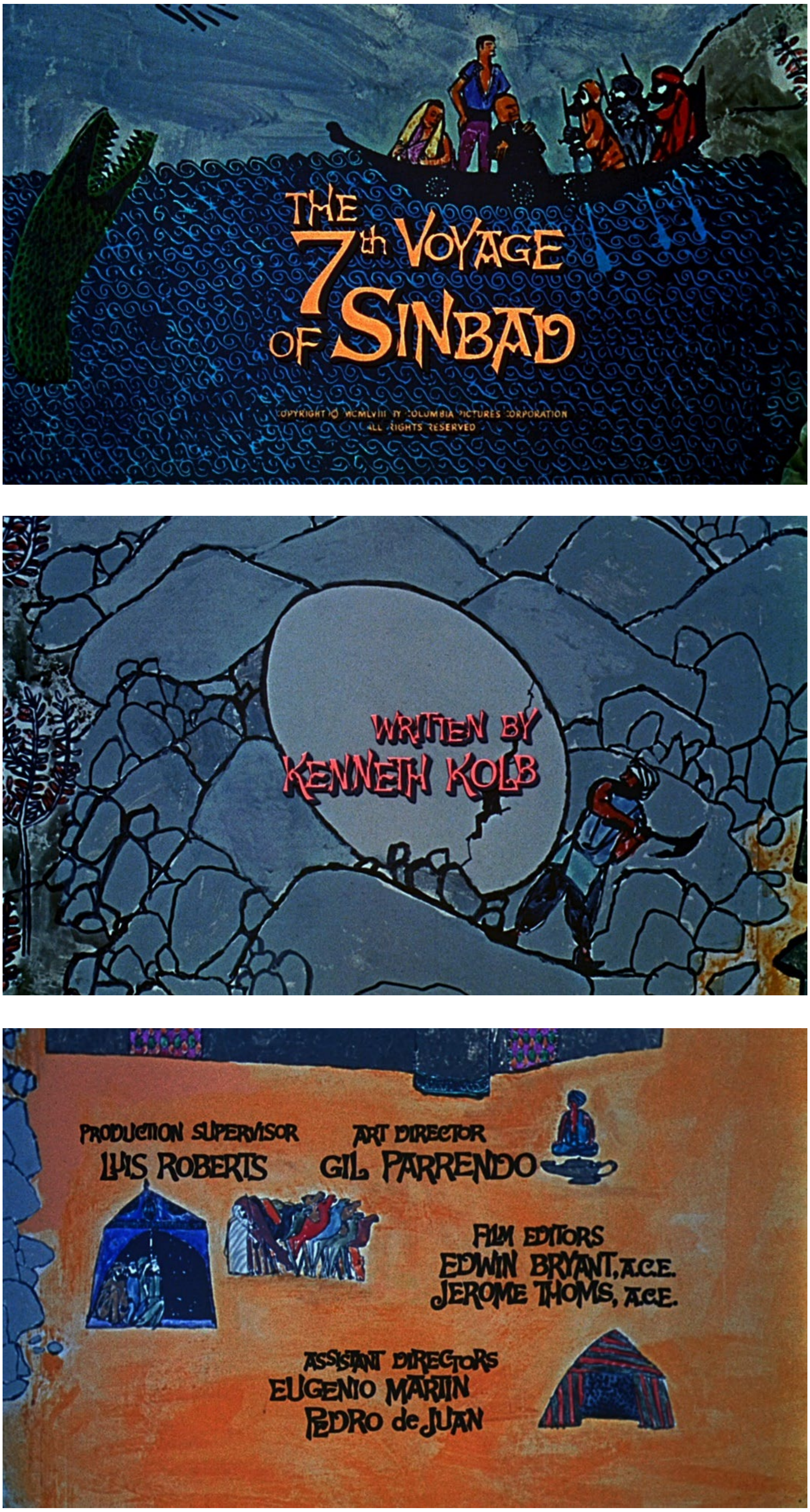


Figure 82. Alfred Hitchcock misses a bus as *North by Northwest*'s title sequence ends.

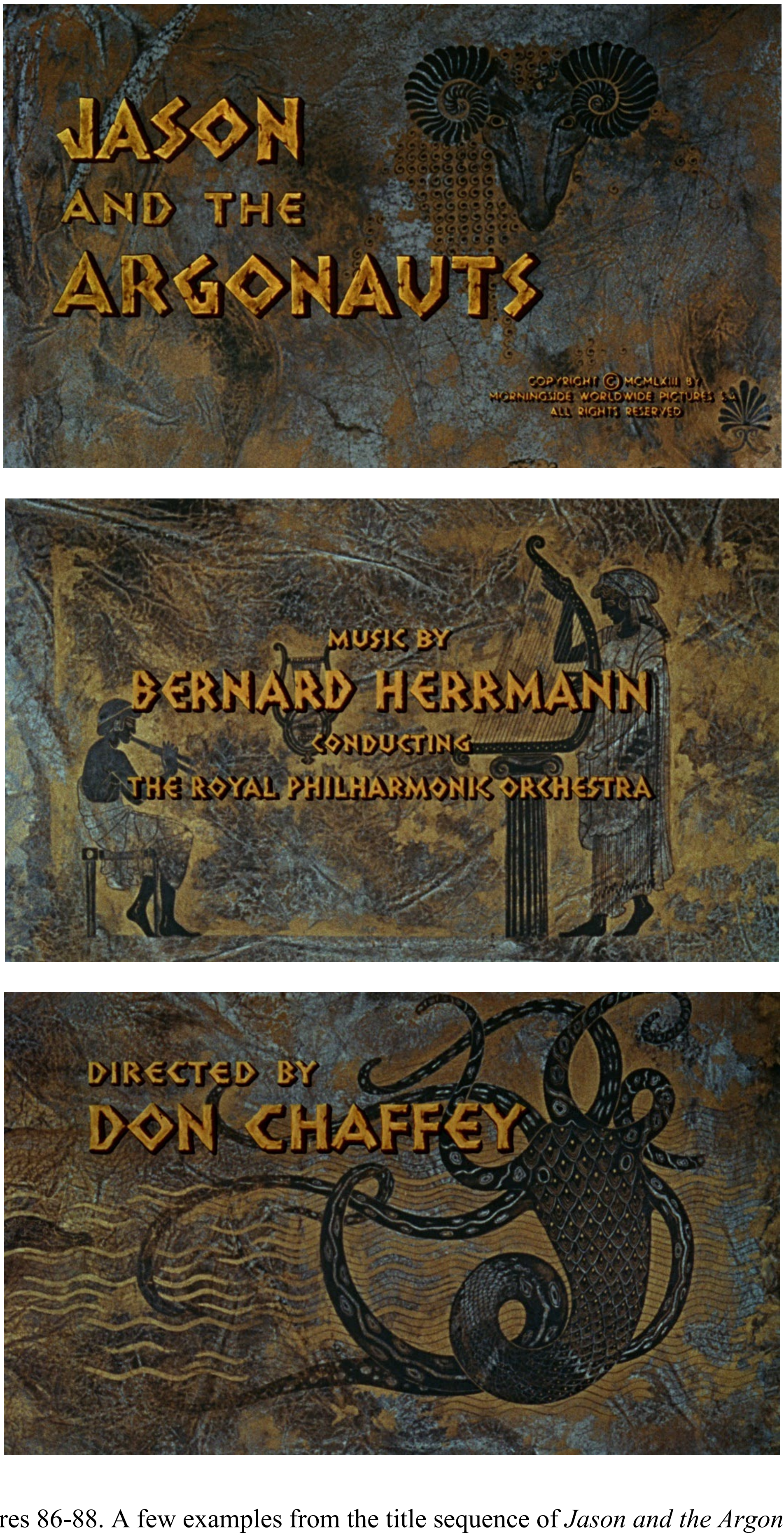
Some films choose to interlock music with animated visualisation, either mosaic or in motion. In the fantasy genre of film, *The 7th Voyage of Sinbad* (Juran, 1958) and *Jason and the Argonauts* (Chaffey, 1963), to name two examples, use brass and percussion to survey murals of the narrative that will follow.[313] The camera glides, slides and pans through painted imagery that resemble worlds that existed perhaps 'once upon a time' in a faraway mythical land. These sequences are fascinating, as they could indeed exist on their own if they were not acting as the paratext for a diegetic world that closely resembles its storytelling depictions. [See Figures 83-88]

---

[313] Both of these films were composed by Bernard Herrmann.


THE 7th VOYAGE OF SINBAD
COPYRIGHT © MCMLVIII BY COLUMBIA PICTURES CORPORATION
ALL RIGHTS RESERVED
WRITTEN BY
KENNETH KOLB
PRODUCTION SUPERVISOR
LUIS ROBERTS
ART DIRECTOR
GIL PARRENDO
FILM EDITORS
EDWIN BRYANT, A.C.E.
JEROME THOMS, A.C.E.
ASSISTANT DIRECTORS
EUGENIO MARTIN
PEDRO de JUAN

Figures 83-85. A few examples from the title sequence of *The 7th Voyage of Sinbad*.

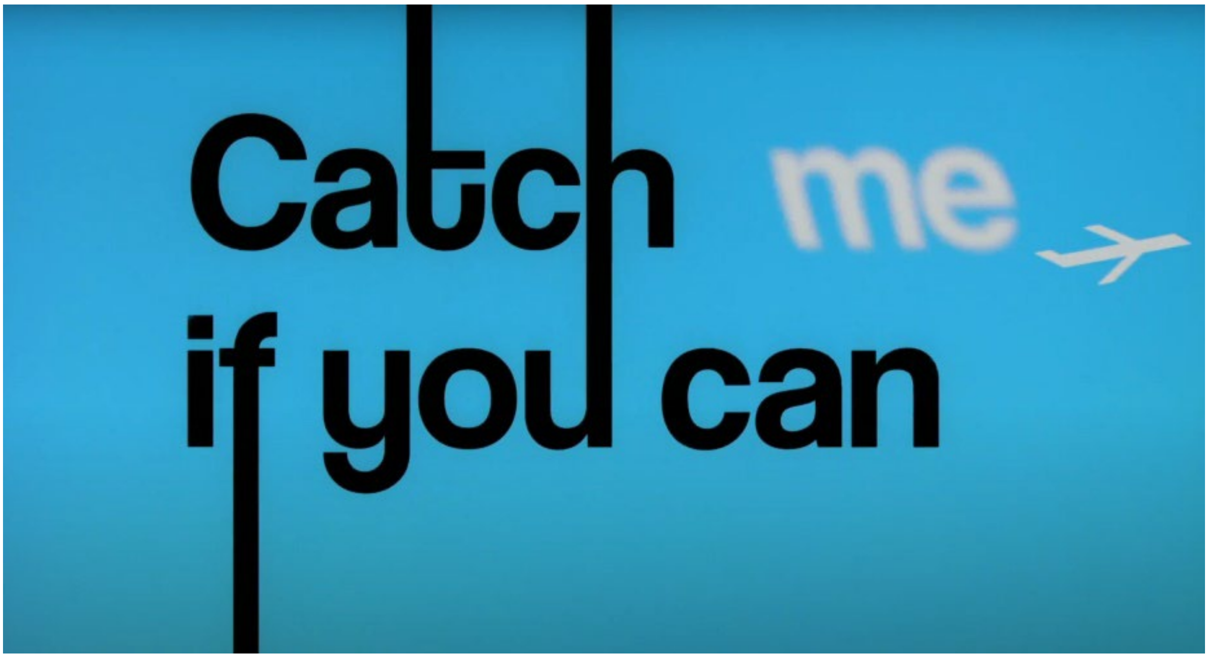

JASON
AND THE
ARGONAUTS
COPYRIGHT © MCMLXIII BY
MORNINGSIDE WORLDWIDE PICTURES S.A.
ALL RIGHTS RESERVED
MUSIC BY
BERNARD HERRMANN
CONDUCTING
THE ROYAL PHILHARMONIC ORCHESTRA
DIRECTED BY
DON CHAFFEY

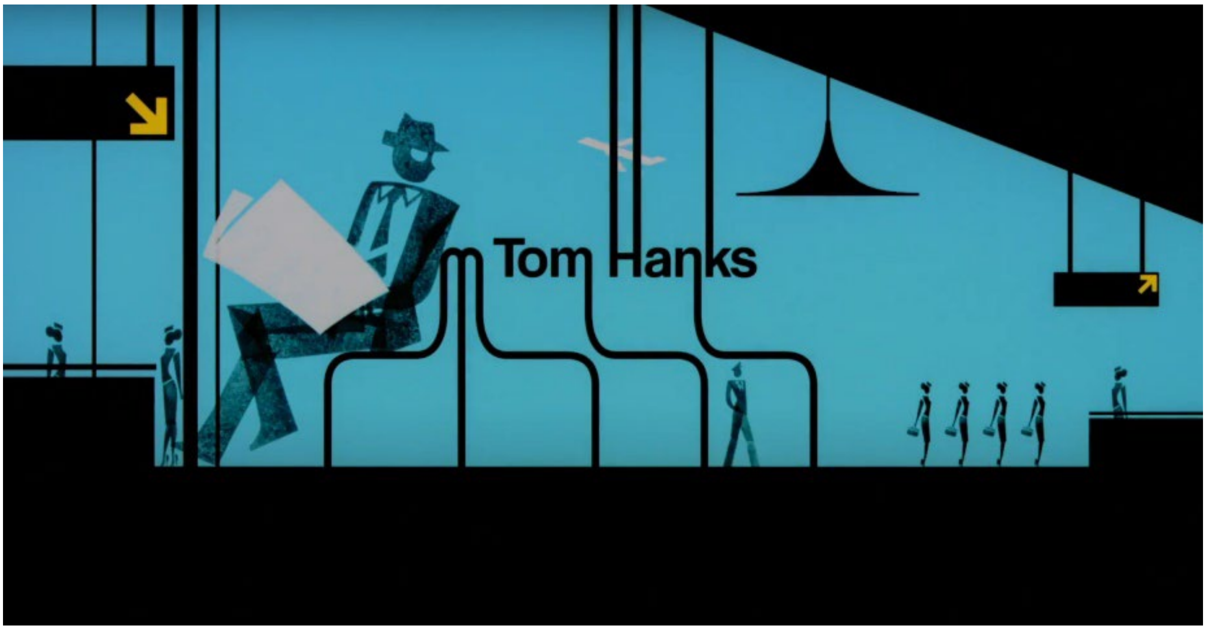
Figures 86-88. A few examples from the title sequence of *Jason and the Argonauts*.

In the title sequence for *Catch Me If You Can* (Spielberg, 2002), the spectator is presented with animation that moves without the camera having to physically drive a narrative. The suspicious, jazz-like instrumentation of saxophone, xylophone and percussion playing call and response quietly (as if hushing) is involved in a visual image of strong blue and black colours (see Figures 89-90). This colouring animates a cat and mouse chase of two silhouetted figures in various animated locations that will be established in the diegetic world of the film as Carl Hanratty and Frank Abagnale Jr.[314] The sequence does not narratively resolve as the film does but placing this paratextual music-image again anticipates the story world which is illustrated through the audio-visual imagery that accompanies the opening film credits.

---

[314] *Catch Me If You Can*'s title sequence transitions from blue and black, to yellow and black, to green and black, to red and black, to purple and black, to white and black, before returning to blue and black, to signal the various settings of the cat and mouse chase in the film's narrative.





Figures 89-90. The opening segment of *Catch Me If You Can*'s title sequence.

The title sequences for the *Star Wars* films, instead of supplying animated visual imagery of the film that will follow, provide a text scroll of the film title and a narrative blurb. The *Star Wars* title sequences operate as a bridge between the paratextual text and the diegetic realm of the narrative world, or paratextual to textual, through audio-visual harmony.[315] To open, the music cues with the title of the film in an emphatic and triumphant manner blasting a B-flat chord. The title card covers the entire film frame, extending that limit of the frame to emphasise the intergalactic scale of the film, but also the scale of this audio-visual image; the film has arrived with a bang (See Figure 91). This is not unique to *Star Wars* as, for example, the title credits for *Gone with the Wind* (Fleming, 1939) also show a title card that is too large for the film frame and must scroll entering stage left and exit stage right with grandiose orchestral music sweeping the titles along.[316] In a 1997 interview, John Williams, the composer of the *Star Wars* music, said that the audio-visual combination of the *Star Wars* title card "must speak to the collective memory."[317] The giant, gold-rimmed title moves straight back (as if moving away from the camera), disappearing into space, synchronising with a brass melody that invokes a sense of grandeur hinting at the intergalactic story to follow.[318] What follows is a blurb of the narrative that will unfold. The text scrolls into the back of the image, into space, rather than linearly from bottom to top like a normal scroll. The opening scroll is part of the space, part of the landscape, rather than existing purely as text to be read (see Figures 92-95). The lack of any cadences played by trombones and trumpets leaves the image of a non-tonal scroll literally floating into galactic space rather than a traditional fade to black to resolve the sequence, an audio-visual harmony within separate audio and visual disharmonies.[319] The audio-visual paratextual space becomes textual as the music and scroll are absorbed into the narrative space, introducing the opening images of the diegetic world in each film entry in the franchise. This audio-visual fanfare repeats in each film entry, closely tying them all within a coherent semiotic field.

---

[315] This analysis is concerned with how music operates in the paratextual/textual shift of the *Star Wars* title sequence. For an extensive music analysis of the *Star Wars* main titles, as well as the film series' use of leitmotifs, see James Buhler, "Star Wars, Music, and Myth," in *Music and Cinema*, ed. James Buhler, Caryl Flinn, and David Neumeyer (Hanover, New Hampshire: University Press of New England, 2000).

[316] The title card for '*Gone with the Wind'* uses a font style that includes paint splotches on the right side of the letters as it sweeps left, as if it has been blown 'with the wind'.

[317] Buhler, "Star Wars, Music, and Myth," 34.

[318] James Buhler describes the letters of the title as "giant black...rimmed with gold...opaque" so the stars cannot be seen behind. See, ibid., 33-34.

[319] James Buhler mentions the lack of a tonal cadence makes the *music* open-ended and ambiguous as the text scrolls into space. However, due to this synchronisation of music and image, this lack of tonic cadence justifies the behaviour of the non-tonal scroll. See, ibid., 36-37.

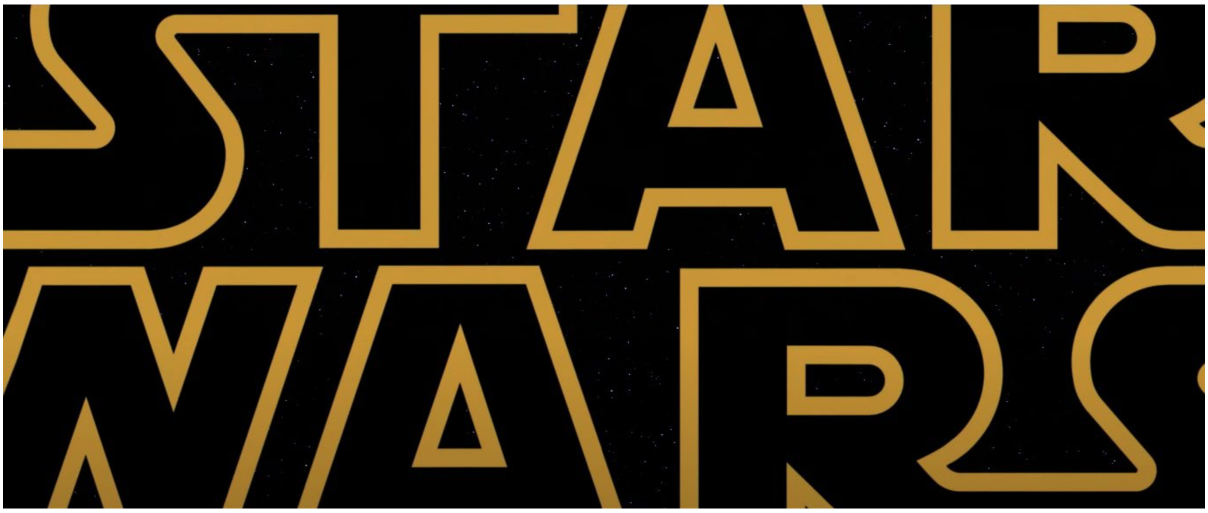
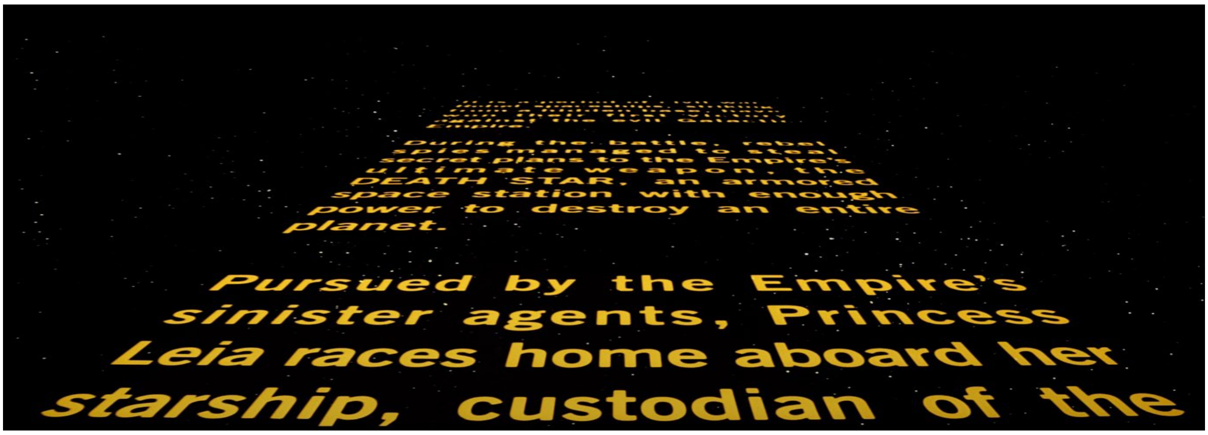


Figure 91. The title card for *Star Wars* begins larger than the film frame emphasising the intergalactic scale of the film and audio-visual image.


During the battle, rebel
spies managed to steal
secret plans to the Empire's
ultimate weapon, the
DEATH STAR, an armored
space station with enough
power to destroy an entire
planet.
Pursued by the Empire's
sinister agents, Princess
Leia races home aboard her
starship, custodian of the

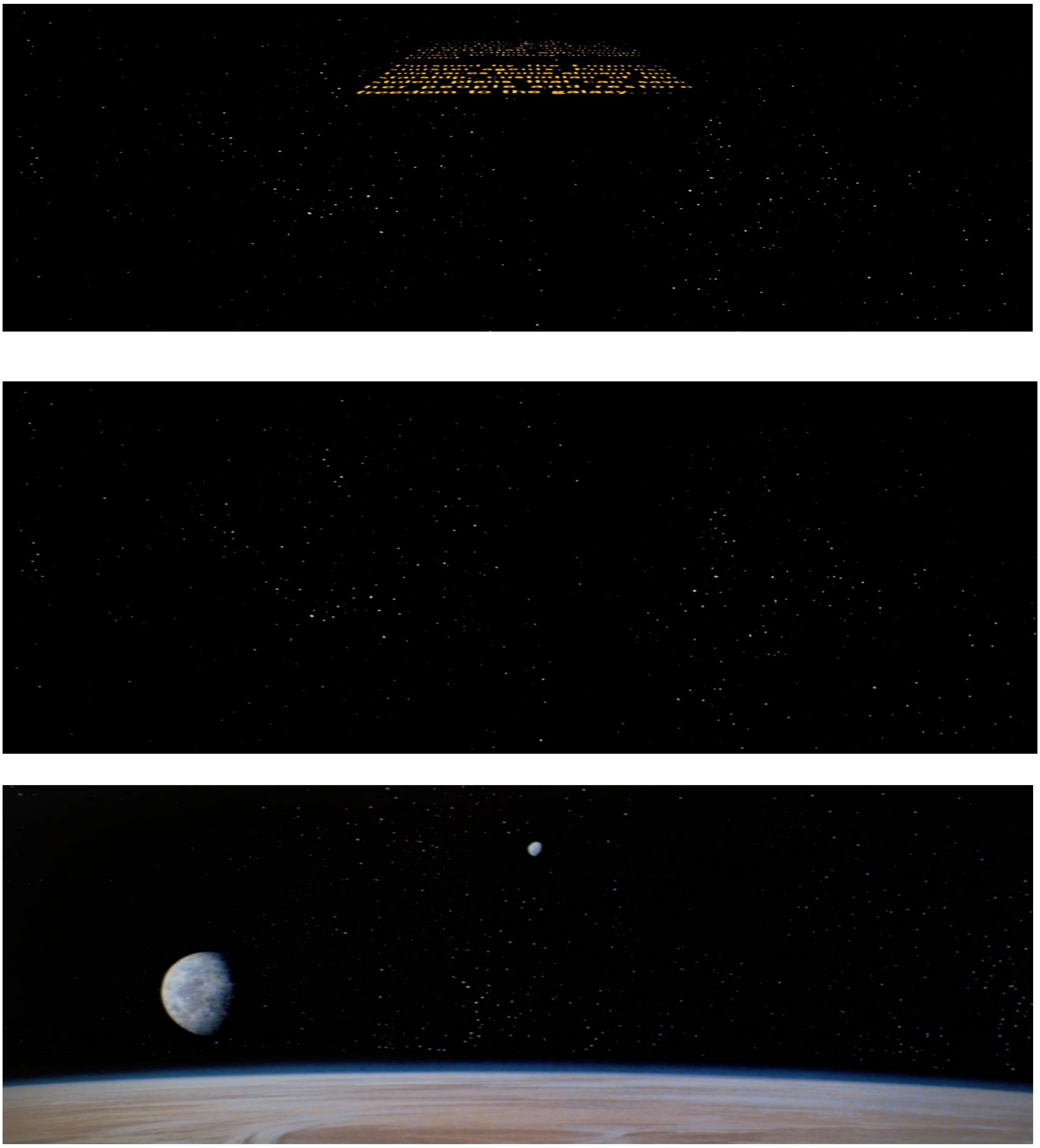
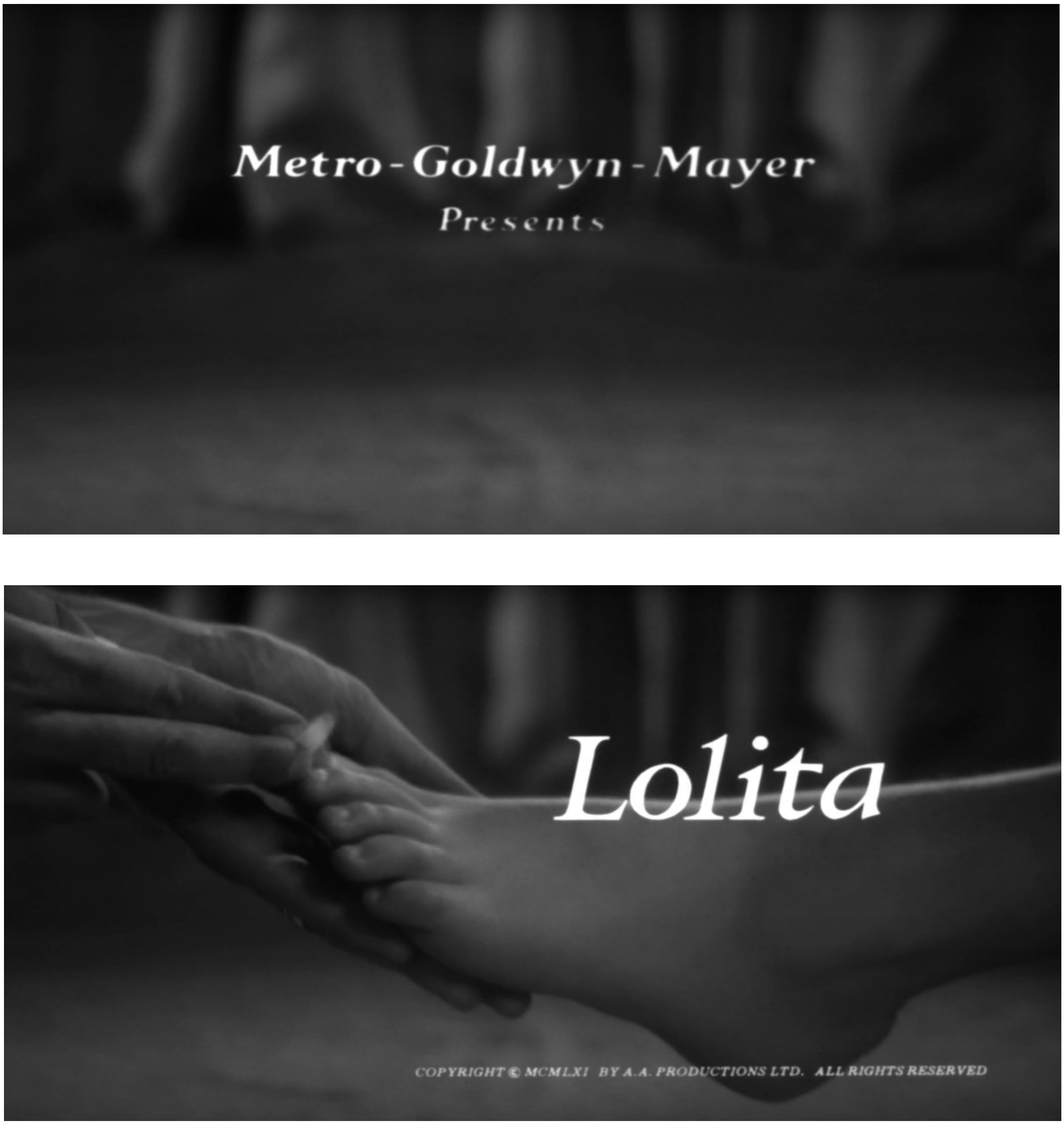


Figure 92-95. The opening scroll for *Star Wars* scrolls back into the image of space and becomes part of the narrative landscape.

Title sequences can also exist within the diegesis of the film world. In *American Beauty* (Mendes, 1999), the music begins as paratextual, with xylophone and percussion moulding the opening credits, but the duration of the music lasts longer than the titles, entering the diegetic world as Lester introduces his life in monologue and a typical morning of his mundane life. In *Taxi Driver* (Scorsese, 1976), the opening credits overlay Travis driving his taxi with his close-up gaze situated in a shot reverse shot setup of the Manhattan street haze and nightlife lighting, while a romantic saxophone melody engulfs the visual space. Again, the music's duration outlasts the film's main titles, and the first scene commences with Travis receiving a taxi driver license. The title sequence would have to have been located somewhere later in the story time of the film if Travis has not yet driven a taxi, which makes the title sequence an episode in his life to come in the film narrative. Similarly, the title sequence of *Lolita* (Kubrick, 1962) begins with an out of focus shot of a bedroom before a young female's foot enters the frame, suggesting an image that is part of the diegetic world. As soon as the foot enters the frame, a romantic piano melody cues, but only to be interrupted by an older male hand to supply a pedicure, prompting the film's title card (see Figures 96-97). This audio-visual grammar has a similar theme to the film, that is, an older man's obsession with a 12-year-old girl, and suspected paedophilia narrated by the suspect, which is why the imagery has an unsettling undertone that jars with the romantic music.


Metro-Goldwyn-Mayer
Presents
Lolita
COPYRIGHT © MCMLXI BY A.A. PRODUCTIONS LTD. ALL RIGHTS RESERVED

Figures 96-97. An out of focus shot of a bedroom is intruded by a foot and hand entering frame as the title card for *Lolita* appears.

In some films, the title sequence is delayed, and instead of having a paratextual function, cues after an initial film sequence. Such sequences function as a prologue. *Out of Africa*'s (Pollack, 1985) title sequence does not appear for more than five minutes into the film's duration. Delaying the sequence for that amount of time situates the title music-image in the diegesis of the film, as the narrative continues to play out with imagery of a train transporting Karen to her "…farm in Africa." Both the *Mission: Impossible* and *Marvel* franchises also delay their title sequences, both appearing as paratextual sequences as they do not contain audio-visual imagery from the diegetic world but introduce the audio-visual themes for the film.

Interestingly, the title sequence for *Halloween* (Carpenter, 1978), with its imagery of a jack-o'-lantern glowing in darkness interlocked with descending perfect 5th intervals played on piano in the odd 5/4 time signature, is paratextual; while the 2018 reboot, *Halloween* (Gordon Green, 2018), has a delayed title sequence, using an almost identical title sequence. The difference in music-image cues lies in the film series' story world. The original 1978 film is the first in a string of *Halloween* films, and the first to introduce the story world and signature character and villain, Michael Myers. The title sequence acts to introduce the story world, whereas in the 2018 reboot the story world already exists. The reboot directly follows the original film some 40 years later with two crime podcasters Aaron and Dana going to visit Michael Myers at a psychiatric hospital following Michael's killing spree 40 years earlier in the original film. Aaron holds up Michael's iconic mask in front of Michael during their encounter, shouting "say something" which is interrupted by the title sequence. The delayed title sequence (with its subtle variant of a regenerating jack-o'-lantern) acts as a revival of Michael's habits, but also regenerates the film series. In the delayed title sequence music-image, the film is not just introducing the film that will follow but commenting on a story world that is larger than the internal duration of the film's narrative. [See Figures 98-102]

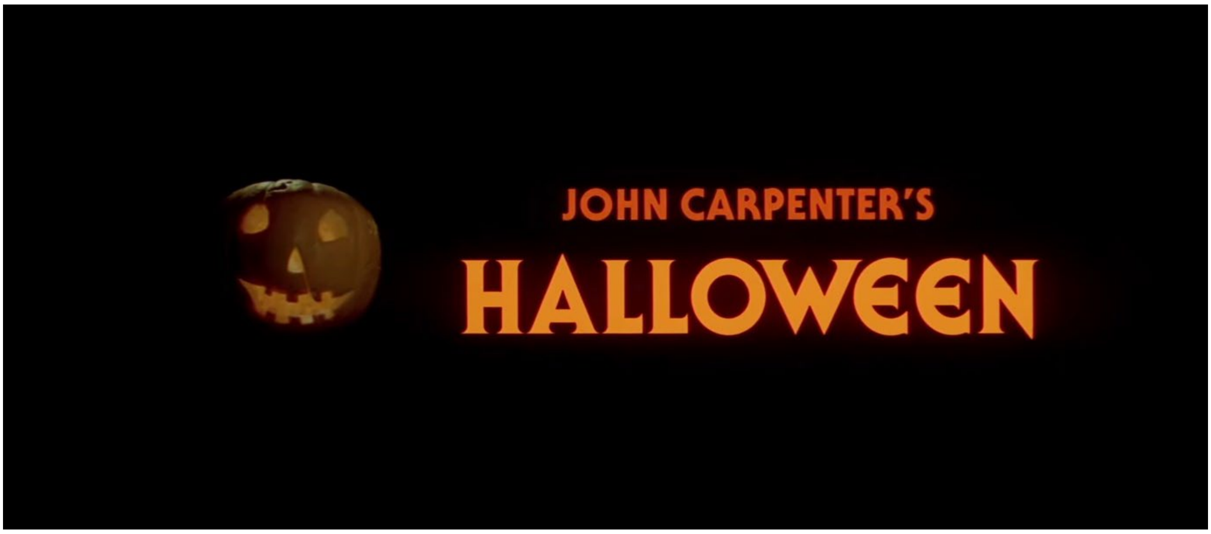
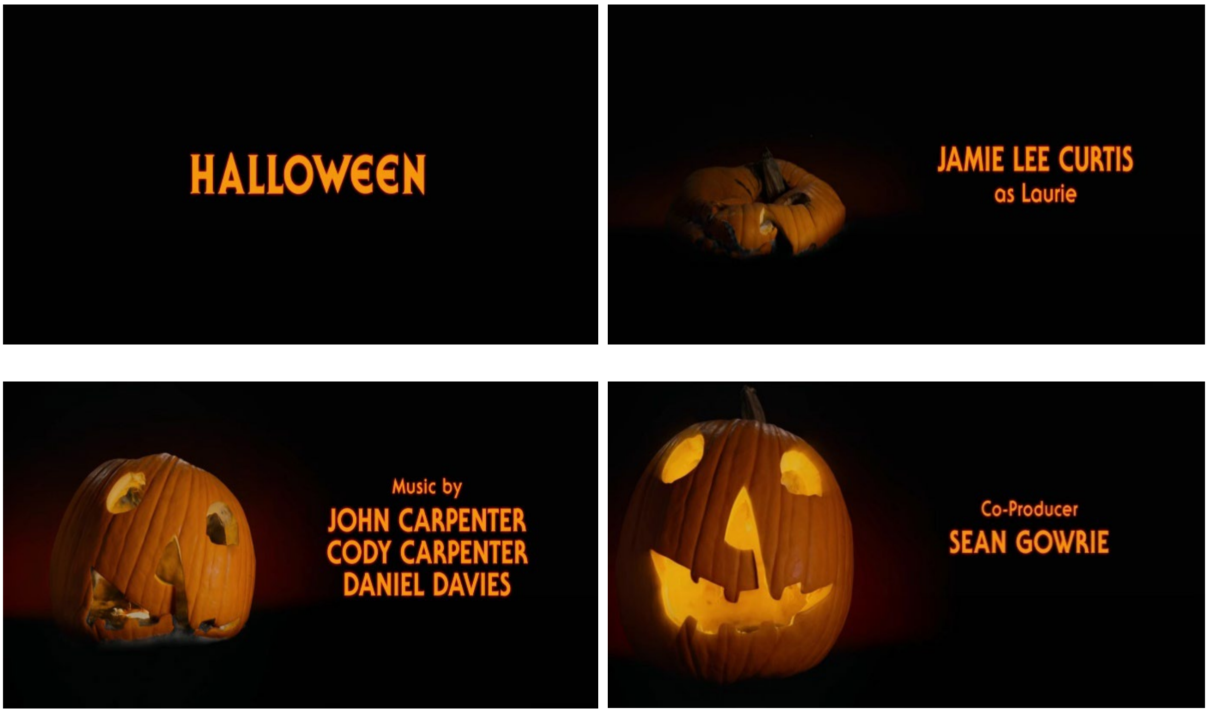

JOHN CARPENTER'S
HALLOWEEN

Figure 98. The title card for *Halloween* (Carpenter, 1978).


HALLOWEEN
JAMIE LEE CURTIS
as Laurie
Music by
JOHN CARPENTER
CODY CARPENTER
DANIEL DAVIES
Co-Producer
SEAN GOWRIE

Figures 99-102. The jack-o'-lantern comes to life, reborn, in the title card for *Halloween* (Gordon Green, 2018)

Conclusion

Music-images exist because they are heard, but to be heard they must be cued. If a music-image exists only when cued, it is experienced only at certain audio-visual interlockings in a film. This chapter has explored the several ways in which an audio-visual analysis of a given sequence generates different meanings to an engagement with the pure codes of musical form, and the pure codes of visual form.

There are points in a film in which music-images may not necessarily be acknowledged (or may remain unheard), but this does not mean they are not encountered by the spectator. In various examples taken from films such as *Wild Strawberries* and *The Thing*, music-images may not be immediately recognised but still cued, and meaning in that sense is always a potential outcome.

In the extended analysis of *The Hours*, chosen because I found the film's narrative to be complex, music-images occurred at the unexpected moments of shared emotional states between the three main characters across three different time periods. The analysis of *The Hours* also revealed that when the same music is used at different points in the film, it evokes a memory of the previous time it occurred, connecting multiple music-images thematically.

When music repeats, however, the music-image is different in that very repetition, as the new music-image bears memory within its aural and visual codes. There is a newness every time a music-image generates meaning, as opposed to isolating the music from the music-image in a replay of pure musical codes per se, however the thematic music-image exists from previous music-images. The thematic music-image which bears memory raised the question of when the first possible music-image in a film occurs. Music-images can be cued even before the first scene, in the paratextual realm of title credits. Through the exploration of various paratextual title sequences, these early music-images can introduce narrative themes and characters, but also worlds.

There are various points in a film when there are music-images and also when there are no music-images, but this discussion of ‘when’ music-images occur demonstrates that film music exists as part of the construction of an interlocked audio-visual image of experience. To separate sight and sound would be a disservice to both. A discussion of ‘when’ music-images occur stretches the ontology of film music. Film music does not just exist, it begins and ends within precise parameters of a film which lay the very foundation of its existence.

Chapter 6: Where is the Music-Image?

A common mode of analysis in Film Music Studies is to locate ‘where’ the film music is coming from, that is, restricting film music to a spatial entity. However, this often obfuscates an analysis of what music does to the spectator’s experience of the film. Emilio Audissino points out that too much attention has been made to discovering 'where' film music comes from in any given film rather than its function.[320] This chapter instead explores whether the location of film music generates a certain audio-visual image of experience. Locating the music source in film is not analysis unless the location matters to film narration as film music does not exist in isolation. Instead, does locating the music change the experience of the film? Film music does not exist as *film* music unless it is part of the constructed narrative image of experience (as explained in previous chapters), and neither does defining the film diegesis.

The location of film music exists only in diegesis, a characteristic of sound that Andrew Knight-Hill argues only exists in experiencing film as an audio-visual phenomenon.[321] Shifting the analysis of diegetic film music from the characteristics of sound to the experience of the spectator considers the involvement of the spectator in the meaning of narrative image. This narrative image is constructed (an idea teased out through David Bordwell’s Principles of Narration in this chapter) with the location of film music within its diegesis. Experiencing film and music as a diegetic, audio-visual phenomenon is a function of the music-image. Therefore, this chapter follows the chapters on the nondiegetic music-image, specifically addressing the music-image within the diegesis. Diegetic film music exists in the construction of the image but departs from the nondiegetic music-image as it carries different narrative meaning located within an audio-visual space. Identifying the location of film music within an audio-visual space cannot be done without the music-image. Daniel Yacavone uses interesting language to distinguish nondiegetic from diegetic music in using the word 'experience': "The underscore of a film is typically referred to as nondiegetic because it is deemed not to be a part of the characters' experience of their world. Rather, it is an articulation of that experience as an act of narration construed as far less visible (if visible at all) than the reality this presented."[322] To reverse this distinction, diegetic music then is part of the characters' experience of their world. However, I argue in this chapter that diegetic film music is still an articulation of that experience as an act of narration. Since the music-image is the semiotic ambiguity that holds film music and film image together through spectator experience, to speak of diegetic film music through this model is to interlock its musical codes with the image.

---

[320] Audissino, *Film/Music Analysis: A Film Studies Approach*, 157.

[321] Andrew Knight-Hill, "Sonic Diegesis: Reality and the Expressive Potential of Sound in Narrative Film," *Quarterly Review of Film and Video* 36, no. 8 (2019): 653.

Complicating the traditional distinction between diegetic and nondiegetic musical fields, this chapter redefines diegetic film music using a wide range of film examples through the discussion of visible and invisible source music, film music’s relationship with the stage, and concluding with the sourceless musical sources of the film musical.

What is diegetic film music?

Étienne Souriau first introduced the term diegesis to cinema in 1953, deriving from the Greek word for ‘narration’, meaning the recital of facts.[323] In Film Music Studies, the term ‘diegesis’ first appeared in Claudia Gorbman’s article titled *Narrative Film Music*, released in 1980 in the inaugural issue of *Yale French Studies*. Gorbman here translates Étienne Souriau's *L'Univers filmique*, published in 1953, which defines diegesis as all that belongs to the narrated story and world proposed by the film's fiction. She then translates Gérard Genette’s definition of diegesis from Genette’s 1966 book *Figures* as “the spatiotemporal universe referred to by the primary narration.”[324] Gorbman then combines the two definitions to introduce diegesis to Film Music Studies for the first time: "the narratively implied spatiotemporal world of the actions and characters."[325] When applying diegesis to film music for the first time, Gorbman asserts that diegetic film music is music that “(apparently) issues from a source within the narrative.”[326] Narrative has much to do with the construction of cinematic space where this music is located. Referencing Lev Kuleshov's experiment (the Kuleshov effect), Gorbman supplies the idea of a “supposed” reality that constructs the diegesis.[327] The spectator creates this diegesis through the derived or experienced narrative.[328] The construction of diegesis makes narrative not a predetermined ‘story-space’ (that might, for example, be located originally within the script), but an experientially *constructed* space.

---

[322] Yacavone, "Spaces, Gaps, and Levels: From the Diegetic to the Aesthetic in Film Theory," 22.
[323] Metz, "The Cinema: Language or Language System?," 97-98.

Redefining diegetic film music

Diegetic film music needs redefining because a central element in the discussion of diegesis in Film Music Studies of the past 40 years, since Claudia Gorbman mentioned film diegesis in 1980, is to locate ‘where’ film music is; that is, the diegesis is commonly restricted to a discussion of the location of the source.[329] In my analysis of film music in this thesis, I argue that the music-image is not derived from its location, that is, the coordinates of ‘where’ the music is, but in the specific field of meanings generated by that location. Source music emits from the same cinema speakers as nondiegetic music, making it part of a narrative and semiotic system that constructs a film image. Obviously, it constructs a very different meaning to that of nondiegetic music as there is an added layer of meaning to being aware of its spatiotemporal location. I argued in Chapter 3 that the main components of the music-image experience are harmony/disharmony, tone, time, and space. These components do not vanish with the supposed mystery of music's location but are intrinsic to the construction of the music-image.

---

[324] Gorbman, "Narrative Film Music," 195.
[325] Ibid.
[326] Ibid., 197.
[327] Ibid., 195-96.
[328] Ibid., 196.
[329] Ibid., 195.

The most obvious example of diegetic film music is music that is visibly located within the narrative space; in other words, the music source can be seen within the film frame.[330] This is what most Film Music Studies scholarship defines as 'source music'. To be consistent with the definition of 'underscore' from Chapters 3 and 4, I will reference Karlin and Wright again here to define 'source music'. Karlin and Wright define 'source music' as "music that appears to come from the scene; music the actors [thus, characters] hear."[331] Royal S Brown describes this source music as bound within the diegesis, such as a radio, a phonograph, a person singing, an orchestra playing – all of which can be heard by characters within the film.[332] There are two ways of achieving this acknowledgement: either showing an object with a known function to play music, or a character performing music on a musical instrument. Film music analysis tends to favour validating the realism of narrative space regarding the input of 'source music' and 'underscore'.[333] In the introduction to his book on finding a diegetic middle ground (neither diegetic nor nondiegetic), Morris B Holbrook argues that diegetic music is both "produced within the film as part of on-screen action or mise-en-scène" and reinforces the "realism and verisimilitude [truth of what is real] of the mise-en-scène."[334] Holbrook further creates the binary distinction of source/underscore as realistic depiction/dramatic development and objective/subjective.[335] Michel Chion also upholds this distinction:

---

[330] Stanley Cavell remarks on what it means to know a sound source is what it is. He uses the example of a phonograph playing a record, and saying while listening, "That's an English horn," to someone who does not know what an English horn looks or sounds like. In order to identify the sound source that is heard, there should be prior knowledge of its existence. See, Cavell, *The World Viewed: Reflections on the Ontology of Film*, 18.

[331] Karlin, *On the Track: A Guide to Contemporary Film Scoring*, 39.

[332] Royal S. Brown, *Overtones and Undertones: Reading FIlm Music* (Berkeley; Los Angeles; London: University of California Press, 1994), 67.

[333] Eisler, *Composing for the Films*, 228.

> "All music in a film, especially pit music, can function like the spatiotemporal equivalent of a railroad switch. This is to say that music enjoys the status of being a little freer of barriers of time and space than the other sound and visual elements. The latter are obliged to remain clearly defined in their relation to the diegetic space and to a linear and chronological notion of time."[336]

This dichotomous categorisation of film music stems from a desire to locate it, that is, to say where it comes from in the film. Why is it that traditional film music scholarship is so interested locating film music? As I have argued, this commonly applied dichotomy is limited in so far as it must separate musical materials from the visual materials of the image; in my own rejection of this position, I agree with the position of Royal Brown, who argues that it is wrong to assume that film music has a binary function based on where it is located.[337]

---

[334] Morris B. Holbrook, *Music, Movies, Meanings, and Markets: Cinemajazzamatazz* (New York; London: Routledge, 2011), 4.
[335] Ibid., 5.
[336] Chion, *Audio-Vision: Sound on Screen*, 81.
[337] Brown, *Overtones and Undertones: Reading FIlm Music*, 69.

Principles of film narrative

What is meant by constructed narrative? I adopt the notion of ‘constructed narrative’ from David Bordwell’s ‘Principles of Narration,’ which can help understand how diegetic film music exists in the construction of film narration through my model of the music-image. Ben Winters cites Bordwell when attempting to define the relationship between film music and narration in “the musical performance of ‘source music’ [as] exist[ing] in both the *syuzhet* and the *fabula*….”[338] Bordwell links the construction of narration to the imagination of the spectator, using the Russian Formalist terms ‘Fabula’ and ‘Syuzhet’ – the spectator makes sense of the narrative world through causal or spatial or temporal links.[339] ‘Fabula’ translates normally as ‘story’, requiring the spectator to “construct the story of the ongoing inquiry while at the same time framing and testing hypotheses of past events.”[340] Without the fabula, it is just space in which there is no cause and effect; otherwise it is merely a series of images. In the fabula, there is now a pattern of time, space, and causality: “the fabula is the pattern which perceivers of narratives create through assumptions and inferences.”[341] The spectator assumes, anticipates, and worries until the moment space and causality is set; narrative is constructed through time and experience, which means the spectator has a role to play in this construction. The ‘syuzhet’, which translates as ‘plot’, is the “actual arrangement and presentation of the fabula in the film.”[342] The syuzhet is therefore a kind of patterning, or the bridge between image at face value and fabula. Syuzhet patterns cinematic form and style, the ingredients to the medium of film (including film music), in various ways before reaching the fabula. The difference between syuzhet and style is that syuzhet embodies the film as a dramaturgical process, whereas style embodies film as a technical one.[343] The reason a film can have 'diegetic' music is based on the patterning of audio-visual style and syuzhet (plot) that informs a construction of narrative, or fabula.

---

[338] Winters, "The Non-Diegetic Fallacy: Film, Music, and Narrative Space," 231.
[339] David Bordwell, "Principles of Narration," in *Narration in the Fiction Film* (London: Routledge, 1986), 49.
[340] Ibid.
[341] Ibid.
[342] Ibid., 50.

The function and effect of diegetic music on the spectator does not require it to be located. Since it is present within the frame, diegetic film music is intrinsic to narration. And yet, narration is only one part of diegetic music. Thinking of audio-visual space as story-space includes diegetic film music as part of image construction that generates one channel of interpretative meaning through semiotic systems and inference.

Visibly sourced film music

I want to begin the exploration of film music's function when inside the narrative space with music that is visibly located within the space. Being visibly present is the most straightforward acknowledgement that diegetic film music is within the space. If the music is 'seen', as Anahid Kassabian asserts, being performed within the space, it is 'visual' source music.[344] Seeing the visual source music, the spectator assumes that the characters sharing the space can hear it. The spectator relates the musicological and spatial components of the music with the characters' actions which affects the meaning a music-image generates. This example is the most stable ontology of 'diegetic' film music. The music affects the characters within the frame, which in turn constructs the music-image within the narrative. In the following discussion, I provide several examples of film music which appears visually in the film frame while the music is playing, but this should not be confused with music that is outside the film frame, and yet still sounding within the diegetic space encompassed within the frame. In this example, it is very rare for music to appear onscreen for the entirety of musical duration, that is, there is always at least one cut during the duration of the music, either to someone listening or watching situated in another frame (thus requiring the cut). A classic example is Harpo Marx playing the harp (or in the case of *Go West* (Buzzell, 1940), playing the loom) in all but a few of his onscreen appearances.[345] In *Go West*, *Horse Feathers* McLeod, 1932), and *A Night in Casablanca* (Mayo, 1946), the camera never cuts away from Harpo playing, and only changes camera angles periodically (see Figures 103-105). However, in *A Night at the Opera* (Wood, 1935), the camera cuts to the adults and children watching and listening (see Figures 106-108). The difference relates to narration. In the first three films, Harpo has a minimal to non-existent audience within the film space, and so performs the music to the spectator only; in *A Night at the Opera*, the top deck of the ocean liner is packed with an audience of adults and children surrounding him as he plays. I mention these instances because their extended visual attention with little interruption in the film form or by bystanders means the visual source music is rooted in performance. These instances exist to entertain, to exhibit the talent of the character or performer, in this case Harpo Marx's talent for playing the harp. Visual source music predates film, going back to theatre and opera with music performed onstage rather than in the orchestra pit, for example in Wolfgang Mozart's *Don Giovanni*, premiering in 1787.[346] A musician performing music onstage next to actors who have dialogue might hamper audience comprehension of the scene, and so music was used as foreground that replaces speech.

---

[343] Ibid.

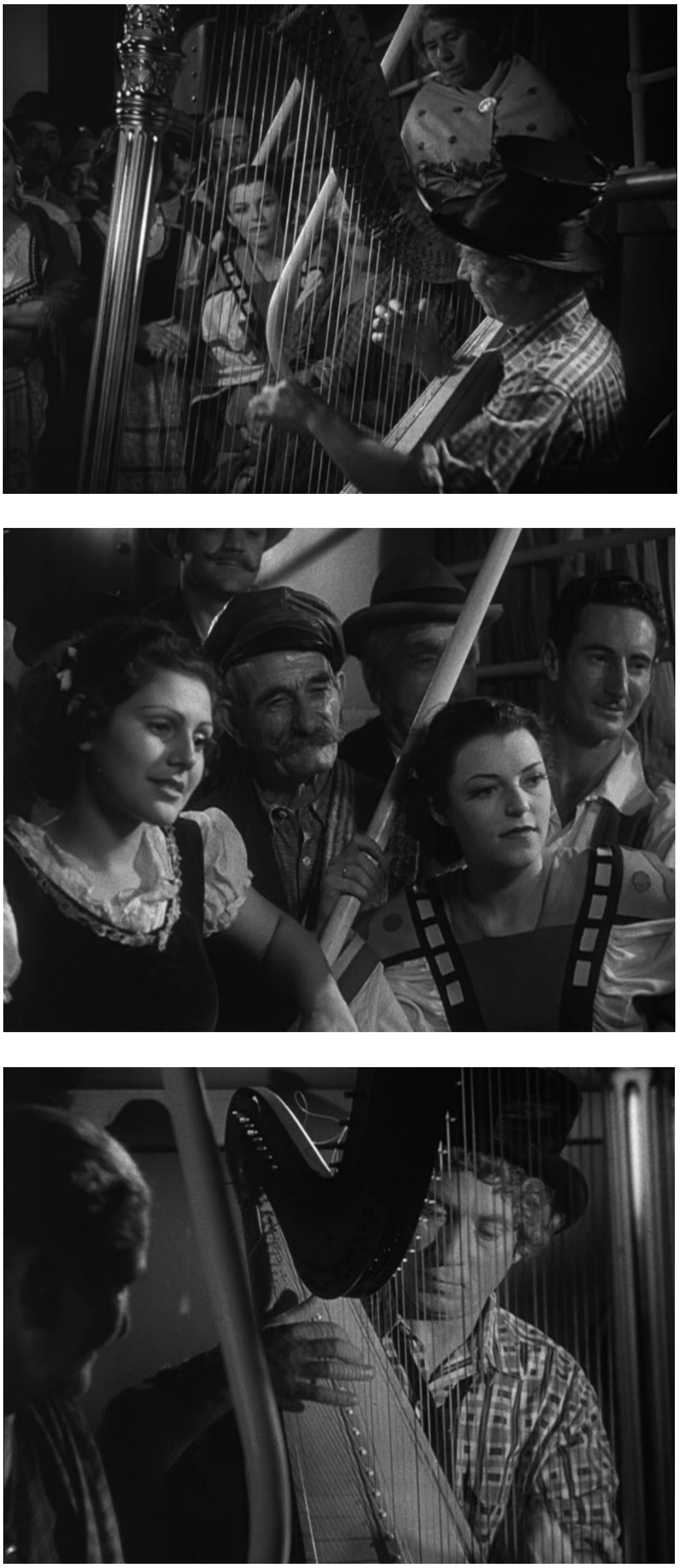
[344] Anahid Kassabian, *Hearing Film: Tracking Identifications in Contemporary Hollywood Film Music* (New York: Routledge, 2001), 44.

---

[345] The few films in which he broke his routine of playing the harp was in *Go West* (he plays a loom, similar in timbre to a harp), *At the Races*, and *Duck Soup*.

[346] Axel Berndt, "Diegetic Music: New Interactive Experiences," in *Game Sound Technology and Player Interaction: Concepts and Developments*, ed. Mark Grimshaw (Hershey, PA: IGI Global, 2011), 22.

Figures 103-105. Harpo Marx is visibly sourced performing for the entirety of the music’s duration in *Go West* (top), *Horse Feathers* (middle), and *A Night in Casablanca* (bottom).



Figures 106-108. The camera cuts between Harpo Marx performing and the audience listening, in *A Night at the Opera*.

Naturally, when the first instance of film sound entered the diegetic world in 1927 with *The Jazz Singer* (Alan Crosland, 1927) through the new technology of the Vitaphone, synchronising sound with the screen world, it would be a performance.[347] James Buhler and David Neumeyer suggest that in the silent film era, diegetic music was referred to as cue music instead of source music as it was performed within the theatre space, along with diegetic sound effects.[348] Thus it should be said that diegetic music was at that time not yet diegetic music, but rather a sound effect. For the first audiences of *The Jazz Singer* to witness diegetic sound and music within the film space, and not the theatre space, it had to be visualised. Therefore, when Al Jolson first sings and dances in the film, audiences needed to see it to believe it. Al Jolson and the orchestra are onscreen, on the stage of a nightclub, with the camera located where the bar patrons are sitting, making the 'cue' now a performance for both the cinematic spectator in the theatre space and the film spectator in the film space. The camera cuts only three times for a total of a few seconds during the performance, but the point of this is that the music now rests in the spatial parameters of the fabula, constructed by the spectator. Paul Mazey points out that the "…diegetic score complements the appearance that the characters show to the [screen] world…," and this is done through music's relationship with the film form, not just the spectator, to create intrinsically filmic narration through image construction.[349] [See Figures 109-112]

---

[347] *The Jazz Singer* was not the first attempt at sound onscreen, but it was the first feature length film with synchronised sound and diegetic music.

[348] Buhler, "Music and the Ontology of the Sound Film: The Classical Hollywood System," 20.

[349] Paul Mazey, "Restrained Airs: The Diegetic Surface and Nondiegetic Depth of British Film Music," *Journal of British Cinema and Television* 16, no. 4 (2019): 441.

Figures 109-112. As Al Johnson performs in a shot lasting 63 seconds (top left), the camera cuts away for 4 seconds (top right) and 3 seconds (bottom left), before cutting back to Al Johnson (bottom right), as the music rests in the fabula.

The moments of visual source music remaining completely visual (that is, remaining onscreen for the entire musical duration) only create a relationship between performer (or audible object) and spectator, and thus have a limited engagement within narration. In the next section, I explore what happens to the music-image when the music rests within the construction of film narration.

The invisible source

The most common form of film music that can be sourced within the film space generally does not appear on screen for its entire duration. This brings the music-image experience in line with nondiegetic film music as both forms are not seen but still heard. However, the invisible source is not absent entirely, and is still understood by the spectator to be within the diegetic space due to appearing briefly or being acknowledged by a character or an interaction with an object within the film frame.

An example of diegetic music as narration occurs in the film *Samson and Delilah* (Thornton, 2009). Samson watches village residents on his front porch perform music as a live band (guitar and drums); as time passes, first with Samson's interest waning, and then as day becomes night, the music from the live band continues to play offscreen. A shot follows of Samson in close-up on his bed turning on the radio beside him. He lays in bed holding his ears as the live music outside competes with the radio, both offscreen. The radio comes from Samson's right, and the live music from his left. The audience knows this as the music and radio are muffled when he holds both ears; but he lets go of his right ear and hears the radio, then unblocks his left ear to hear the music outside (see Figures 113-114). There is a visual acknowledgement of the live band and the radio emitting sounds within the space, but they emit this sound when they are not necessarily onscreen. Samson, by holding and letting go of his ears, narrates the experience of music within the diegesis by attempting to drown out certain sounds and choose what he, and subsequently the spectator, listens to. These two musical sources cannot be onscreen at the same time so need to exist across the film's montage of space, even though we get a close-up of Samson in bed with neither onscreen. Samson's experimentation with controlling what he listens to envelopes the music within the image of narration.

Figures 113-114. Samson lets go of his right ear to hear the radio (top), and then unblocks his left ear to hear the band play outside (bottom), in *Samson and Delilah*.

There is another occurrence of competing music within the narrative space of the film. The band finishes playing for the night and Delilah sits in a car and inserts a radio cassette of Ana Gabriel performing the romantic Mexican song *Mi Talismán*. We then cut to Samson, who upon hearing that the band has finished performing plugs a microphone into the band's amplifier and places it next to his radio, blasting the electrifying rock song *Warlpiri Woman* by the Indigenous Australian band, Lajamanu Teenage Band, dancing freely. Delilah hears the music from the car and looks over her shoulder to spy on Samson dancing to his music. What is notable about this instance is that the camera cuts between Samson in a wide angle, Delilah in a close-up from within the car, and another shot showing her perspective of Samson from within the car. Initially, Samson's music drowns out Delilah's car radio, grabbing her attention, but then her music slowly rises in volume and complements instead of competes with Samson's music, as she stares beyond the camera, with the soft and radiant orange glow of a campfire kissing her face as she radiates wonder. The music from both Delilah's cassette and Samson's radio are not visually sourced for their entire song durations, yet they are still heard. Matching Samson's earlier encounter, Delilah narrates the musical experience with the camera in her proximity, subjecting her to reverse shots of her perspective of Samson. When offscreen, the music can still be acknowledged existing within the film space through character interaction, but the music can also be heard and therefore narrativised. Delilah swoons and is captivated by Samson's performance, which affects audio-visual space within her car; both Samson's music and Delilah's music, as well as their spaces, are bridged in a romantic connection. [See Figures 115-118]

Marshall

Figures 115-118. Delilah watches Samson dance as both Delilah’s and Samson’s music bridge in *Samson and Delilah*.

The example from *Samson and Delilah* shows that diegetic music can motivate narrative form and narration. On this type of motivation, Claudia Gorbman argues that:

> “Offscreen sound [and music] …typically motivates camera movement and/or cutting to new quadrants of space. As the camera eye searches out the sound source, cinematic space “naturally” unfolds. Diegetic music fleshes out film space, and variables in recording, mixing, and volume levels further determine the quality, the "feel," or framed/lived space in a given film. In films with stereo sound, diegetic music as well as dialogue and sound effects, of course can articulate space with all the more directional precision."[350]

In *Samson and Delilah*, the music forces the shots to cut to the sound source, as Delilah searches for the music source.

In an example from *Rear Window* (Hitchcock, 1954), the camera never leaves Jefferies’ apartment. The film is narrated by Jefferies who is disabled in a wheelchair, and so the camera has no choice but to physically move with him and scan spaces outside his apartment window. A songwriter lives in an apartment to the top-right of Jefferies’ view and apartment, and in the evening of the second day in the film, the songwriter is seen sweeping his apartment floor; every so often he has musical inspiration and experiments with several chords on his piano. The camera in this scene begins from the right and pans left to halt at his apartment. He sweeps, stops to play a music chord on the piano, then continues sweeping. The camera as he resumes sweeping resumes panning left and tilts up to observe the rest of the apartment block. In this instance of camera movement, the same chord is heard to be struck again, however this time it is images of the apartment block being observed from Jefferies’ apartment perspective rather than observing the sound source (see Figures 119-

---

[350] Gorbman, *Unheard Melodies: Narrative Film Music*, 25.

121). The music obviously stops as the songwriter is just experimenting in short increments, however he resumes a few minutes later. This time, the piano is cued on Jefferies asking his girlfriend Lisa, “That would be a terrible job to tackle…just how would you start to cut up a human body?” and continues to play while Jefferies and Lisa observe the soon-to-be-discovered murderer of his wife, Lars Thorwald returning to his apartment with a rope. Ben Winters calls these audio-visual instances ‘musical wallpaper[s]’, where the music within the space is likened to wallpaper in a film scene, that it may lend to characterising the nature and feeling of the space in which the action takes place.[351] This 'wallpaper' analogy derives from Igor Stravinsky, a composer critical of film music.[352] Stravinsky’s description of diegetic music as ‘wallpaper’ is not to diminish its significance, but to acknowledge music’s connection to visual form. Winters continues: "...the music we hear must still be part of the diegesis that is narrated, since it is attributable to the same narrative source."[353] Nondiegetic film music is usually acknowledged in Film Music Studies as the only source affecting narrative drama, the dramaturgy of the nature of the scene as mentioned in Chapter 3, but hearing *any* music lends to the narrative drama, and *knowing* the music is within the space carries a certain meaning to that space.

---

[351] Winters, "Musical Wallpaper?: Towards an Appreciation of Non-narrating Music in Film," 43.
[352] Ibid., 53.
[353] Ibid., 44.

Figures 119-121. The camera pans left from the songwriter visibly playing chords on the piano, but the music is still heard when there are images of the neighbouring apartments in *Rear Window*.

If the source of diegetic music is invisible, and thus offscreen, it is part of the gaze in which the drama is read. In the restaurant scene in *Shaun the Sheep Movie* (Richard Starzak, Mark Burton, 2015), Shaun and his fellow flock of sheep are dressed in disguise as humans. They sit at a table awkwardly trying to fit in, looking intently at their menus while a pianist plays Frederic Chopin’s *Grande valse brillante* on the piano to act as what Ben Winters would regard as wallpaper for the narrated space. However, their disguise is blown when havoc is unleashed. One sheep is chased around the restaurant by a staff member, and they jump on and over the piano, striking random notes as they do so. The pianist comedically shrugs his shoulders and changes the Chopin tune to a more ragtime piece, increasing the pace and changing from the leisurely ¾ waltz time to a pulsating 4/4-time signature. The camera cuts to the action of the sheep being chased by various restaurant staff, and the piano becomes part of the gaze that observes the chaotic energy of the chase action around the restaurant. In this instance, diegetic music, when offscreen but visually sourced, participates in the construction of music-image.

Is diegetic film music actually diegetic?

Up to this point, I have mentioned invisible sources that are visibly established to build the bridge from (visible) performance to invisibility. Such sources are often delayed in their visual revelation, complicating the narrative experience. Sometimes music is audible, assumed rather than acknowledged within the narrative space until it is visually revealed. However, sometimes this point of revelation is unclear. Rick Altman argues that any diegetic sound (after he mentions theme music, assumingly nondiegetic) remains subordinate until it receives a narrative function, requiring its volume to increase and rise above other tracks of less importance.[354] Jeremy Barham indirectly disagrees with this position, stating that understanding the diegesis of a given source of music in a film is "crucial to how a represented world is creatively established, in other words to the construction and experience of a diegesis."[355] For this chapter, Barham's position reinforces the idea that diegetic music can and does contribute to narrative function. Diegetic music, whether visible or not, lends to the construction of the music-image, and understanding its location is crucial to understanding the story-world.

It is not clear why music emanating from, for example, a radio within the film frame should be analysed separately from the rest of the narrative space-time when analysing diegetic film music in a given film.[356] That music emanating from the radio (or any other diegetic object) should be included as part of the larger narrative image that generates meaning. I therefore agree with Andrew Knight-Hill's position that a binary distinction between diegesis and nondiegesis is "[not] appropriate" if all sound, diegetic and/or nondiegetic, constructs diegesis.[357] 'Diegetic', as a term, is applied inconsistently to film, Knight-Hill argues as, for example, there is a "focus on the distinction between the Platonic definition of *telling* and the stratified literary definition derived from Gerard Genette in which different narrative voices operate at various levels of literary diegesis," masking the role music plays in that very construction of diegesis.[358] This is why diegetic film music is a neglected field in contrast to studies of nondiegetic film music. However, as has been argued in this chapter, diegetic music is intrinsic to narrative and the semiotic systems operating within the film, affecting the spectator's sense of time, space, and tone.

---

[354] Rick Altman, "24-Track Narrative? Robert Altman's Nashville," *Cinémas* 1, no. 3 (1991): 113-14.

[355] Jeremy Barham, "Music and the moving image," in *Aesthetics of Music: Musicological Perspectives*, ed. Stephen Downes (New York: Routledge, 2014), 225.

Knight-Hill argues that diegesis only exists in experiencing film as an audio-visual phenomenon: "If the reading of mimesis [and diegetic] lies not in the characteristics of the sound itself but within the listener this requires a transition from striated and distanced observation to the smooth engaged and material experience of the audio-visual phenomenon."[359] While Knight-Hill uses the word 'subjectivity' in his audio-visual analysis, aligning the concept of diegesis with the spectator's unique encounter, in chapter 2 I use the paired notions of ‘experience’ and ‘semiotic ambiguity’ to conceptualise the film image. In Knight-Hill’s as in my model of spectatorship, the spectator 'accepts' that film cannot be analysed “in isolation from its audience.”[360]

---

[356] Take, for example, Jeremy Barham's analysis of the contrary claims of film music by David Bordwell and Kristin Thompson, and Roy Mottram in separate chapters of the same book, *Film Sound: Theory and Practice*. See, ibid., 226. David Bordwell and Kristin Thompson in their chapter 'Fundamental Aesthetics of Sound in Cinema' describe nondiegetic music as mood music that has "no relation to the space of the story." See, David Bordwell, and Kristin Thompson, "Fundamental Aesthetics of Sound in the Cinema," in *Film Sound: Theory and Practice*, ed. Elisabeth Weis, and John Belton (New York: Columbia University Press, 1985), 199. Roy Mottram in his chapter 'American Sound Films, 1926–1930' argues that "music is part of the whole fabric of expression, more expressive than words alone, and far more than merely a background for the dramatic action. It actually becomes a second level on which the narrative develops meaning." See, Ron Mottram, "American Sound Films, 1926–1930," ibid., 223. The two chapters may not be in direct disagreement, but it is striking to see that two such contrasting analyses of film music's relation to diegesis appear in the same book.

[357] Knight-Hill 652

[358] Knight-Hill, "Sonic Diegesis: Reality and the Expressive Potential of Sound in Narrative Film," 652.

[359] Ibid., 653.

Film requires the activity of spectatorial perception, as the image circulates through a continuum of spectatorial experience. Sound, an intrinsic part of film form, is similarly not pure, and thus is intrinsic to film’s expressive potential. The spectator assumes the music is offscreen only through prior knowledge. What does this knowledge do to film experience? In a literal sense, it enlarges the world, expanding it beyond the frame, taking form within the spectator’s cognitive and affective engagement. When diegetic music challenges the ontology of narrative space-time, is all film music nondiegetic in the sense that it is involved in the construction of an image, that it always relates to the spectator who exists outside the diegetic world, in some way? In the Claudia Gorbman sense of located/sourced music within the narrative space, film music is not always diegetic, but film music is and can be diegetic in the music-image sense as it aids in the construction of diegetic image, even though it is not necessarily sourced. For example, if there is a cut to a shot of a red rose, and the camera slowly tracks in with romantic string music heard in the cue of this shot, the music can be said to be cued because of this rose; the music attaches to and in some sense explicates the meaning of the rose. However, there is clearly no music emitting from the rose, as it has no speakers on the surface of its petals. 'Construction of space' and ‘sourced within the space' can both mean 'diegetic', but not in terms of the binary definition of diegesis as it is commonly understood. This complex operation of diegetic music complicates the relationship between music and spectatorial experience. Since the music-image relies on spectatorial experience for its existence and functionality, there are no boundaries to defining diegesis, and the hard line between diegetic and nondiegetic weakens. Many filmmakers have played with this dilemma comedically, such as Mel Brooks playful diegetic experiments in *Blazing Saddles*, explored in Chapter 1. However, what if the frame containing the red rose cropped out the presence of a radio beside it? The easiest way to know whether it is 'diegetic' is if it is acknowledged by inhabitants within the space. This example proves that the diegetic status of film music can *change* at any moment in a film depending on how the spectator understands the scene. The contingency of locating the source of film music creates a problematic analysis of 'where' the film music is. Attention needs to focus on the *experience* of the film at the given moment, that is, including screen time in which the music is not visually sourced. When film music has a delayed visual reveal within the diegesis, the analysis is closer to that of nondiegetic film music, which interprets the construction of the film image to generate narrative meaning.

---

[360] Ibid., 654.

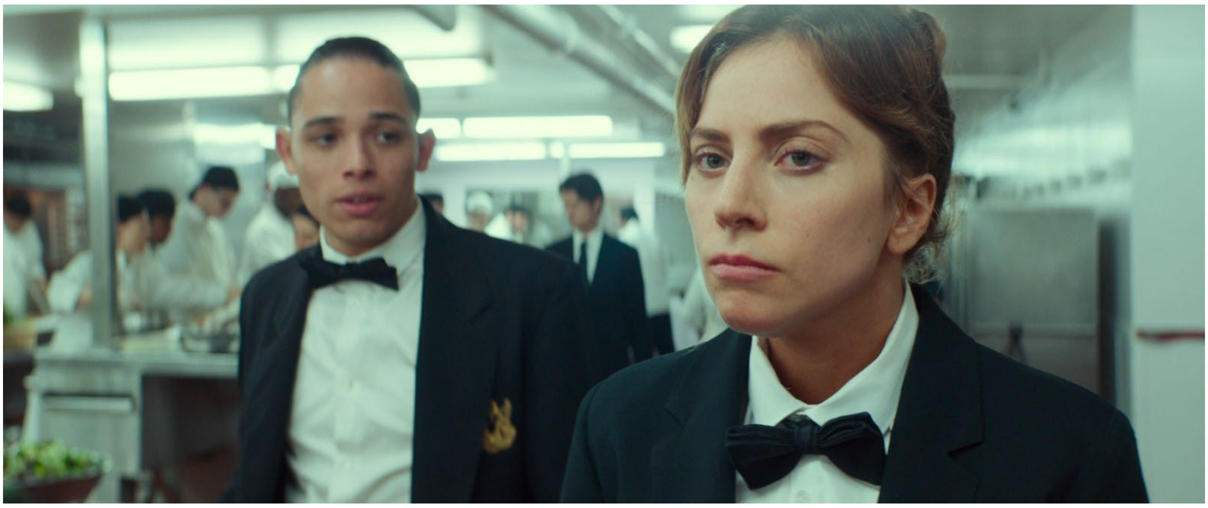
One of the most common ways for diegetic music to have a delayed visual reveal, from invisible to visible, is within the edit, and usually in transitioning scenes. In postproduction of filmmaking, a very popular sound edit is the J-cut. The J-cut is a technique used by sound editors to "bring the sound [and music] earlier than the images…to avoid straight cuts and this softens the transition between shots" and "to draw our attention to the next shot."[361] This technique can be used for dialogue, ambient sound, sound effects, and film music. An example of this in the most recent version of *A Star is Born* (Bradley Cooper, 2018) occurs when Ally quits her job to meet Jack backstage at his live concert. As soon as Ally's boss in the kitchen tells her that she is late to work, a fast, clicking rhythm from drumsticks is heard. The anticipation and drive for Ally to go to the concert with her passively aggressive reaction to her boss in close-up is constructed in this image. The camera cuts to behind her so that her face is not seen while the music plays, ultimately constituting the expression of her decision.

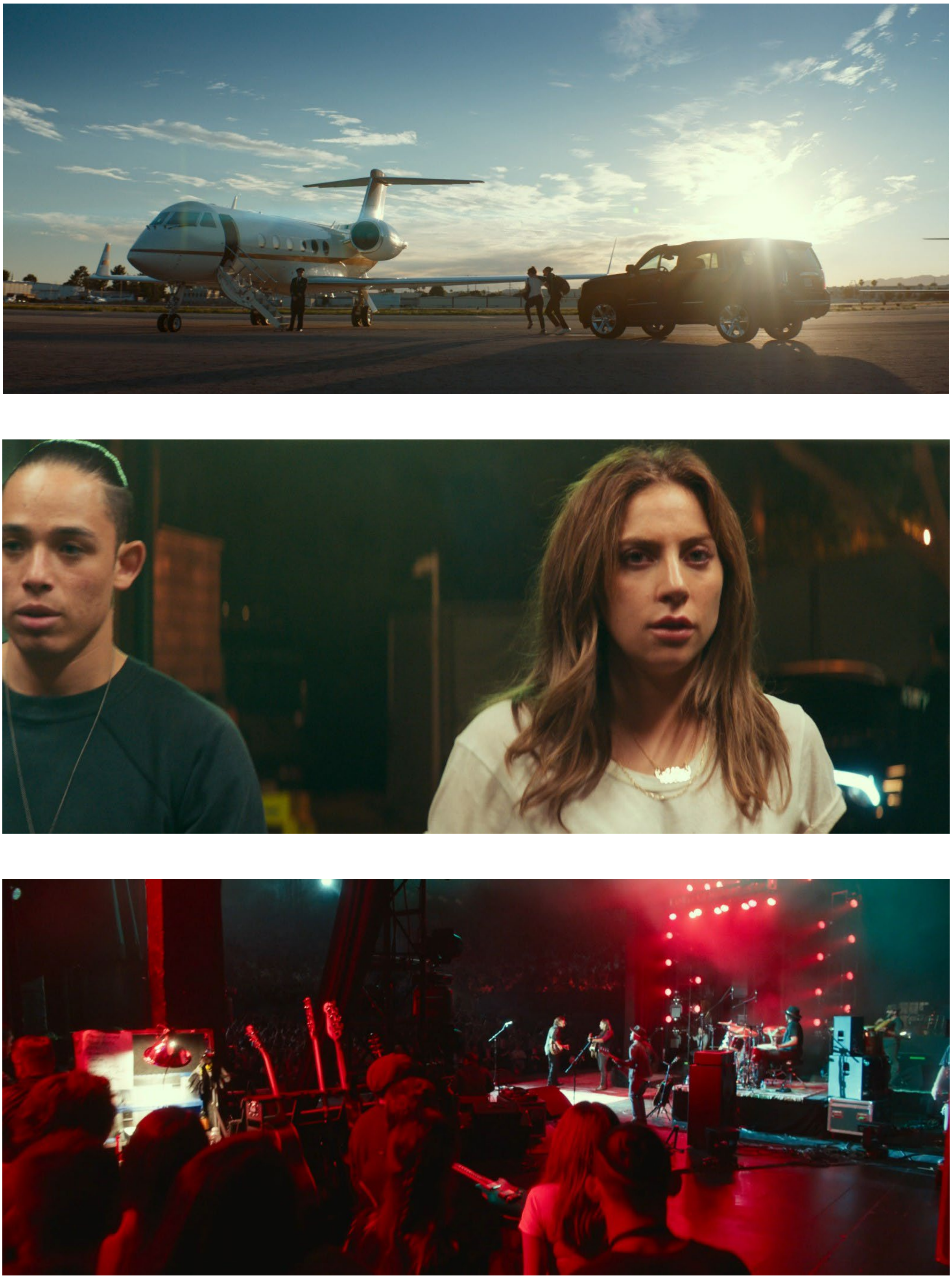
---

[361] Pablo Romero-Fresco, *Accessible Filmmaking: Integrating translation and accessibility into the filmmaking process* (London: Routledge, 2019), 204.

Ramon enters stage right and asks, “are we doing this?” to which Ally turns her head as an acoustic guitar responds to the rhythmic patterning of the drumsticks with a fast ostinato in the lower register. When the guitar cues, Ally laughs as if to acknowledge the music cue in prompting her decision. Following a series of cuts across the ’nondiegetic’ music, we now hear cheers from an invisible crowd, opening and extending the audio-visual space. Throughout this sequence, the music narratively orchestrates Ally’s journey to the source of the music, while the rhythm of the musical beat affects and structures the frenetic energy of the editing and action.

The *A Star is Born* example is a very simple sequence narratively, but a complex one temporally. That is, it is difficult to comprehend Ally taking the short duration of the song being performed (and still being performed once she arrives) to quit her job, ride to the airport, fly in a plane, then arrive onstage (see Figures 122-127). This journey would have taken hours. However, the music elides the conventional temporality of the journey to reconfigure it as an expanded filmic space-time. In the transition from invisibility to visibility, there is a rearrangement of time and space, but this is only discovered through the diegetic acknowledgement of the music.





Figures 122-127. While a song plays, Ally quits her job, travels with Ramon by car, plane, and arrives onstage to the source of the music, in *A Star is Born* (2018).

The film is a stage

In *The Broadway Melody* (Beaumont, 1929), sisters Hank and Queenie are slowly becoming Broadway stage musical stars. However, chaos ensues as both fall in love with the same man on the road to success. The film's narrative undoubtedly revolves around music (even the title of the film confirms this), but most of the film takes place off-stage, backstage. Up until this point in the chapter, I have mentioned examples of diegetic music as performance or offscreen but have reserved narrative-focused diegetic music until now. As Rick Altman points out, there are many films which include a "significant amount of diegetic music, and yet we will refuse to define as musicals…."[362] A film musical, to Rick Altman, is a very different kind of filmmaking, one that can be easily understood by the cinemagoer in terms of what they expect to experience, and easily understood by the filmmaker as what the studio wants.[363] The American film industry describes a musical as a film with music at its centre, but the conception of a music-image I've proposed in this thesis would conclude this of any given film with music, as music is intrinsic to the meaning and effect of the film image.

In the case of the musical as a genre, I would argue that the film must be explicitly narratively driven by music; it needs to affect the fabula and exist within the immediate space-time of the diegesis. The example of *The Broadway Melody* has music at its narrative centre, but most of the action happens off-stage, without performed music; I would therefore classify *The Broadway Melody* as a 'back-stage musical'. This genre of film narrative derives from the theatre, pre-dating film. When sound was first introduced to cinema in the 1920s (as seen with *The Jazz Singer* earlier in the chapter), most films situated music at the narrative centre, *The Broadway Melody* being one of them.[364] My task in this section is not to attempt to define and explain the genre of the musical, but to situate the musical as a genre form that explicitly engages the structure and itinerary of music-images.

---

[362] Altman, *The American Film Musical*, 14.
[363] Ibid., 12.
[364] Ibid., 131-32. Rick Altman provides a list of the 'Fairy Tale Musical', the 'Show Musical', and the 'Folk Musical' at the Appendix of his book *The American Film* Musical. See, ibid., 371-78.

Backstage musicals derive from the stage. When music is performed, it is in the orchestra pit, while the narrative comes alive on stage. Nina Penner states that when there is music and dancing on screen, it is referred to as a musical 'number'.[365] In the case of the backstage musical, whenever there is a musical performance on stage, this is a number, as if watching a musical on Broadway. Although *The Broadway Melody* was a rare example of a backstage musical released merely months prior to the events of The Great Depression, the backstage musical is a *product* of the Great Depression in the 1930s with locations very similar to the musical theatre, and that of the cinema, and the sets on stage are usually aesthetically expressive, with a live show or concert of a performing orchestra or live band.[366] This is essentially the way in which filmmakers first conceptualised how to incorporate music into films. The backstage musical, which birthed the wider genre of the film musical, became the staple for the diegesis.

The backstage musical showcases a self-aware film narrative. It is self-aware because the narrative would not exist without diegetic music emitting from the orchestra pit. In a typical musical theatre, there is a stage, an orchestra pit, and the interior of the theatre which includes the audience seats. The camera is predominantly situated where the audience would be seated in the theatre. Take, for example, the film *Yankee Doodle Dandy* (Curtiz, 1942), tracking the life of entertainer and performer George Cohen. Whenever there is an onstage performance in the film, the camera is positioned as if sitting in a seat in the audience (unless the narrative is offstage), as if the camera is an audience member whose eyes are panning with the narrative action or following the singers and dancers as they perform (see Figures 128-131).[367] *Yankee Doodle Dandy* could easily exist as a stage musical or a film, but audiences instead do not necessarily have to attend the physical space of the theatre. Since the physical stage and the orchestra is visible to the audience, the audience accepts it as a performance, even if the music is for a shot or two offscreen; that is, the music bleeds into the image. Placing the camera in an audience seat creates the illusion of a theatrical performance.

---

[365] Nina Penner, "Rethinking the Diegetic/Nondiegetic Distinction in the Film Musical," *Music and the Moving Image* 10, no. 3 (2017): 3.

[366] Altman, *The American Film Musical*, 61.

---

[367] In some of the onstage performances in *Yankee Doodle Dandy*, the camera begins in the position of an audience seat, but sometimes cuts into a medium shot of the onstage action or the camera moves on a crane from the audience seat closer to the stage. The camera, however, always faces the stage and is never physically positioned on the stage.

legraph
2378
RD

THE
FOUR
COHANS

Figures 128-131. During onstage performances in *Yankee Doodle Dandy*, the camera remains positioned where the theatre audience would be seated.

In theatre, there is a diegetic membrane between screen/orchestra pit and the audience. The stage is the diegetic world while the orchestra pit, offstage, accompanies the action nondiegetically, as the narrative does not acknowledge its presence. However, film can change this binary arrangement of diegetic/nondiegetic music. Jeff Smith points out that:

> “Whereas film theorists have long ago acknowledged the aesthetic and formal possibilities of offscreen sound, music tends to be interpreted through a rather different lens by the viewers. Because the use of nondiegetic score is so common in

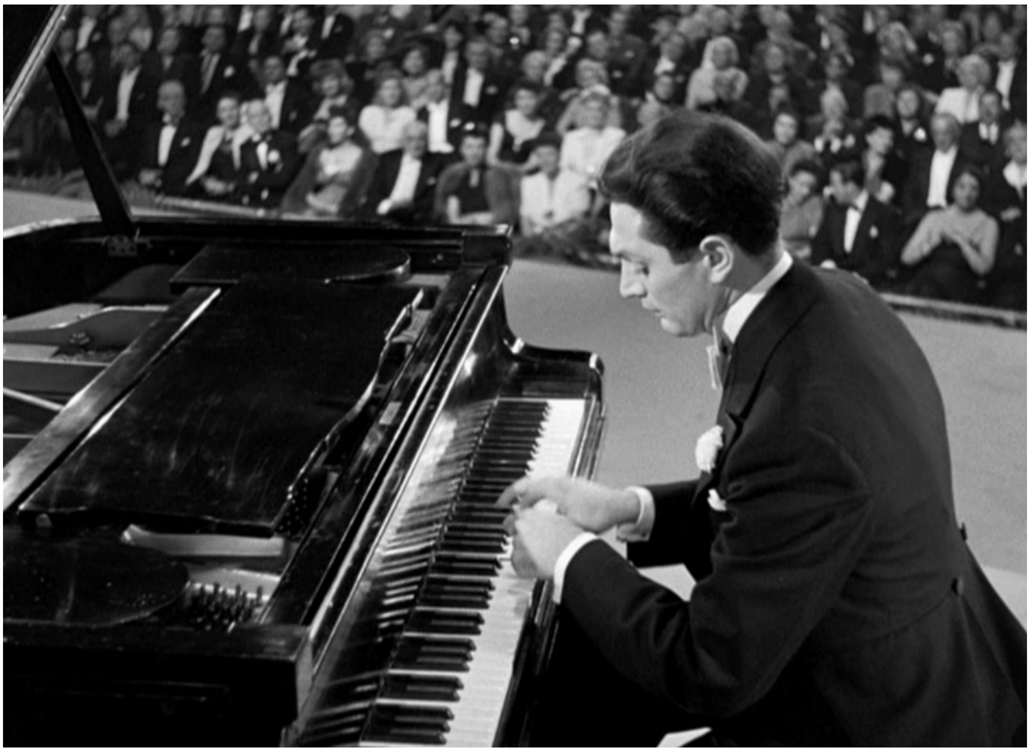
> fiction filmmaking, one usually assumes that any music not clearly located in the story world is likely to be nondiegetic."[368]

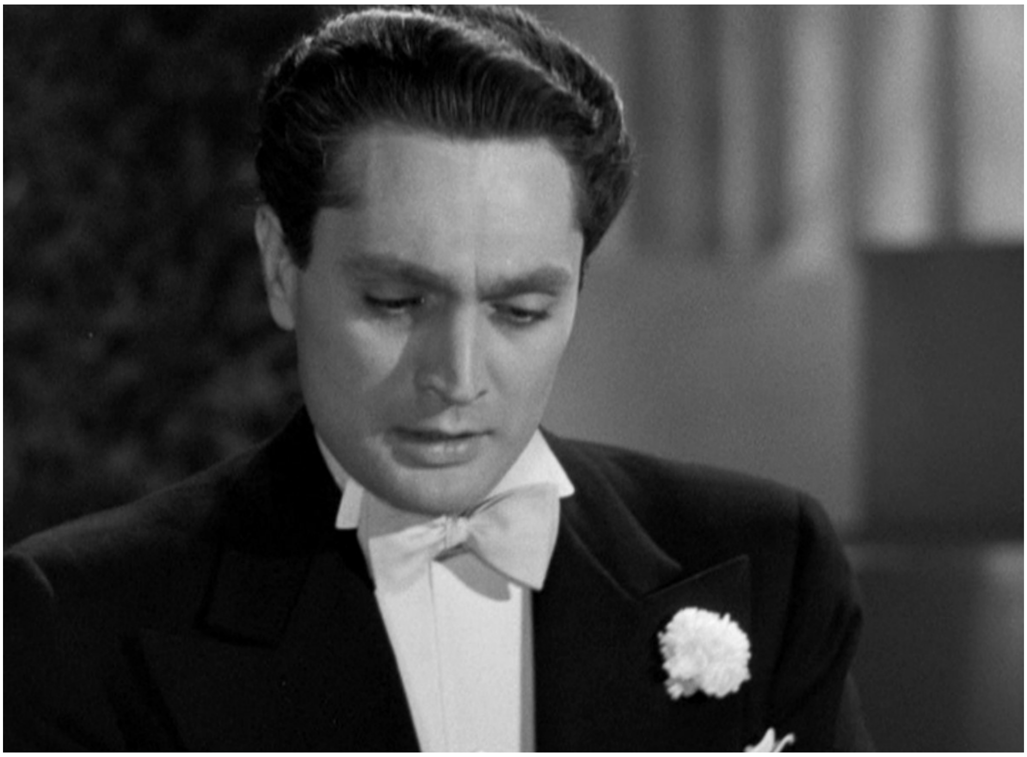
Smith's observation is critical; in theatre, music offstage functions nondiegetically. However, in film, the diegetic membrane is the film screen, and everything within the camera space is diegetic, making the orchestra pit part of the diegesis. The stage performance existing within the film makes the music part of the narrative. Returning to the examples of *The Broadway Melody* and *Yankee Doodle Dandy*, a shift in narrative focus occurs, causing the stage to become part of the larger world of the film existing offstage. In theatre, on the other hand, the narrative can never leave the stage. When experiencing the music in this instance, the music shifts from nondiegetic to diegetic in the form of an image that has 'permission' to leave the stage. Cinema abounds with this slippage between diegetic and nondiegetic space.

Since the theatre stage exists within the space of a film, the camera is allowed to move away from the seating area of the theatre. The camera can shift its gaze to other vantage points in the audience, backstage, outside the theatre complex, or onstage. Placing the camera onstage shifts the narrative focus from performance to performer. The seating area becomes a point of interest or concern (depending on the narrative situation) for the performer. In *Rhapsody in Blue* (Rapper, 1945), a biopic about the life of American composer George Gershwin, there are several stage performances to showcase the musical talent of Gershwin. Most of these performances (including Gershwin's most famous work, *Rhapsody in Blue*, which shares the title of the film) are shot from the theatre audience's perspective. In the performance of his Concerto in F near the end of the film, the camera occasionally moves onto the stage either to present a close-up of Gershwin playing the piano or a rear view of the audience (see Figures 132-134). In the narrative, at this point in his life, Gershwin suffers from a brain tumour that affects his playing; the shift of the camera therefore provides the point of view (and emotional import) of friends and family in the audience. Moving the camera onstage allows a shift in filmic address from passive observation of performance to the intimate life of the performer.

---

[368] Smith, "Bridging the Gap: Reconsidering the Border between Diegetic and Nondiegetic Music," 14.



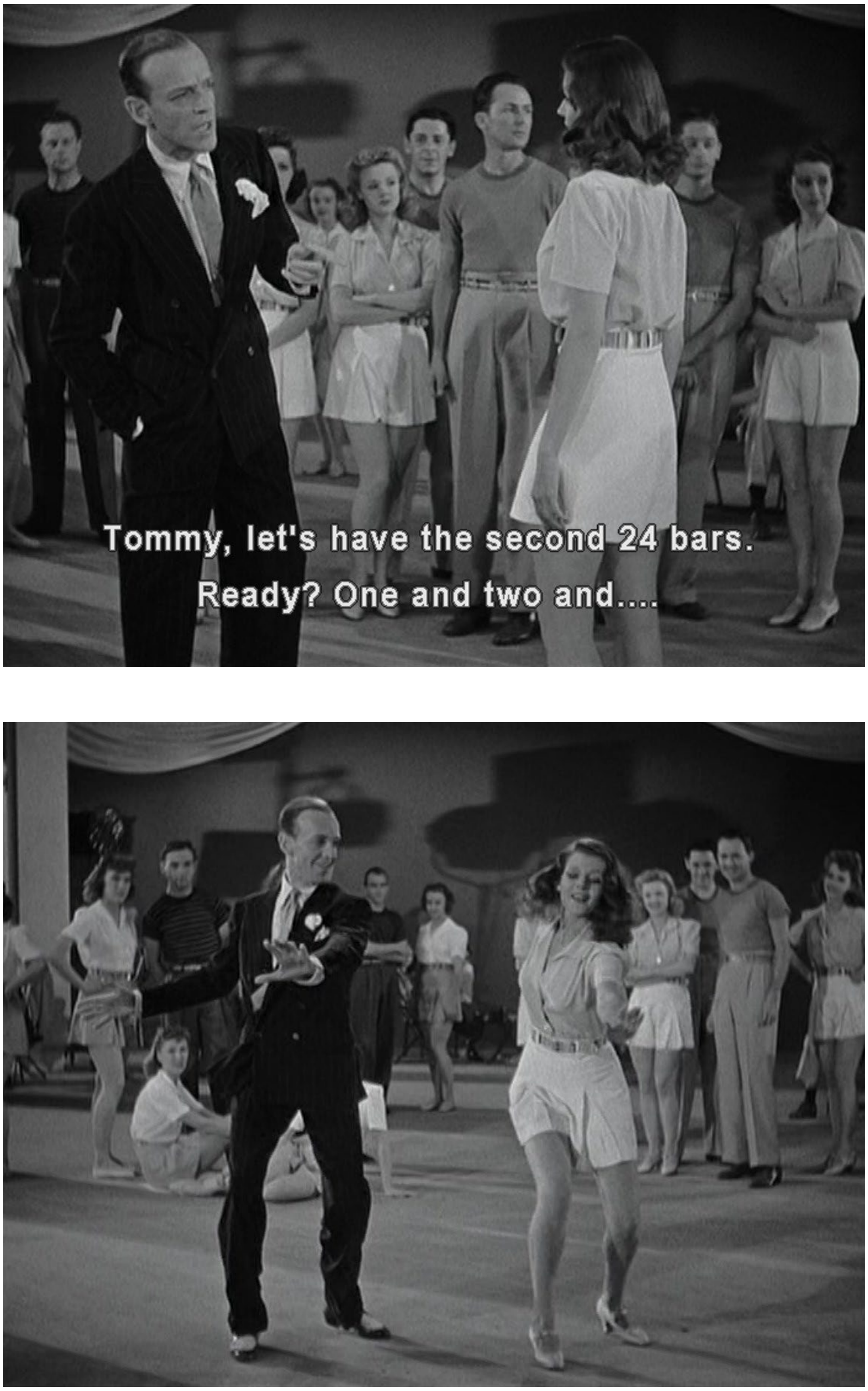


Figures 132-134. The camera in *Rhapsody in Blue* is able to move from the position of the theatre audience to onstage.

To return to the notion of musical performance as ‘numbers’, Nina Penner explains that, even offscreen, the music during a number is considered diegetic with the onscreen performance.[369] When the camera is allowed to leave its theatre seat and move wherever it wants to go, the diegetic music is still a reference point to its source, whether that be from the orchestra pit, onstage, or outside the theatre, even offscreen. For film experience, this reference point works within the image created. For example, in an opening scene of *You’ll Never Get Rich* (Lanfield, 1941), when Robert (Fred Astaire) dances with Sheila (Rita Hayworth) during a rehearsal of the number "Boogie Barcarolle", Robert asks the piano accompanist offscreen “Tommy, let’s have the second 24 bars!” The source of the music is not visible, but it has been acknowledged by Robert within the space. The acknowledgment therefore requires the spectator to infer the dance as one channel of performance and meaning (see Figures 135-136).

---

[369] Penner, "Rethinking the Diegetic/Nondiegetic Distinction in the Film Musical," 3.

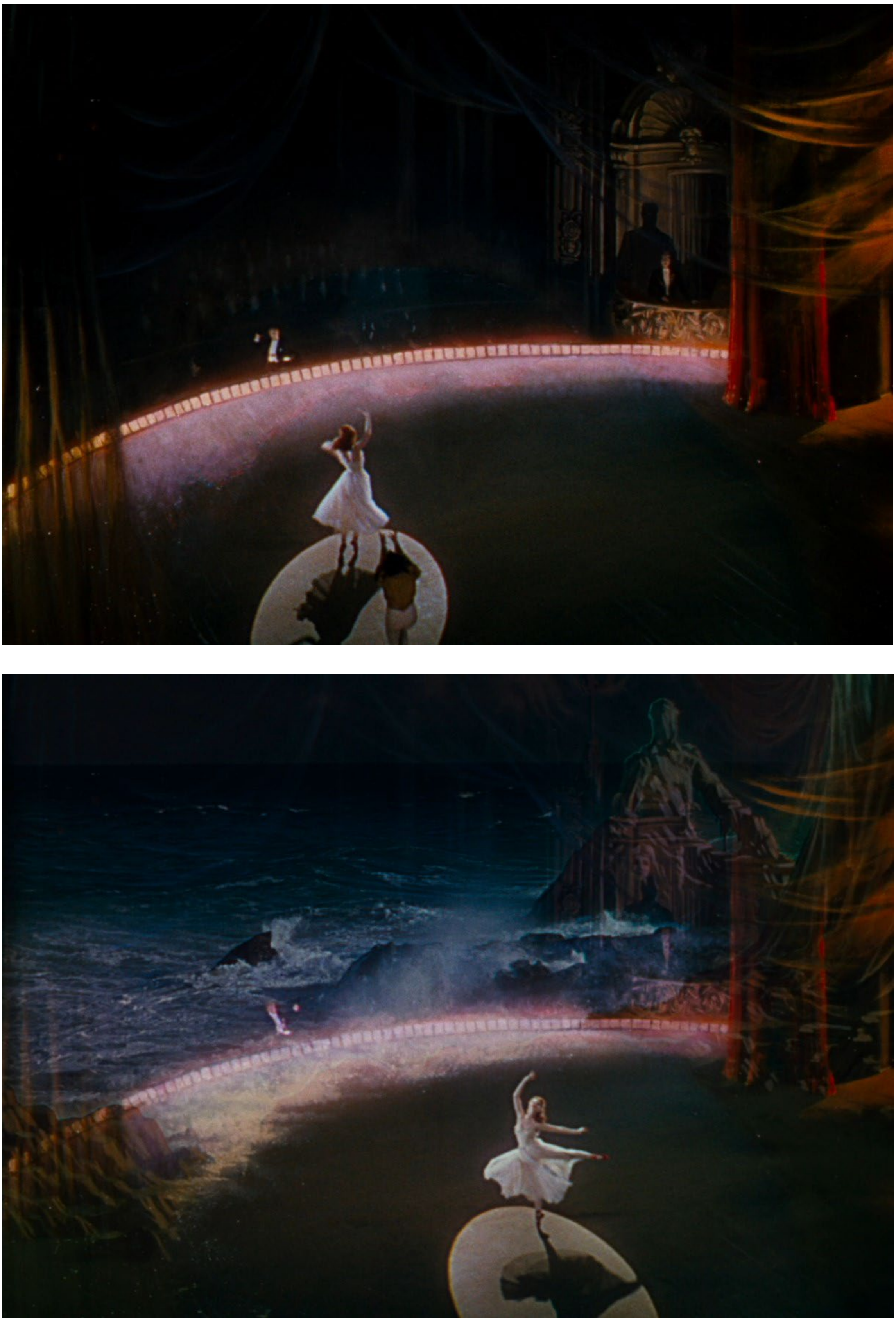

Tommy, let's have the second 24 bars.
Ready? One and two and....

Figures 135-136. Music is acknowledged offscreen while Robert and Sheila dance in *You'll Never Get Rich*.

In the central ballet sequence in *The Red Shoes* (Powell and Pressburger, 1948), a "fourteen-minute surreal confection of movement, music, and colour which is allowed to play out uninterrupted" provides the camera exclusive access to the performance.[370] In the sequence, the camera shifts to a brief extreme high angle shot from the rear corner of the stage, facing an audience silhouetted from the stage floodlights. In this peculiar shot, the music transitions from romantic to sombre as muted brass and double bass infiltrate the musical register as Vicky (the central protagonist and dancer) is released from her dance partner to dance alone. The silhouetted audience then dissolves into the imagery of turbulent waves crashing against a rocky shoreline (see Figures 137-138). The timbre of the music in the brass and strings is matched with the harrowing tone of the surreal imagery. Vicky dances as if on a cliff next to this image while onstage. The sounds of crashing waves dissolves to the sound of applause as Vicky dances offstage. The music has brought the audience into the narrative, and into the ballet which is part of the stage. Furthermore, shifting the camera from the viewpoint of the audience to the viewpoint of the stage has brought the audience into the narrative. In this sense, we can say that the diegetic world of film is an infinite stage. There are many examples of onstage music effecting an offstage narrative, including *A Star is Born* and *The Red Shoes*. Others include *Straight Outta Compton* (F. Gary Gray, 2015), a biopic which explores the origins of the Black American rap group N.W.A, and *Sing Street* (John Carney, 2016), a coming-of-age tale of an Irish schoolboy forming a school band to impress the love of his life who lives across the road from the school. Both films, as well as others I have discussed, explore narratives in which diegetic stage music invades offstage narrative.

---

[370] Linda Ruth Williams, "The shock of The Red Shoes," *New Review of Film and Television Studies* 15, no. 1 (2017): 10.



Figures 137-138. In an extreme high angle shot from the rear of the stage, the audience dissolves into imagery of a turbulent seaside, in *The Red Shoes*.

Sourceless musical sources

In *Cleo from 5 to 7* (Varda, 1962), Cleo is awaiting her medical results and a potential diagnosis of cancer, so the film tracks two hours of her life from 5pm to 7pm in the evening. In the beginning of the film, Cleo has tarot cards read to her to predict her future, and the fortune teller tells her she will die of cancer. Cleo is a professional singer, and after the first half of the film's runtime filled with anxiety and self-hatred towards her image, she goes home after a lengthy taxi journey for singing practice with a piano accompanist. *Cleo from 5 to 7* is part of the French New Wave (though considered under the sub-category Left Bank as the 'true' New Wave filmmakers were from French film magazine *Cahiers du cinéma*) in the 1960s. The film employed many artistic nuances from the new wave such as "location shooting, the creation of dead time…spatial and temporal realism [such as the film focusing on Cleo for 2 hours] …jump cuts…long takes…and mobile framing."[371] Agnes Varda is clearly interested in exploring how the camera relates to Cleo and the space she inhabits. When Cleo, roughly halfway through the film after anxiously awaiting her medical results and attempting to ward off the pain by staring at her beauty in mirrors, rehearses a song about death with her lyricist the film completely shifts gear. In this musical sequence, the camera begins in a mid-shot of Bob (the pianist) at the piano, Maurice (the lyricist) leaning on the back of the piano, and Cleo at the side of the piano singing as she reads the lyrics, though furthest from the camera. The film explores its mobile framing by swinging around to the left around Bob towards Cleo, leaving her alone in frame. The camera stays focused on her face, zooming into a close-up. All that is heard in this instance is Cleo's voice and the piano. Once the edge of the film frame is limited to the tip of her shoulders, a stringed orchestra begin to accompany her and the piano, yet there is no stringed orchestra in sight – visually, it is assumed Cleo cannot hear the string accompaniment. However, as the strings cue in, Cleo takes her eyes off the lyrics and addresses the camera. The background is blacked out; it appears as if Cleo is no longer in the room. In this sequence, from what source does the music emanate? The spectator is aware of the piano in the room (shown in a previous shot), but where are the string instruments? In this instance of staring into the camera, Cleo hears and is accompanied by music that is not in her space; thus, the diegetic music is sourceless. [See Figures 139-142]

---

[371] Kelley Conway, "Cultivating the New Wave Spectator: Cleo from 5 to 7," in *Agnes Varda* (Urbana; Chicago; Springfield: University of Illinois Press, 2015), 39.

Figures 139-142. The camera swings around the piano into a close-up of Cleo's face, who addresses the camera, in *Cleo from 5 to 7*.

What happens if the musical number leaves the theatre stage? Once the camera has overcome the barrier of the orchestra pit between the stage and the audience, anywhere in the film diegesis becomes the stage. When the musical number is situated within the theatre space, the orchestra pit and music is diegetic, even to the performers. When the stage leaves the theatre, the orchestra is also left behind. Where, then, does the music in the musical number come from? In such instances, again we see the traditional film frame overcoming the diegetic membrane of the orchestra pit, moving closer to the stage, and situating the orchestra within the film frame. In the musical number which occurs outside the theatre, the film world is no longer screened from the spectator. The character that hears and reacts to external music references cinema itself, collapsing the barrier between the screened 'narrative' world, the world of the apparatus, and the world of the spectator.[372] Since the stage is now the film image, the orchestra pit is now within the screen; in the illusion of being source music, the music is sourceless. The source is the film image itself, devoid of narrational or diegetic content; the musical in this sense can be described as a 'pure' film genre.

Idealism as a diegetic bridge

Rick Altman suggests that the barrier between screen world and spectator world commonly breaks down during a musical number when characters idealise something or someone.[373] On the theatre stage, a character or set of characters may turn to the musical number as a source of imagining an ideal scenario or imagining their thoughts. These characters would turn to the orchestra pit for help in the pursuit for the ideal. However, inside the narrative space, the orchestra does not exist, much like nondiegetic music in a film. When the stage leaves the theatre, this dichotomous relationship still occurs in the world/stage and nondiegetic/orchestra. The character(s) resort to the nondiegetic/orchestra to engage in the ideal/imaginary. Since the nondiegetic/orchestra is part of the film/theatre, and the character(s) can hear the music, the character(s) are resorting to film/theatre, breaking the barrier between film/theatre and world/stage. The spectator implicitly acknowledges that the character(s) can hear music if they dance, sing, and react to its presence, even if that music is sourceless.

---

[372] See Chapter 2 for an extended breakdown of the music-image automaton, which includes the spectator in the cinema apparatus.

[373] Altman, *The American Film Musical*, 63.

Emily Petermann notes that at face value musical numbers blend nondiegetic and diegetic sound.[374] For this blending to be seamless in film experience, musical numbers must take charge of dialogue. Petermann continues: "The often fluid transition from dialogue to musical number marks not a change in diegesis, but another mode of expression within the story world - musical numbers are instrumental in furthering the plot yet serve different function from dialogue."[375] Musical numbers take care of dialogue by incorporating its function within the fabula. Altman calls this the ‘audio dissolve’ technique used in musicals, which super imposes sounds to pass from one sound to another.[376] Petermann, referencing Altman, further suggests that when there is no orchestra present there needs to be a bridging device between realism and idealism, speech and song or dance and music.[377] In *West Side Story* (Spielberg, 2021), when Tony leaves the dance and learns of Maria’s name, he mutters “Maria” to himself, then incorporates ‘Maria’ into song lyrics when music appears: “The most beautiful sound I’ve ever heard…Maria.” Tony’s internal monologue of idealising his love for Maria has turned into song. In order to transition smoothly from the narrative world to the ideal in song, there needs to be this seamless audio dissolve to synchronise sources.[378] Ben Winters suggests that spectators have “little difficulty in believing that music occupies the same narrative space as the characters, whether or not a source is visible.”[379] There is little difficulty in believing, for example, that Tony can hear stringed music in the side streets of Upper West Side Manhattan that perfectly matches the harmonies of his voice. Experimental musicals such as *The Umbrellas of Cherbourg* (Demy, 1964) and *Les Misérables* (Hooper, 2012) present believable dialogue as fully sung.[380]

---

[374] Emily Petermann, "The Film Musical as a Subject for Word and Music Studies," in *Silence and Absence in Literature and Music*, ed. Werner Wolf and Walter Bernhart, Word and Music Studies (Leiden; Boston: Brill, 2016), 223.

[375] Ibid.

[376] Altman, *The American Film Musical*, 63.

[377] Petermann, "The Film Musical as a Subject for Word and Music Studies," 227-28.

It is important to note when the synchronisation of narrative world and a world idealised by a character or characters within that narrative world occurs. Emilio Audissino states that the narrative level of a film, where these two worlds collide, is about storytelling.[381] A musical number needs to have a narrative motivation if it is to use the film world as a stage. Emily Petermann points out that many critics ‘disparage’ musicals for their lack of realism.[382] What Petermann means by this is that any regular person cannot sing or dance like a professional, and so audiences have trouble believing characters breaking into song and dance when they envision an ideal scenario.

‘Integrated’ musicals have songs that are not ornamentations (decorative add-ons) but are fully ‘integrated’ in the film plot.[383] In order for the songs to be integrated into the plot in a more believable manner, film plots can centre on show business, “which requires little in the way of a transition between narrative and number: the performance of songs fits within the realistic frame of the story if characters are professional singers and dancers.”[384] Some examples of these integrated musicals include *Gentlemen Prefer Blondes* (Hawks, 1953) and *Singin’ in the Rain* (Kelly and Donen, 1952).

---

[378] Altman, *The American Film Musical*, 65.

[379] Winters, "The Non-Diegetic Fallacy: Film, Music, and Narrative Space," 229.

[380] In a recent scientific study of how the human brain experiences speech and melody, the results show the human brain balances speech and music in different auditory regions. By applying these findings to the operation of listening to a song with words and music, the brain can decode the information of written word and music simultaneously, crossing the content of each in a semantic and semiotic balance of meaning. This allows in a musical number for dialogue to be incorporated in song if it is balanced in the music. See, Philippe Albouy, Lucas Benjamin, Benjamin Morillon, and Robert J Zatorre, "Distinct Sensitivity to Spectrotemporal Modulation Supports Brain Asymmetry for Speech and Melody," *Science (American Association for the Advancement of Science)* 367, no. 6481 (2020): 1043. There may be the question: Is it a musical if no one sings? Singing to music is the finest acknowledgement of music heard within the narrative space. How do you acknowledge sourceless music if it is not visible? Sing and dance. The body is used as a tool of listening and reacting. It is a narrative affect, per se.

[381] Audissino, *Film/Music Analysis: A Film Studies Approach*, 78.

[382] Petermann, "The Film Musical as a Subject for Word and Music Studies," 227.

[383] Ibid., 224.

When the characters are not in show business, the musical requires another device to plausibly link dialogue with song, or bodily movement with dance. Petermann notes:

> “...the lack of realism in characters singing and dancing means that what happens in the numbers is generally in some kind of an imaginary sphere – whether the fantasy world of performance, of dreams and desires, of memory, of hypnosis, or set in a world that is a fairy-tale domain altogether.”[385]

The sourceless music allows the non-musical speech to be free of the constraints of everyday life, representing an escape from reality into the Richard Dyer idea of utopia: “a non-place, the imaginary one of many possible worlds created by the characters’ imaginings.”[386] Examples of this could be: Dorothy, an innocent country girl imagining a technicolor fantasy world ‘somewhere over the rainbow’ in *The Wizard of Oz* (Fleming, 1939); and Maria, a governess summoning the hills to be alive with music through a series of sweeping wide shots of the rolling hillside surrounding Salzburg in the opening of *The Sound of Music* (Wise, 1965). The idealised utopia, whether imagined or referenced, bridges nondiegetic music with the diegetic film world, replacing the orchestra of the musical theatre, creating a sourceless diegetic musical source within a created film image.

---

[384] Ibid., 227.

[385] Ibid.

[386] Ibid., 235.

La La Land (Chazelle, 2016)

La La Land (Chazelle, 2016) is a film musical that encompasses all the elements of diegetic and nondiegetic musical form discussed in this chapter. *La La Land* draws on classical and modern musical forms and approaches, and contains both musical numbers and diegetic music that does not necessarily conform to the genre. Indeed, the film contains moments of diegetic performance, invisible sources, the film stage, and sourceless musical sources, providing a rich expression of the possibilities in composition of what I have called cinema's music-image.

The musical numbers in *La La Land* are about ideal scenarios: falling in love while pursuing career dreams. These dreams help create what Rick Altman refers to as the dual-focus narrative through two central characters, Mia and Sebastian (Seb). Mia aspires to be a Hollywood film actress while Seb wants to open his own jazz club in Los Angeles. They come together by chance, fall in love, and ultimately choose pursuing their dreams over their love for one another. This idea of the dual-focus narrative is not new to *La La Land*, according to Altman, but it is wrong to limit this style of narrative exclusively to the film musical.[387] The pursuit of dreams in the film requires Mia and Seb to dream, literally, with musicals numbers entering a dreamscape. Ultimately, Mia and Seb are both aware of the process of filmmaking, narratively and self-reflexively (aware of the camera looking at them). Mia and Seb turn to film to make sense of their dreams, which consumes their will to act. *La La Land* bridges nondiegetic and diegetic music through the dreamscapes of idealism.

The opening musical number, 'Another Day of Sun,' takes place on a Los Angeles freeway, blending diegetic and nondiegetic music; with its image-sound composition, the sequence immediately introduces the spectator to the musical genre. The camera moves throughout the space curious as to the music sources (seemingly) in a single shot. In the Blu-Ray commentary for the film, cinematographer Linus Sandgren remarks, “The camera could move more than the dancers…another part of the dance crew that is dancing,” implicitly suggesting that the music orchestrates both the diegetic world and the image that constructs that image.[388] The dancers and the camera respond to music that is not visible, yet begins from a car radio, to transition diegetic music to sourceless idealism. The number ends where it begins, cars honking in a traffic jam, suggesting the number, which involves everyday LA residents, is “in some kind of imaginary sphere….”[389] On the topic of the imaginary sphere, at the end of the sourceless musical number (with a brief moment of a diegetic band), the camera tilts up and then back down. In the bottom of the frame, the two cars at the beginning of the shot (before the tilt up and down) are not the same ones we see at the end of the shot; the cars have now changed. These new cars are Mia’s and Seb’s (the two main characters in the film). [See Figures 143-145]

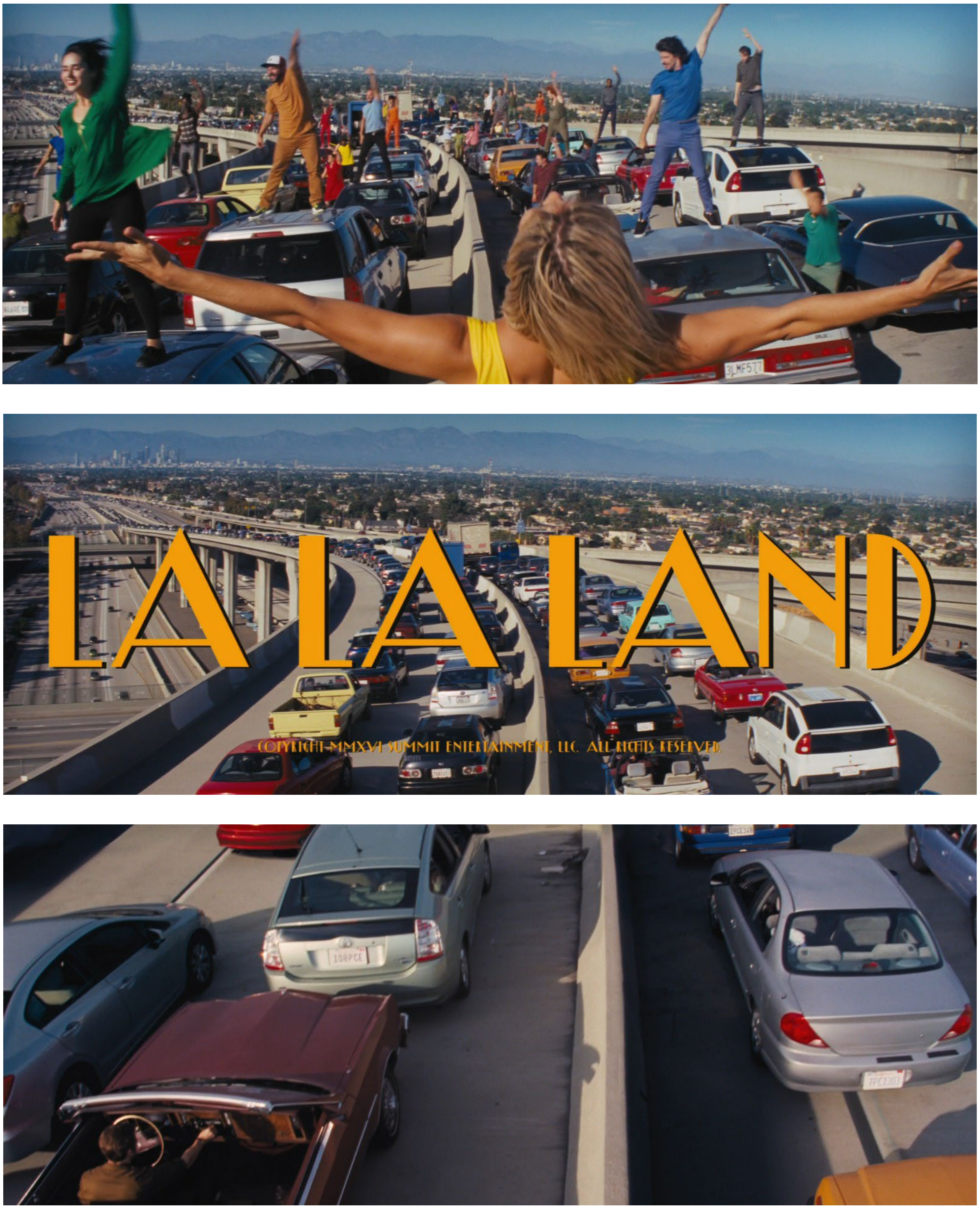
---

[387] See Altman, *The American Film Musical as Dual-Focus* Narrative, Altman, *The American Film Musical*, 16-27.

---

[388] Damien Chazelle, "La La Land," (Australia: Summit Entertainment, 2017).

[389] Petermann, "The Film Musical as a Subject for Word and Music Studies," 227.


LA LA LAND
COPYRIGHT MMXVI SUMMIT ENTERTAINMENT, LLC. ALL RIGHTS RESERVED.

Figures 143-145. At the ending of the ‘Another Day of Sun’ sequence, the camera tilts up and then back down discovering Mia and Seb’s cars which were not there previously, in *La La Land*.

Another musical number in the film which uses a sourceless musical source is the Planetarium sequence. Following a screening of *Rebel Without a Cause* (Ray, 1955), the main characters are inspired to visit LA’s Griffith Observatory where a scene from the film takes place, which triggers a dreamlike sequence. The space of the planetarium is devoid of diegetic sound; all that can be heard is nondiegetic music. Mia and Seb begin to dance to the music's 3/4 time, with a pendulum contraption in middle frame acting as a metronome in one bird’s eye shot (see Figure 146). Mia and Seb end up in the planetarium room, look directly at the camera and float into the clouds among the stars and planets (see Figures 147-149). Their look into the camera connects the film production and the film world, to express the ‘out of [their] world’ of their narrative idealism.

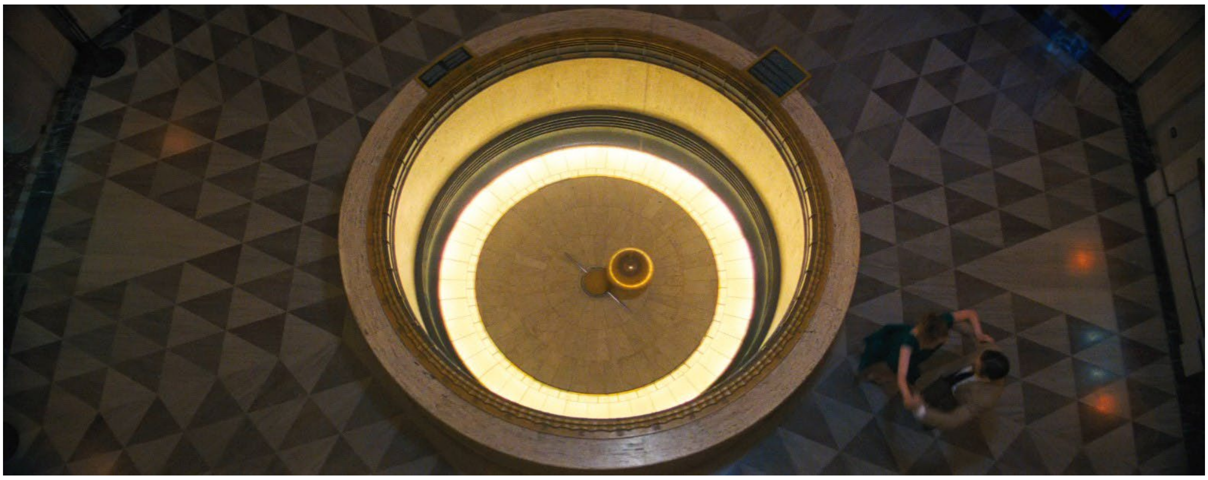
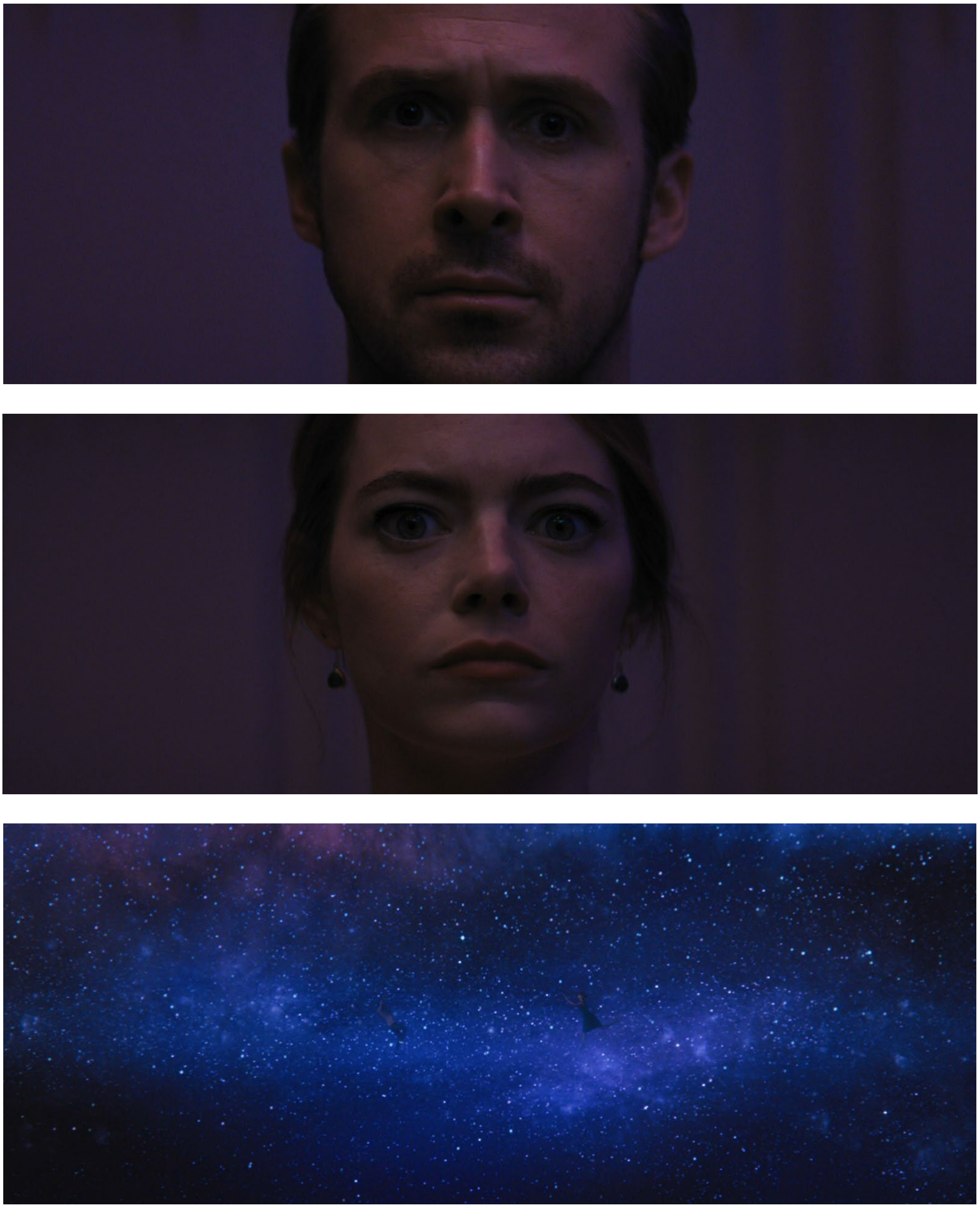


Figure 146. Mia and Seb dance with a pendulum contraption in middle frame, in *La La Land*.



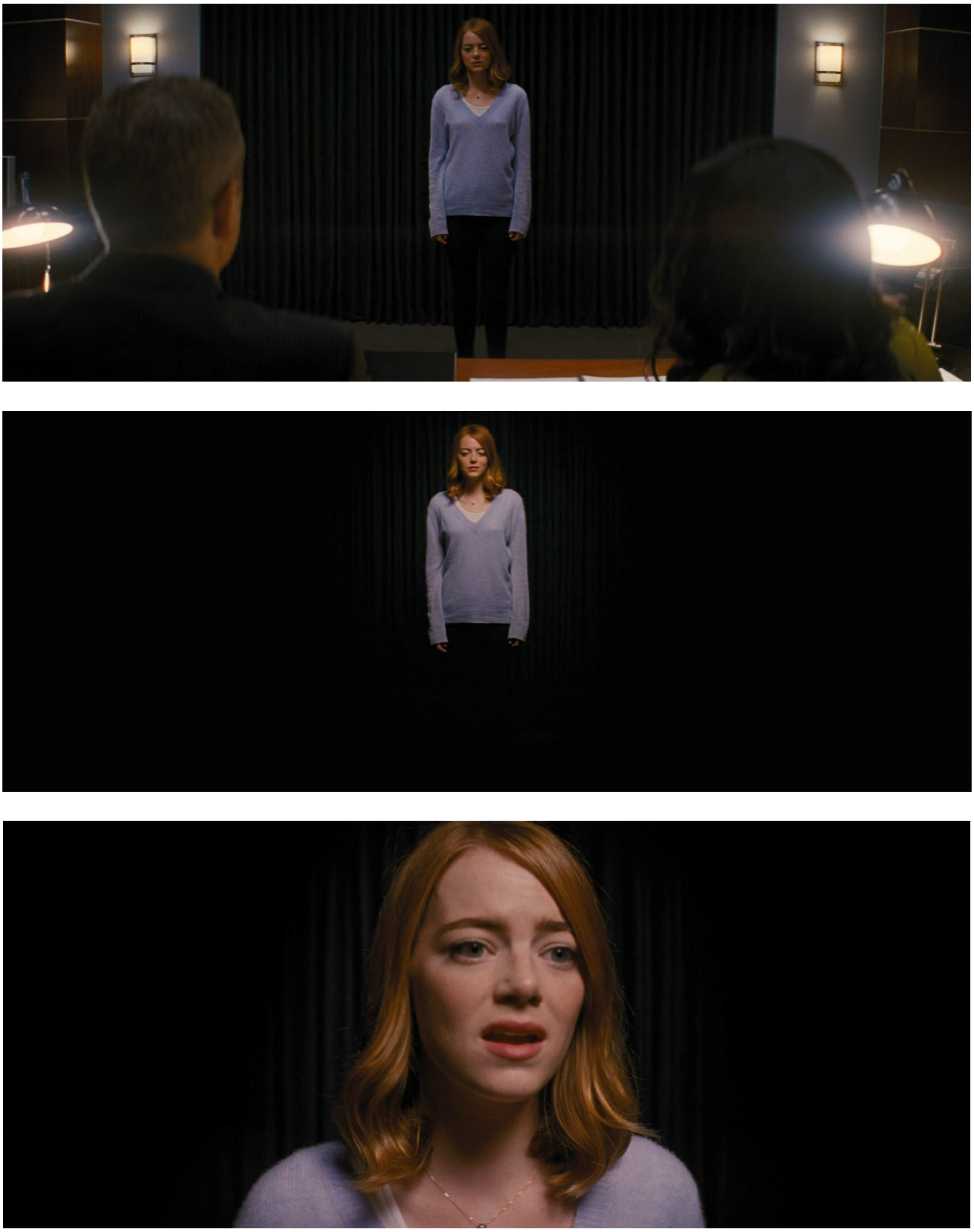
Figures 147-149. Seb and Mia look directly into the camera before floating among stars and planets, in *La La Land*.

La La Land's 'Start a Fire' concert is a live performance with diegetic music throughout the sequence. The music does not become sourceless because this situation is realistic in the sense that it is not ideal for either Seb or Mia, as Seb's commitment to his band ultimately ends their relationship. This is an example of an invisible source as the camera cuts between the perspectives of Seb on stage and Mia in the crowd of people in the audience while the music is known to be sourced within the space. The sequence is not to showcase the live performance, but to inform the beginning of the 'fall' in their relationship, as the sequence immediately following is the 'Fall' (autumn) sequence. In the 'Fall' sequence, Mia and Seb realise their dreams are getting in the way of their love for each other. When Seb shouts at Mia "You're an actress, what are you talking about!" after Mia says, "Since when did you care about being liked?" the vinyl player stops playing the record.

Another invisible source example from the film is when Seb takes Mia to the Lighthouse Café to share what a jazz performance looks and sounds like. The camera cuts between the performers on stage, and Seb and Mia. However, it is a performance framed for Mia and Seb, as after a few close-up shots of different instruments performing, the camera tracks back from the stage to behind Seb and Mia who are watching on. The 'Audition' sequence is the closest the film comes to an audio-visual performance. Mia begins in a wide shot, where the camera is positioned behind casting agents in an office, then (as the lights dim) slowly tracks into her face as Mia turns her monologue audition into song, transitioning from the world of real to the world of the ideal (see Figures 150-152).[390]

---

[390] The 'Audition' song performed during Mia's audition for a breakthrough film role is an ode to those who have dreams.



Figures 150-152. Mia begins in a wide shot in the audition room before the lights dim and the camera tracks around her face, in *La La Land*.

Conclusion

When music is located within the film world, it has a complex, rich and multi-faceted function and series of effects. When music is completely visible for its musical duration, it serves as a performance, between spectator and musician/source. When the musical source is visibly identified and cuts to a shot that does *not* visualise the sound source, the image necessarily contains narrative intention or consequence. The history of the stage musical has impacted the way in which music is presented within film worlds. With the absence of the physical audience and the pit orchestra (which forms the diegetic barrier to the stage and narrative world), the screen acts as a diegetic barrier. When the stage becomes the world of the film, unbounded from the theatre space, the possibilities of music's spatiotemporal location to generate meaning is unrestricted and boundless.

Chapter 7: The Breakdown of the Diegesis

In Chapter 2, I note that the diegesis holds the music-image together. I also contend that the spatiotemporal parameter of film music (its location in space and time) only coheres within the spectator's experience of film. In this process of experiencing a film, if the spectator closes their eyes, only aural stimulus is received, preventing the spatiotemporal location of the film music within the film to be known and accessed as experience. In Chapter 6, I examined the development of diegetic film music as it becomes, increasingly, *source-less*, and the implications of the phenomenon of source-less music within the diegetic space-time. In so doing, I challenged the binary between diegetic and nondiegetic film music, film spatio-temporalities, and filmic experience. Essentially, I ask what happens when film music cannot be placed strictly in either of these domains of the diegesis or the nondiegesis. Returning to Deleuze's taxonomies and terminologies for understanding how film images work, I argue that film music is able to float freely within the spatiotemporal realms of the music-image. Deleuze's critical formulation of his perceived crisis of the action-image – which I would expand to incorporate the intrinsic function of music as the music-image – puts our existing definitions and understandings of diegesis at odds with filmic experience.[391] Using the music-image as the basis of narrative construction, in this chapter I trace particular ways in which films break down film music diegesis. More specifically, I attempt to understand the way in which film music can contain a nonbinary spatiotemporal quality within the diegesis. In the first section, I explore how the music-image can break down, allowing diegetic and nondiegetic film music to exist outside restrictive spatiotemporal definitions and converge, using an example from *The Conversation* (Coppola 1974). In the second section I explore the way in which film music can freely float within the spatiotemporal parameters of the film in the construction of the film image. However, I want to stress that the concept of the nonbinary diegesis I formulate in this chapter does not apply to every music-image construction (indeed, most music images infer a conventional diegesis/nondiegesis binary), but the subjective experience of film presents the nonbinary diegesis as a radical form of potential in the construction of the film image.

---

[391] The phrase 'crisis of the action-image' in cinema is used by Gilles Deleuze in Cinema 1. See Gilles Deleuze, *Cinema 1: The Movement-Image*, trans. Hugh Tomlinson and Barbara Habberjam (London: The Athlone Press, 1992), 197-215. Using the concepts of movement and time from Henri Bergson, Deleuze tracks the history of cinema to pinpoint certain cinematic movements that follow or change trends of action. Recalling from Chapter 2, Deleuze defines the three movements of cinematic image as: the perception image, the affection image, and the action image. In other words, a character can look at something (perception), the spectator sees their reaction (affection), and this reaction leads them to act (action). The spectator shares these three movements of human perception, relating to the perception, affection, and action of the image through the spiritual automaton that shifts percepts into concepts.

The breakdown of the music-image

In the analysis of sourceless musical sources I presented in the last chapter, characters (in musicals) inside the narrative space interact with nondiegetic film music. Nondiegetic film music, or underscore, is, as defined in Chapter 3, "supporting music, which the audience hears but the actors *do not*."[392] And yet, in the several examples of film musicals I analyse in Chapter 6, characters do hear the 'supporting music'. Diegetic film music is "music that appears to come from the scene; music the actors [thus, characters] hear."[393] Yet the characters in the musicals I reference in Chapter 6 'hear' music that is commonly defined as nondiegetic. Diegetic film music can be sourceless, breaking down the binary notion of diegetic/nondiegetic film music. Film scholars and musicologists ever since the concept of the diegesis was introduced to Film Music Studies have situated film music either within or outside the narrative spatiotemporal realm. In the music-image model I propose, film functions as an audio-visual image with one channel of meaning in which music and visual image interlock in semiotic ambiguity. It is thus only *within* the music-image that the diegesis can be defined. What I mean by this is that music is understood to be within or outside of the narrative spatiotemporal realm of film due to conceptualising audio and visual stimuli of the image as having narrative space-time parameters. As audio-visual stimuli turn into concept, diegesis is created and defined within film experience. There are instances, however, when the spectatorial experience of the film image results in a free-form spatiotemporal environment. Furthermore, the music-image and its ontological stability suffers a crisis of form and function within what Deleuze calls the crisis of the action image, which broadly refers to the notion that action is broken from intentionality. This crisis occurs when there is no certainty, from either character or spectator, of the ontological *position* (the diegetic situation) of the music.

---

[392] Karlin, *On the Track: A Guide to Contemporary Film Scoring*, 39.
[393] Ibid.

The Conversation demonstrates a radical breakdown of the traditional music-image, resulting in the unchaining of diegetic and nondiegetic film music within their respective spatiotemporal environments. In the final scene, sound-surveillance expert Harry Caul tears apart his apartment in search of a planted sound bug, only to be left trapped in an image of inaction as the end credits roll. In the final shot, Harry returns to his saxophone to play a jazz tune. Musically, “Harry plays [the saxophone] woefully,” as the “diegetic sound of the saxophone merges with the nondiegetic piano.”[394] The sonic ambiguity reflects the larger frame of a film narrative founded on paranoia. Christian Keathley writes that many American films of the late 1960s and early 1970s respond to the Vietnam War and its aftermath in a deeply pessimistic way.[395] Harry Caul is the product of a society that has broken down due to the traumatic experience of war waged by the United States in Vietnam, creating an image of deep ambiguity.[396] In the film, Harry mishears a conversation he is recording for a job which he believes is evidence that the lives of the two people he is eavesdropping on are at risk: “he'd *kill* us if he had the chance.” This traumatises Harry, disabling him from making key decisions and leaving him in social isolation. Sound orchestrates Harry’s life, sitting at the centre of the film’s narrative with Harry as the principal character. After Harry records a conversation in a busy Union Square in San Francisco, he only cares about obtaining the recording of sound, not what he sees. This allows the film to arrange images sonically as seen when Harry is in his workshop the next day imagining the conversation taking place prompted by the recording. Sounds provide the image in the film, but also provide perceptions Harry must act upon. Julio d’Escrivan argues that film sound can organise the visual world interpretatively; however, since Harry mishears the conversation, this leads him and the film into a state of uncertainty due to misinterpretation of image.[397]

---

[394] Juan Chattah, "Background and Analysis of the Score for The Conversation," in *David Shire's The Conversation: A Film Score Guide* (Lanham; Boulder; New York; London: Rowman & Littlefield, 2015), 159.
[395] Christian Keathley, "Trapped in the Affection-Image: Hollywood's Post-Traumatic Cycle (1970-1976)," in *The Last Great American Picture Show: New Hollywood Cinema in the 1970s*, ed. Alexander Horwath and Noel King Thomas Elsaesser (Amsterdam: Amsterdam University Press, 2004), 293.

Every moment leading up to the final scene of the film reads as negative consequence, such as when the assistant (Stett) to Harry’s client (the director) for the job tells Harry, when he refuses to hand over the recording, “someone may get hurt,” followed by Harry noticing Ann and Mark (the conversationalists) in the office building. Harry worries for Ann and Mark’s lives and takes the tapes back to his warehouse. In the dénouement of the film, Harry reimagines the conversation after seeing the news that the director has been killed in a car accident. In this newly imagined version, Harry hears, “he'd kill *us* if he had the chance”, altered from “he'd *kill* us if he had the chance”, and the spectator hears this too.[398] The spectator hears what Harry hears and so shares Harry’s guilt for not hearing the conversation properly the first time. Sound (hearing the conversation a different way) has thus orchestrated the image and the emotion (guilt) that comes with it. Harry and the spectator have been affected by a transition in sound rather than visual form.

---

[396] Ibid., 295.

[397] Julio d'Escrivan, "Sound Art (?) on/in Film," *Organised Sound* 14, no. 1 (2009): 70.

[398] Bruce Isaacs, "Cinema’s Autonomous Image in Michelangelo Antonioni and Francis Ford Coppola," *Sydney Studies in English* 38 (2012): 40.

In the final scene, hoping his paranoia has been put to rest, Harry sits in his apartment playing his saxophone to a recording of a jazz soundtrack played on a vinyl record player. Interestingly, Juan Chattah points out that Harry misinterprets the soundtrack, "failing to recognise the main tune of the music and continuing to solo over it."[399] The phone rings, interrupting his playing, but there is no answer. He resumes playing until the phone rings again and an anonymous voice (though sounding like Stett) telling him he is being listened to: "We know that you know. For your own sake do not get involved any further. We will be listening to you." To provide Harry with proof he is being listened to, his saxophone playing is played back to him over the phone. Before destroying his apartment in attempt to locate the sound bug, the camera circles around Harry in a handheld medium shot of him looking around his apartment, wondering where the bug would be. This is a disorientating camera movement while the nondiegetic piano plays an unsettling, unresolving ostinato that modulates every phrase with a single bass note pulsating in no melodic direction. Like the music, Harry searches aimlessly for the bug. As a sound surveillance professional, Harry feels ashamed for not hearing the conversation correctly and has retreated to his apartment to console himself with his saxophone. He has retreated from the public sphere for the supposedly safe private space of his apartment, but here sound is used against him. His private space (and life) has been exposed to the public, which traumatises him; hearing a recording of *himself* rather than a recording of others places him in danger. In a profession of surveilling others, he has become the victim of surveillance. To try and restore the natural order of his life and profession, he destroys his apartment in an attempt to find the sound bug.[400]

---

[399] Chattah, "Background and Analysis of the Score for The Conversation," 154.

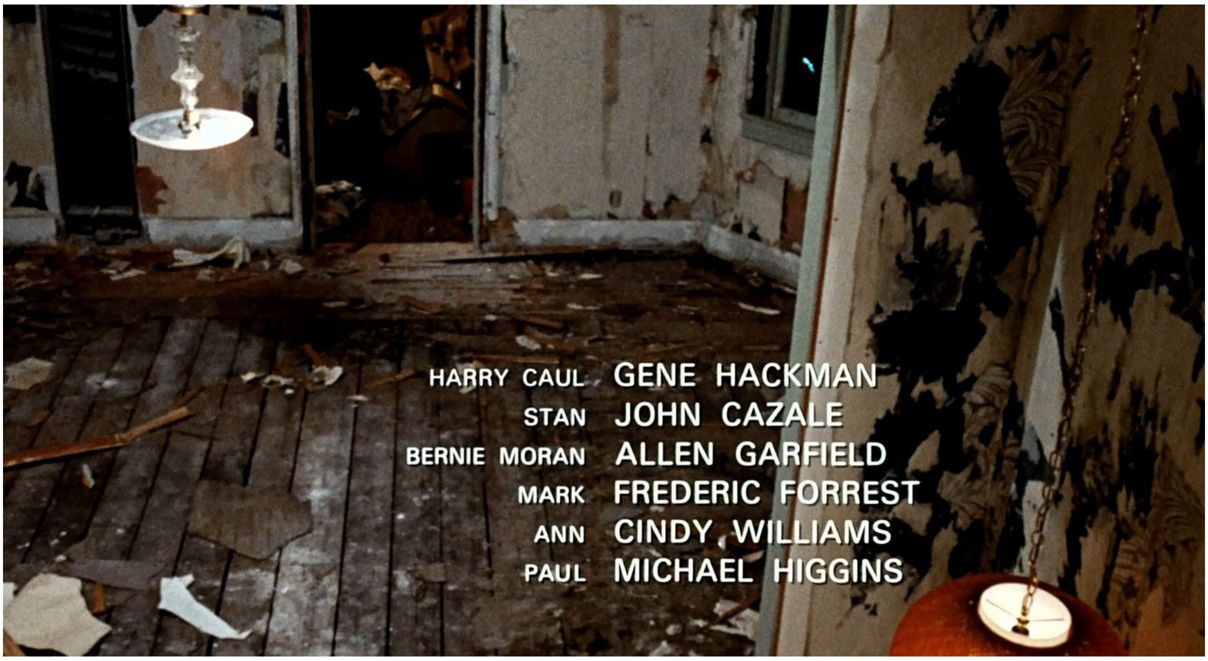
A common question asked about the ending of *The Conversation* is: where is the sound bug/microphone in Harry's apartment? Many answers are possible, but this question only exists in the visual sense if Harry can see and find the bug. In Bruce Isaacs's examination of the same sequence, he concludes with the line: "what does the spectator hear…From what source…does cinema's image emanate?"[401] Concluding on the question of film music's location sidelines film experience. If these questions were reapplied to "what does this do to film experience," the ontology of the film image completely changes. Harry ultimately does not find the bug in the visual sense as he positions himself away from the destruction he caused. Following Isaacs's analysis of the film's final image, which observes the camera in a high angle from the ceiling of the apartment observing the aftermath of Harry's destruction, there is saxophone heard in the image, but Harry is not seen playing just yet. The camera then pans to the left to show Harry sitting and playing his saxophone in the left of the frame, with the destruction occupying most of the frame to the right.[402] The camera, however, pans back to the right, then back again to Harry before the film fades to black. Earlier in the sequence, Harry is accompanying jazz music on the vinyl record player, but in the final shot the record player is absent. Instead, the nondiegetic piano from earlier has joined the music cue of Harry playing. Instead of asking if Harry can hear the piano music and whether he is accompanying it instead of his record player which leads to a dead end, the audio-visual assessment of the shot instead generates an image of surveillance. The high angle and oscillating movement of the camera replicates the function of a surveillance camera, and Harry is surveilled within this footage (see Figures 153-157).[403]

---

[400] Keathley, "Trapped in the Affection-Image: Hollywood's Post-Traumatic Cycle (1970-1976)," 299.
[401] Isaacs, "Cinema's Autonomous Image in Michelangelo Antonioni and Francis Ford Coppola," 44.
[402] Ibid., 42-43.

---

[403] Walter Murch refers to the movement of the final shot in the film as 'oscillating'. See, Michael Ondaatje, "Third Conversation," in *The Conversations: Walter Murch and the Art of Editing* (New York: Alfred A. Knopf, 2008), 174.


HARRY CAUL GENE HACKMAN
STAN JOHN CAZALE
BERNIE MORAN ALLEN GARFIELD
MARK FREDERIC FORREST
ANN CINDY WILLIAMS
PAUL MICHAEL HIGGINS

Figures 153-157. Stills from *The Conversation*'s final image of surveillance as the end credits begin to roll.

In the final image, Harry is now the *victim* of surveillance. Chattah points out that the nondiegetic piano music moves from "percussive attacks manipulated electronically" to reflect Harry's psychological disintegration, to a sudden halt before emerging again as "peaceful and undistorted" in the final shot.[404] The music has moved from atonality to a clear key signature of A minor and clear tempo. This music operates as nondiegetic, not appearing in frame but creating a dramaturgy of nature within the image construction. However, as soon as Harry's saxophone playing cues, the two tracks overlap but do not interfere or disrupt each other. Instead, the two tracks harmonise, even if Harry's playing is "woeful."[405] The overlapping of the two tracks creates one track in the shared harmonies of the music, suggesting that music has reorganised itself into one shared image, making nondiegetic and diegetic music converge.

---

[404] Chattah, "Background and Analysis of the Score for The Conversation," 157-58.
[405] Ibid., 159.

Throughout the film, Harry has been traumatised by sounds, and after destroying his apartment, he continues to hear the nondiegetic piano in the image of never-ending surveillance. Like the characters in the musical numbers explored in Chapter 6, Harry can hear the construction of the film image as the nondiegetic music does not exist inside the space of his apartment. Harry is not just being heard by the sound bug, but by cinema itself. Walter Murch, sound editor of *The Conversation*, acknowledges that Harry Caul's sound surveillance craft is akin to the process of filmmaking.[406] In the final shot of the film, this image construction has come full circle as Harry becomes victim to his own profession.[407] In the transgression of diegetic and nondiegetic space-time, Harry's trauma has broken down the traditional diegetic film image, resulting in a state of radical aural ambiguity. There is no certainty of the diegetic status of the music, allowing the piano and saxophone to harmonise and flow freely within the image.

At the ending of *The Conversation*, Harry Caul is unable to act on the suspicion of being surveilled in his apartment. The inability to act (leading to the Deleuzian crisis in action) leaves Harry in a final shot in which nondiegetic and diegetic music harmonise and converge within a single image form. There is a relation between the music-image and the action-image in that the inability of Harry to act leaves the spectator unable to construct a clear music-image.

The Conversation raises the question of what happens in a situation where the character cannot act? It leaves the character in a state of crisis (as experienced in the film's final scene), but it also leaves the image in a state of crisis, as the image is left in unmotivated, directionless stasis. Deleuze argues that “actions are linked to perceptions, and perceptions develop into actions”; when a character cannot act upon a situation, the link is broken.[408] When percept cannot be linked to concept, the image perceived is in a state of ambiguity, which is precisely what occurs with the intrinsic ambiguity of the music-image in the final sequence of *The Conversation*.

---

[406] Ondaatje, "Third Conversation," 155.

[407] Keathley, "Trapped in the Affection-Image: Hollywood's Post-Traumatic Cycle (1970-1976)," 295.

The crisis of the action-image, and subsequently the breakdown of the music-image, does not occur in any given film or image. David Bordwell, for instance, examines classical Hollywood cinema’s storytelling as “invisible,” suggesting that such an image is never in crisis.[409] In Classical Hollywood cinema, the space-time of the narrative world is stable, and brings little or no attention to itself in telling a story. Deleuze even remarks that in the classical cinema, images are linked by rational cuts, which he argues changes throughout cinema’s history and turns toward movements of irrationality.[410] In the case of *The Conversation*, the filmmaking brings attention to itself through the interaction between diegesis and nondiegesis. The rational sound-image (Deleuze’s causal relation) breaks down, setting the intervals of images free.[411] For this breakdown to occur, Deleuze notes that the crisis of the action-image is determined by narrative:

> “…the crisis which has shaken the action-image has depended on many factors which only had their full effect after the [Second world] war, some of which were social, economic, political, moral and others more internal to art, to literature and to the cinema in particular. We might mention, in no particular order, the war and its consequences, the unsteadiness of the 'American Dream' in all its aspects, the new consciousness of minorities, the rise and inflation of images both in the external world

---

[408] Gilles Deleuze, "On the Movement-Image," in *Negotiations 1972-1990* (New York: Columbia University Press, 1995), 51.
[409] Bordwell, "An Excessively Obvious Cinema," 2.
[410] Deleuze, *Cinema 2: The Time Image*, 277.
[411] Ibid.

> and in people's minds, the influence on the cinema of the new modes of narrative with which literature had experimented, the crisis of Hollywood and its old genres."[412]

For *The Conversation*, Christian Keathley groups the film in the 'post-traumatic cycle' of American film during and following America's involvement in the Vietnam War from the late 1960s to the early 1970s. These films are deeply pessimistic responses to the war with protagonists suffering from post-traumatic stress which disrupts their ability to act, thus disrupting automation of the image. Harry Caul is trapped in the image of affection, and "affection is what occupies the interval [between perception and action]…it surges in the centre of indetermination…between perception which is troubling in certain respects and a hesitant action."[413] It is important to note that humans 'perceive' the world with not just the visual sense, but all senses, including the experience of sound.[414] Since the film involves a narrative of recording and recreating images of sound, Harry is traumatised through his profession:

> "He shifts, runs and becomes animated in vain, the situation he is in outstrips his motor capacities on all sides, and makes him see and hear what is no longer subject to the rules of a response or an action. He records rather than reacts. He is prey to a vision, pursued by it or pursuing it, rather than engaged in an action."[415]

Harry's trauma through recording sound has led him to hear sounds that exist outside his realm of recording. In the final scene of the film, he can hear the nondiegetic piano score, which defies its definition as existing outside his narratively constrained world within the image.

---

[412] Deleuze, *Cinema 1: The Movement-Image*, 206.

[413] Ibid., 65.

[414] Thomas Elsaesser, "(Un)Representing the Real: Seeing Sounds and Hearing Images," in *Psychoanalysis and the Unrepresentable: From Culture to the Clinic*, ed. Agnieszka Piotrowska and Ben Tyrer (London: Routledge, 2016), 69.

[415] Deleuze, *Cinema 2: The Time Image*, 3.

Towards a nonbinary diegesis

The previous analysis of *The Conversation* prompts another, more radical, question: can film music potentially float freely within the construction of the music-image? In Chapter 5, I suggested that the language of film montage is as much cinematic ‘reality’ as music dissonance is in *Jaws*, both of which do not exist contemporaneously. If one is to read the image as one interlocked audio-visual meaning, as I have proposed in this thesis, music is always present to the image, even if not locatable within the diegetic space. The spectator’s relation to film generates endless possibilities of experience. Daniel Yacavone cites Ben Winters in suggesting that “what is often thought of as nondiegetic music 'normally belongs (in our imagination) to the same diegetic realm as the characters; it is part of the story's world, not an invisible means by which the story is narrated.”[416] The approach to the ontology of film music (or what I have called music-images) opens the diegesis to what can be imagined by the spectator.

Towards the end of *Terminator 2: Judgment Day* (Cameron, 1991), the T-1000 and The Terminator are having a final stand-off before the T-1000 gets the upper hand and begins ramming an iron girder into The Terminator (see Figures 158-159). During this 'action' of ramming until The Terminator is temporarily defeated, music is cued with each ram, clearly mimicking the iron rod's movements into the Terminator. The musical cue assists in the construction of the image as the minor-tone to the horn music lowers in register with each ram to suggest The Terminator is helpless and losing strength with each ram of the iron rod. The music is nondiegetic as it assists in the construction of narrative. Why would music be mimicking the movements of the iron girder when it could be channelling narrative in a different way? The fact that it does prompts the spectator to wonder whether the music is operating in some relationship to the metal rod. Does The Terminator feel the sensation of the metal girder with the music or without it? The spectator would conventionally read the music as part of the plot, but would the characters read the plot this way? Do the characters need to hear music to understand what is occurring? For instance, if they did hear the music that constructs the image, the characters would believe an image is being constructed in, around, and of them. This ontological distinction of diegesis could break down the entire film and the characters could rebel against this construction. For instance, if Chrissy heard the E-F shark ostinato at the beginning of *Jaws*, she would have rushed back to shore, there would be no shark attack to open the film, and possibly no film.[417] However, Chrissy cannot rebel against the image construction (including the music) of her shark attack as she is part of that very image construction.

---

[416] Yacavone, "Spaces, Gaps, and Levels: From the Diegetic to the Aesthetic in Film Theory," 23.

---

[417] Audissino, *Film/Music Analysis: A Film Studies Approach*, 33.

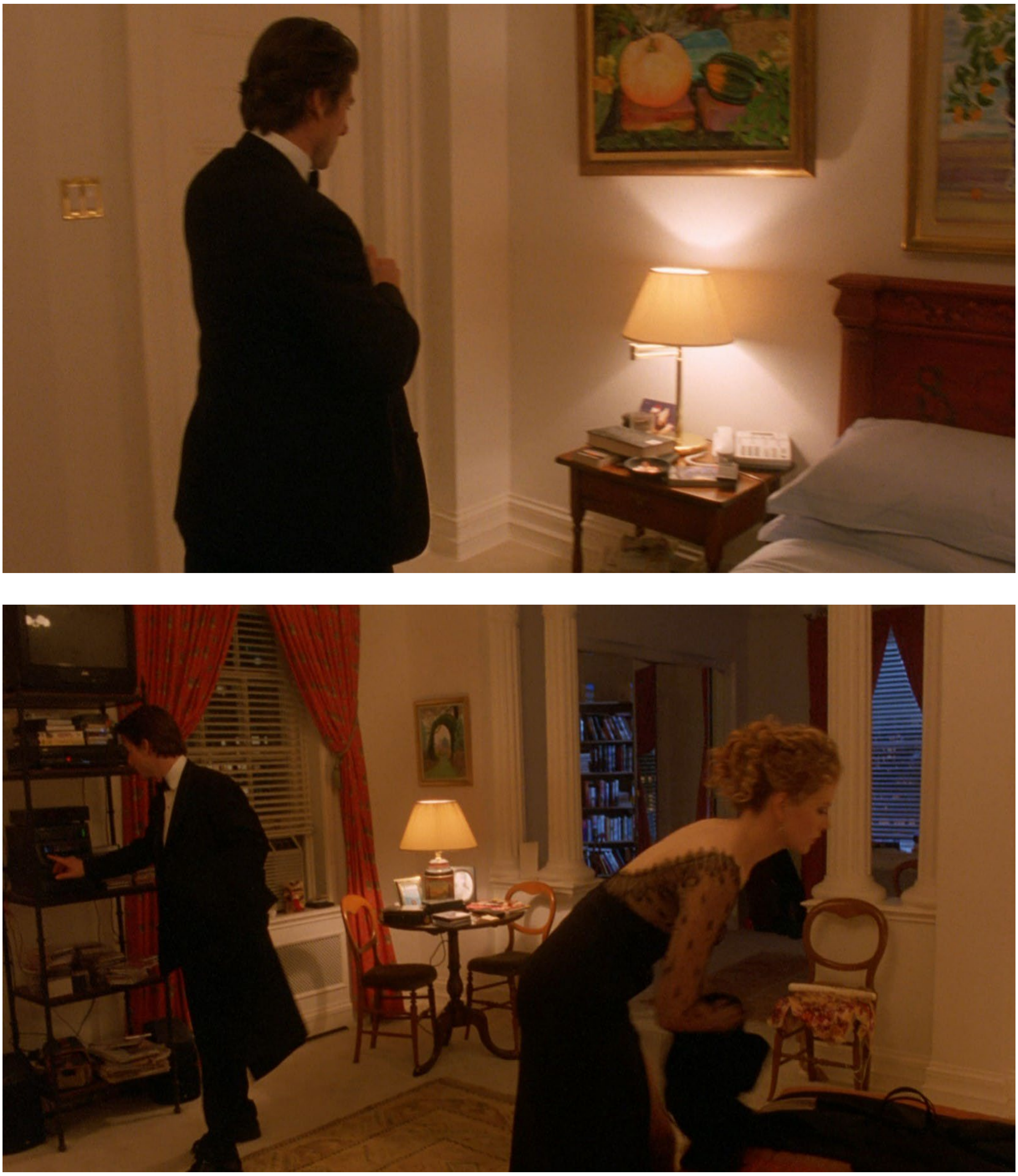
Figures 158-159. The T-1000 rams an iron girder into The Terminator multiple times in *Terminator 2: Judgment Day*.

Film music could be heard by characters in any situation. For example, the reason the comedic gag in *Blazing Saddles* (mentioned in Chapters 1 and 6) works is because the spectator has constructed an image that the character, Burt, does not hear. It is when the live band is revealed that the spectator is shocked. Music sifts through the fabric of the image construction site, allowing characters to hear it or not, but the effect is reliant on the spectator. Giorgio Biancorosso writes: "When music first heard as background scoring ("nondiegetic" music) turns out to be coming from a source inside the fictional world of a film ("diegetic"), the switch may be hilarious, disconcerting, mildly surprising, or all of the above."[418] Biancorosso uses the example of *Bananas* (Woody Allen, 1971). When Woody Allen's character receives a dinner invitation, he initially hears the sound of a harp, only to then discover a harpist playing a harp in his closet. Biancorosso labels these images as 'ambiguous' because the music's diegesis cannot be contained or defined.[419] In *Eyes Wide Shut*, Dmitri Shostakovich's 'Waltz No. 2' plays with the title credits, which is white text on a black screen. This is followed by a shot of a downtown street in Manhattan at night, and Bill and Alice preparing to leave for a Christmas party. The music has all the audio-visual markers to be nondiegetic as it aids in the construction of a festive, waltz-like image of two people in step inside the apartment. However, as they are about to leave for the party, Bill reaches over to the music player to switch it off, and the music stops. The synchronisation of the switch and the abrupt end of the music now suggests that the waltz has been diegetic all along, and that the spectator was deliberately deceived. Bill and Alice seemed to be harmonious, as if partners in a festive waltz, but the diegetic reveal of the music that is switched off abruptly indicates this may not be so. [See Figures 160-163]

---

[418] Giorgio Biancorosso, "The Harpist in the Closet: Film Music as Epistemological Joke," *Music and the Moving Image* 2, no. 3 (2009): 11.
[419] Ibid.

EYES WIDE SHUT



Figures 160-163. Bill and Alice prepare to leave their apartment, then Bill switches off the music player in *Eyes Wide Shut*.

In an excellent reading of the discovery of the monolith in *2001: A Space Odyssey* (Kubrick, 1968), Biancorosso suggests that while the monolith is 'non-realistic', there is nothing to suggest that the music is nondiegetic or closed off from the story world.[420] This is a great insight into experiencing a film that can be applied to any film with nondiegetic music: is all music diegetic until proven otherwise (as nondiegetic)? How does one disprove that music is not playing within the narrative space of the film? In what may be the "most famous anecdote of film music history", there was an argument during the production of *Lifeboat* (Hitchcock, 1944) about whether there should be music in a film set on a boat out in the open ocean.[421] Hitchcock apparently asked, "where is the orchestra?", to which the composer David Raskin responded, "where is the camera?"[422] This argument reveals something deeply problematic in our conventions for thinking about music and visual forms.[423]

It seems the problem in analysing film music has come from this hard distinction between diegetic and nondiegetic in Film Music Studies scholarship. When it comes to analysing what diegetic music does to film experience as opposed to nondiegetic, the scholar is torn in the decision to label the spatiotemporal quality of the film music. The gap between the two realms needs to be brought closer while still acknowledging their differences. This is not to say there is a presupposition that all film music is a blend of diegetic/nondiegetic, but beginning with this thought opens the possibilities for film experience and stories told.

---

[420] Many film music scholars deem diegetic music as realistic, so when Giorgio Biancorosso addresses the realism of the monolith, those scholarly positions are being commented on. See, Winters, "The Non-Diegetic Fallacy: Film, Music, and Narrative Space," 232., for a commentary on Biancorosso's analysis.
[421] Guido Heldt, *Music and Levels of Narration in Film: Steps Across the Border* (Bristol, UK; Chicago, USA: Intellect, 2013), 6.
[422] Robynn J. Stilwell, "The Fantastical Gap between Diegetic and Nondiegetic," in *Beyond the Soundtrack: Representing Music in Cinema*, ed. Daniel Goldmark, Lawrence Kramer, and Richard Leppart (Berkeley; Los Angeles; London: University of California Press, 2007), 188, 201.
[423] Ibid., 201.

The transdiegetic music-image

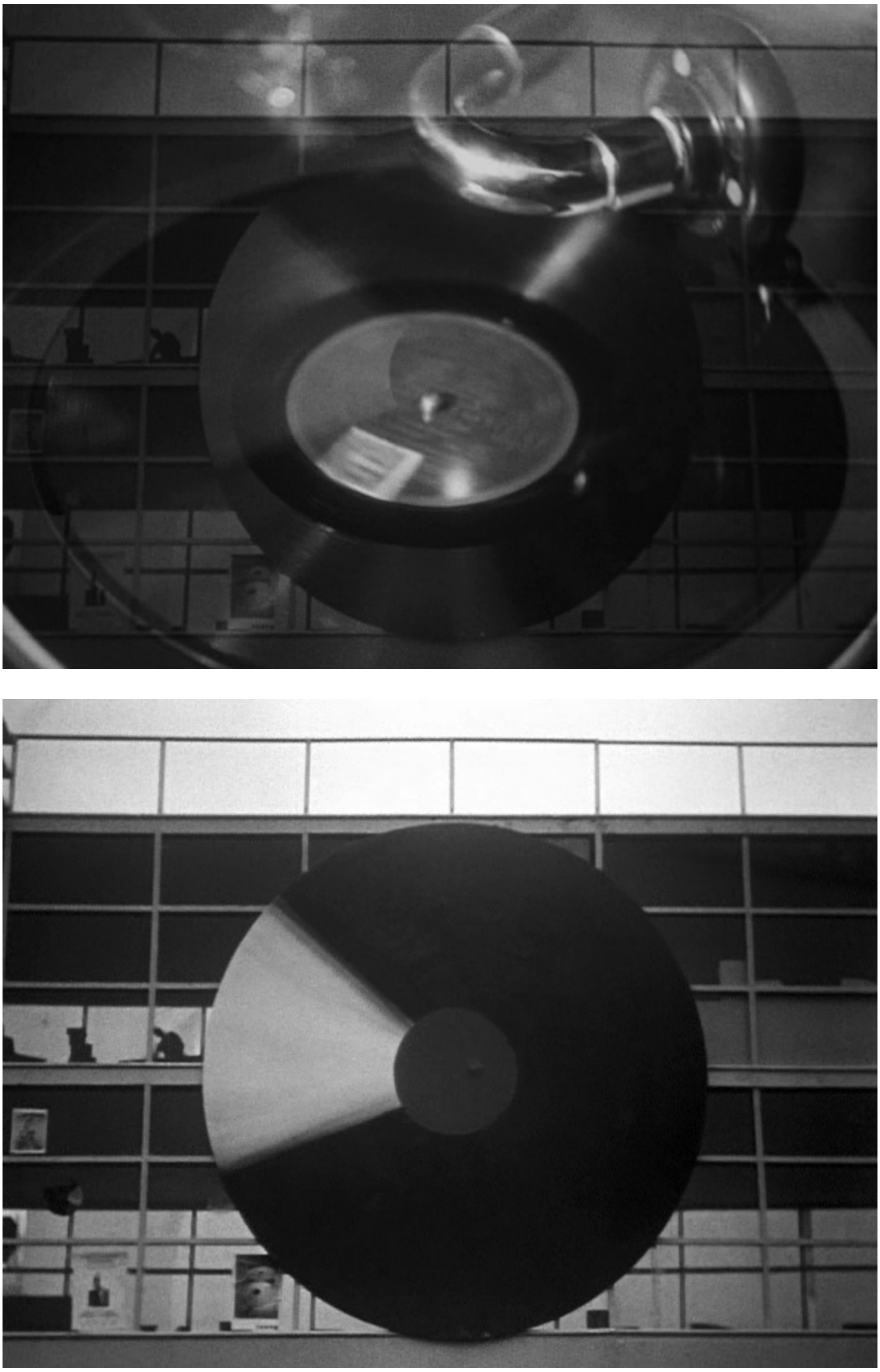
The possibility of film music existing within or outside of the space-time of film opens the possibility for different kinds of film experiences and film meanings. One aspect of film music that I have not thus far explored is film music that exists *both* diegetically and nondiegetically. I do not mean in the same instance, as music cannot be unheard and heard at the same time, but music that shifts between the two realms, completely changing the meaning of the music-image when in either realm. In *The Conversation*, diegetic and nondiegetic music converge into the same image, but there are melodies and tunes that shift between these two spatiotemporal realms that completely alter the meaning of the music-images presented.

In *À Nous la Liberté* (Clair, 1931), Louis escapes prison and begins his freedom as a poor record salesman. A montage begins showcasing Louis' rise to success (to become head of a phonograph company) with a close-up of a record playing music of a trombone accompanied by cornet and harp, performing a soft melody that creates a humble space inside Louis' phonograph shop looking out the front window. From this beginning image, the music is known to be diegetic, it is seen being played, becoming an invisible source as the close-up fades to Louis at the shop counter. As the customer he serves leaves the store, Louis winds up the phonograph beside him, crossfading to a new piece of music (using different instruments) and space. Louis is now dressed more as a wealthy businessman illustrating his success in the duration of time between shots. After assisting a man who dropped his wallet and ushering a woman into his store, he resets the phonograph arm in a close-up, causing the music to restart diegetically. However, the image crossfades again, transitioning to the front of a factory in a match cut (show Figures 164-167). Time and space have changed again, but there is also now no phonograph in the new space. The diegetic music emitting from the phonograph, showcasing the success of Louis' business endeavours, has transformed to an image of a phonograph factory. It is misleading to suppose that the phonograph has jumped in time with the cut, and so it is safe to presume that the music has carried through the edit nondiegetically as the music remains playing the same cue without interruption. The move from diegetic to nondiegetic music with the cut from a phonograph to a phonograph factory has shifted the meaning of the image. At first the diegetic music accompanied the evolution of Louis' success as a businessman, but now the nondiegetic shift has created the metaphor of a phonograph generating the success of a phonograph empire. Colin Roust points out that the music, when it has its nondiegetic shift, leads "to a full jazz band playing a contrapuntally complex foxtrot."[424] Louis is seen in front of the factory as a rich "capitalist fat cat" with assistants and a chauffeur.[425] The move to a full jazz band with a variation in orchestra in the image of his success as a phonograph factory owner highlights a different audio-visual timbre and texture. A shift in audio-visual timbre and texture can change the emotional energy or tension of the image. In *À Nous la Liberté* the shift from diegetic to nondiegetic music using the same piece of music presents Louis' rise in society to a life of success, complexity, and sophistication.[426] Roust argues that in this transition to the shot of the factory, the diegetic status of the music could remain since it could be emitting from phonographs inside the factory, but since the music defies the shift in time and space through the transition, this makes the music exist outside the space and time of the factory.

---

[424] Colin Roust, "'Say it with Georges Auric': Film Music and the esprit nouveau," *Twentieth-Century Music* 6, no. 2 (2009): 143.
[425] Ibid.
[426] Ibid.

L
Enrico DI MAZZEI



Figures 164-167. Louis resets his phonograph, with the audio-visual cue triggering a match cut to the front of a factory in *À Nous la Liberté*.

The above example is a lesser-known variant of diegesis called ‘transdiegetic’ film music. Henry Taylor defines the transdiegetic film music as “a variant…referring to sound’s propensities to cross the border of the diegetic to the nondiegetic and remaining unspecific.”[427] Taylor did not discover this variant; it has been widely implemented in film throughout the 20th and into the 21st centuries but it is quite new to the Film Music Studies dialogue. Robynn Stilwell, writing in the same year as Taylor (2007), proposes her ‘fantastical gap’ theory which sounds quite similar:

> "The diegetic and nondiegetic are conceived as separate realms, almost like two adjacent bubbles, and there seems to be little possibility of moving from one to the other without piercing the skin that explodes the two “universes,” which certainly is one reason for the reliance on the language of “transgression.”"[428]

Even though Stillwell uses the term ‘transgression’ instead of ‘transdiegetic’, Stilwell references a roundtable discussion with James Buhler some years prior in which Buhler mentions that the music on Skull Island in *King Kong* (Cooper and Shoedsack, 1933) is neither nondiegetic nor diegetic, and would rather locate the ‘fantastic’ in the gap between “what we hear and what we see”.[429] Stilwell agrees and coins the term ‘fantastical gap’ to refer to a diegetic crossing that can signify “fantasy (cinematically, a musical number, dream, or flashback)…[and] could also mean, musically, an improvisation, the free play of possibility.”[430] Michel Chion also alludes to this form of musical variant:

> “…the numerous cases in current films where music established as on-the-air freely circulates between the two levels. In *Taxi Driver*, Bernard Herrmann's main theme,

---

[427] Henry M. Taylor, "Discourses on Diegesis: The Success Story of a Misnomer," *Offscreen* 11, no. 8-9 (2007): 3-4.

[428] Stilwell, "The Fantastical Gap between Diegetic and Nondiegetic," 186.

[429] Ibid., 187.

[430] Ibid.

> heard as pit music throughout much of the film, crops up as the music on a phonograph to which the pimp (Harvey Keitel) and his young hooker (Jodie Foster) dance."[431]

Taylor’s concept of transdiegetic music is the closest variant in Film Music Studies to describing a wider field of nonbinary diegesis I have been able to trace. To understand what this action does to film experience rather than just to locate the instances in film when and where it occurs, I want to offer an analysis of several of these instances. *The Third Man* (Reed, 1949) begins with “one of the strangest opening-credit sequences in cinema”, revealing the musical instrument the zither in extreme close-up.[432] However, before the zither begins playing, the sound hole on a bleak white surface looks like a barren landscape. The strings that overlay this image look like the linear geometry that would typically open a Saul Bass opening credits sequence. These strings begin vibrating to the rhythm of the music playing, connecting the two diegetically as the opening credits roll. The strings act as lines on a page for the words in the credits as if the zither plays a role in the presentation of the film. With the entire contents of the visual image being the face of a zither for this opening title sequence, the diegetic zither rhythm seems to exploit the image that could be constructed in the film experience as no hands are seen strumming these strings. The strings seem to be vibrating by themselves to play the music. The music is not limited to this sequence as a quick fade to black then to an establishing shot of Vienna does not pause or stop the music from playing. The zither music continues to play, with the zither landscape being replaced by the streets of Vienna. Michel Chion notes that the music “expresses the opaque and closed feel of an old city that retains its impenetrability despite the disasters of war” through its

---

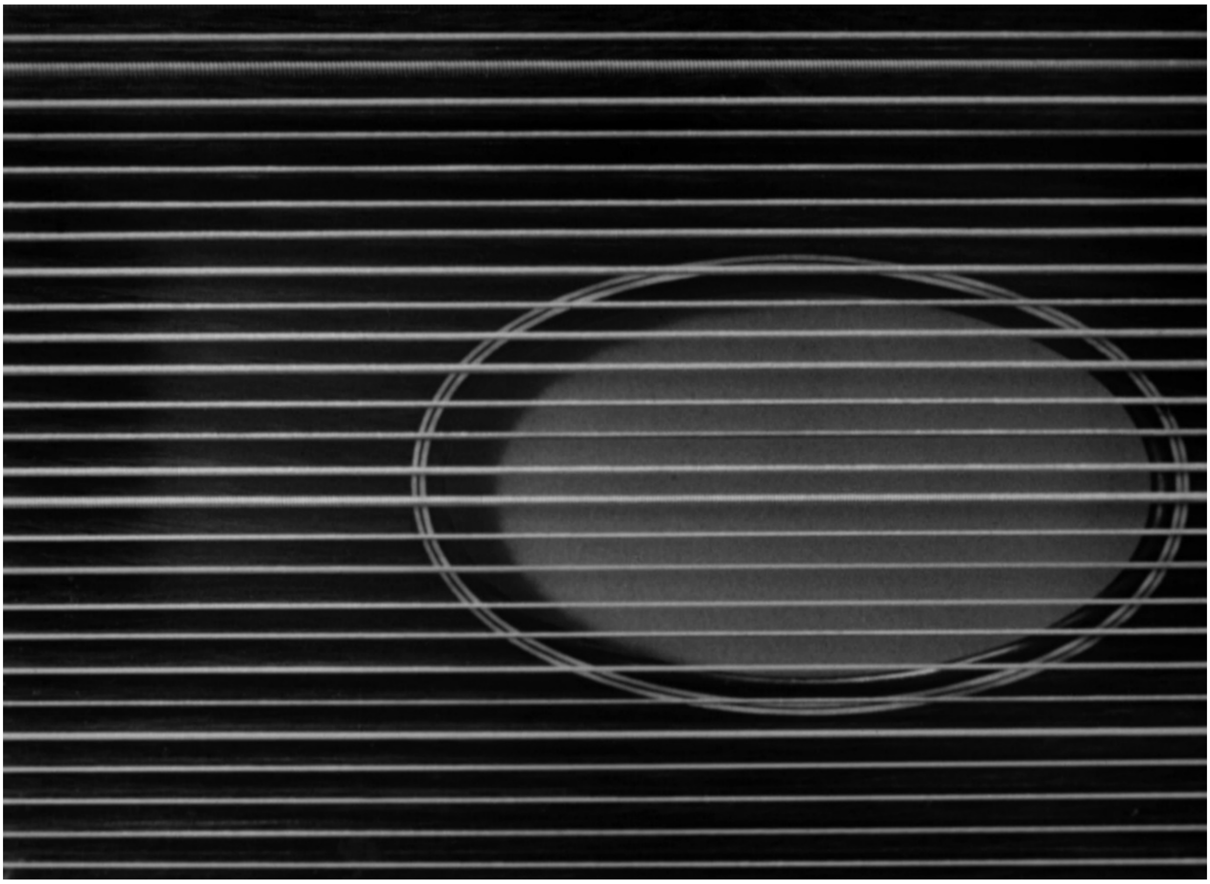
[431] Chion, *Audio-Vision: Sound on Screen*, 81.
[432] "Classicism to Modernism: 1935-1975," in *Music in Cinema*, ed. Claudia Gorbman (New York: Columbia Univeristy Press, 2021), 117.

"particularly energetic and cheery" tone that persists.[433] This audio-visual disharmony is possible through the transition from the opening credits. The zither from diegetic to nondiegetic state retains its prankster quality of controlling the image. Royal S. Brown states that an occurrence of diegetic music becoming nondiegetic can influence the narrative.[434] Brown would agree then that this crossover in diegesis affects the narrative discourse as the zither's audio-visual cues occur whenever Harry Lime is mentioned, and most famously when Harry is revealed out of the shadow in a doorway. The link between the opening diegetic image, the prankster disharmony, and Harry Lime denotes Harry as a significant plot element, and that he could be up to no good. [See Figures 168-170]

---

[433] Ibid.

[434] Brown, *Overtones and Undertones: Reading FIlm Music*, 70.



CAROL REED'S
PRODUCTION
THE
Third Man
BY
GRAHAM GREENE

Figures 168-170. The face of a zither turns from a barren landscape to the streets of Vienna through the zither's music in *The Third Man*.

Towards the end of *Amadeus* (Foreman, 1984), Antonio Salieri takes Wolfgang Amadeus Mozart home after he collapses during a performance of his opera, *The Magic Flute*, and persuades him to finish his Requiem mass. Considering Mozart's bed-ridden illness, Salieri helps Mozart finish the mass. During the composing, Salieri hands some pages over to Mozart to proofread before Salieri moves on to the next section. When Mozart reads a page of the *Confutatis* sequence, a sourceless musical source of the music written on the page is heard as Mozart sings along to the melody. When Mozart hands back the page, the source music ends, assuming Mozart owns control of the music, which he does, but it also implies that the music is metadiegetic, emitting from Mozart's mind. This is important to note, as around three minutes later in the film sequence Salieri finishes scribing the *Confutatis* sequence of the Requiem mass and the sheet music is again ready for Mozart to proofread. After Mozart shouts “show me the whole thing from the beginning”, implying it is finished with all instrument parts combined, instead of a sourceless music source, the shot of Mozart counting in the first bar cuts to a horse carriage at night on the cue of the menacing string ostinato. As the carriage exits stage right of the shot, the shot cuts back to Salieri watching Mozart proofread. It can be experienced as Mozart having constructed this audio-visual image in his mind due to the earlier proofread, but the carriage is eventually revealed to be transporting his wife, Constanze, home, where she greets Salieri and Mozart in the bedroom in which they have been composing the music. The music is constructed within the diegetic space, comes alive in Mozart’s mind metadiegetically, then aids nondiegetically in the construction of the image of Constanze travelling home. The music moves from a construction site to a constructed image through diegesis. [See Figures 171-176]

Good. Show me the whole thing
from the beginning.

Figures 171-176. Mozart constructs an audio-visual sequence of Constanze returning home to where he composed the music in *Amadeus*.

The example from *Amadeus* is not limited to this film as the transition from musical performance to a constructed image in which the source is not present is quite common in sound cinema. Christian Metz suggests that "…a nondiegetic image must in one way be linked to a diegetic image, or it will not appear to be a nondiegetic image."[435] The nondiegetic film music is linked to the image, and so it is premature to suggest that the music could not be part of, or was part of, the film world. Many films freely float music within the film image, inside or outside the spatiotemporal parameters of the film world. For example, in *Three Colours: Blue* (Kieslowski, 1993), the composition of Julie's late husband constantly shifts between being played on the piano or hummed by a character and thick audio-visual texture through nondiegetic orchestra, depending on Julie's mental state. The music transforms texturally and tonally but flows freely within the diegesis to construct different audio-visual meaning via the same composition.

One of the most famous examples of a nonbinary melody comes from the ending of *Close Encounters of the Third Kind* (Spielberg, 1977) where the music does not rest in one diegetic realm. Throughout the film, the extra-terrestrials transmit a 5-tone melody to humans, whether it be through radio, telephones, or other sound communication channels. By the end of the film, the scientists led by ufologist Claude Lacombe understand the melody to be geographic coordinates. The scientists end up at Devil's Mountain in Wyoming where the extra-terrestrial mothership eventually appears and lands. The scientists try to communicate through the 5-tone melody on an electric keyboard, which is organised in the combination: "one; one major second up; one major third down; one octave down; one perfect fifth up."[436] The extra-terrestrials are also given visual aid via a flat screen that associates the tones with colour blocks, to which they respond with colour blocks of their own from their ship. One scientist asks, "What are we saying to each other?" Another responds, "It seems they're trying to teach us a basic tonal vocabulary," and then another says, "It's the first day of school, fellas." The scene is shot in a very simple set up of shot-reverse-shots between the spaceship and various scientists and human characters, however the conversation to spark the shot set up is the unconventional communication through the universality of music (see Figures 177-178). As soon as the ship's door opens, however, the ship stops communicating and an eerie tonal cluster that is "highly dissonant" and a "sound halfway between music and noise" constructs the image of the first visual encounter with an extra-terrestrial creature's ship as one of tense horror.[437] However, instead of one of the creatures coming out to greet the humans, the lost WWII Naval squadron appears and so does the little boy Barry. Moments later, the creatures do appear, and the music is a far more dissonant "noise", re-establishing the anxiety in the image. The tonal clusters become a pedal point (however, higher in register than usual) as a xylophone and strings begin playing a soft melody to construct an image of innocence from the initial horror. That pedal finally evaporates as Roy, the main character is chosen by the creatures to board their ship. The music builds in texture in the stringed melody and crescendos to a peak as Roy makes eye contact with Lacombe, Barry and then Barry's mother Gillian to shout a "this is it" moment of looking back one final time before boarding the spaceship. The harsh white and blue light coming from the interior of the ship creates the barrier between humans and the extra-terrestrials, an unmarked territory of communication and life together. The music halts completely as Roy takes a step onto the ship and that same 5-tone melody plays nondiegetically softly on a flute and clarinet in call and response. The melody that sounded like a cypher on various sound mediums and then an electronic keyboard has been replaced by earthly instruments that are often said to sound like the human voice.[438] The cluster of notes have been recognised as a meeting of understanding between the humans and extra-terrestrials. The image of tense horror and anxiety as to what the extra-terrestrials would be like has transformed into a mutual interlocking of peace and harmony. There is even a brief visual aid to this interlocking as Lacombe performs sign language to a creature to the melody, and the creature responds without the aid of the melody to suggest that humans and extra-terrestrials are no longer apart and 'alien', they can communicate simply and easily. The film ends with the ship taking off into space to the melody played by the entire brass section of the orchestra reaching an emotional peak conveying the sense of pleasure and awe of their contact (See Figure 179). The 5-tone melody, in shifting from a cluster of notes to communicate to an image of peaceful resolution and mutual understanding, demonstrates that the crossover in diegesis reconstructs the audio-visual image for the purposes of narrative experience.

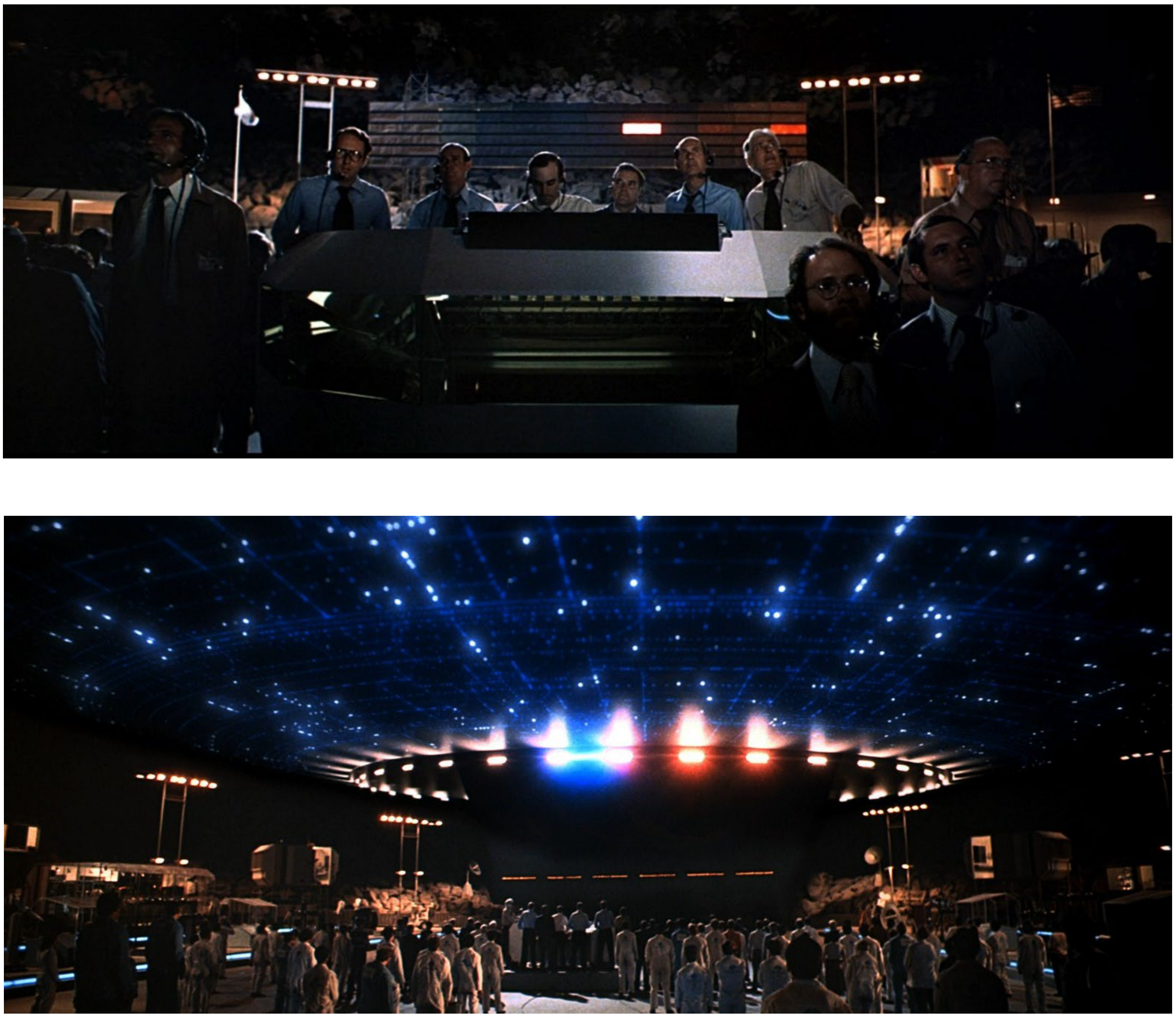
---

[435] Christian Metz, *Film Language: a Semiotics of the Cinema*, trans. Michael Taylor (Chicago: The University of Chicago Press, 1991), 211.

[436] Audissino, *Film/Music Analysis: A Film Studies Approach*, 193.

---

[437] Ibid., 195-96.
[438] Ibid., 200.



Figures 177-178. A shot reverse-shot setup of a musical conversation between the scientists and the spaceship in *Close Encounters of the Third Kind*.

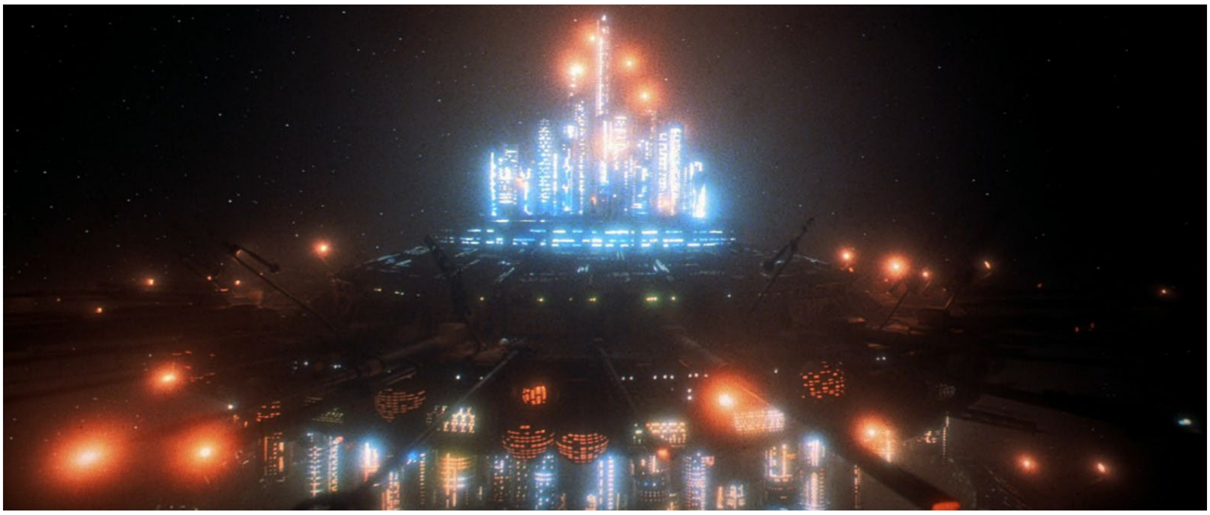


Figure 179. The spaceship in *Close Encounters of the Third Kind* takes off into space to the 5-tone melody played by orchestra.

When nondiegetic film music enters the diegesis

I have mentioned the crossover from diegetic to nondiegetic music in films in the above examples but have reserved some peculiar instances of nondiegetic score entering the diegesis. Analysing an example from *Mr. & Mrs. Smith* (Hitchcock, 1941), Royal Brown acknowledges these moments: "Occasionally, a theme from the nondiegetic music track enters into the diegesis…Indeed, if handled in the right way, the intrusion of recognizable nondiegetic music into the diegesis can have a very unsettling effect."[439] In *Harry Potter and the Philosopher's Stone* (Columbus, 2001), in one of only two instances of diegetic music in the entire film (the other being a harp playing for the three-headed dog, Fluffy), Hagrid plays the film's main theme, Hedwig's Theme, on the flute.[440] This scene occurs immediately before Harry, Ron, and Hermione enter the trapdoor to retrieve the Philosopher's Stone. Harry has a revelation about the stone and about what Hagrid told a hooded figure to receive a dragon's egg in return. The trio notice Hagrid as he can be heard offscreen playing the flute. However, playing the film's main theme diegetically is jarring. How does Hagrid know this tune? It is important to understand that the theme is used nondiegetically throughout the film whenever Harry is associated with magic or has a revelation. The theme is first heard during the title credits, after Professor Dumbledore reveals the baby placed on the front doorstep of 4 Privet Drive is Harry Potter. It is not performed by a flute in the title credits, but by a full orchestra accompanied by a chorus of chanters. Also, the theme is used two scenes after the Hagrid flute encounter, as soon as Harry announces that he, Ron and Hermione are "going down the trapdoor, tonight." The theme is very much associated with Harry, and so it makes sense narratively that Harry notices Hagrid on hearing Hedwig's Theme, but it does not make complete sense spatiotemporally, as it implies Hagrid knows the theme, or happens to be playing the theme unawares of the production of the film, making him the composer of the theme within the film world. [See Figures 180-183]

---

[439] Brown, *Overtones and Undertones: Reading FIlm Music*, 70.

[440] The harp music played for the three-head dog, Fluffy, is not based on existing music like the flute music.

Figures 180-183. Harry, Ron, and Hermione notice Hagrid playing the flute to the tune of the film's main theme 'Hedwig's Theme' in *Harry Potter and the Philosopher's Stone*.

Another example of this jarring nondiegetic to diegetic crossover comes from *The Godfather* (Coppola, 1972) and *The Godfather: Part III* (Coppola, 1990). At the opening wedding scene in *The Godfather*, Vito Corleone and his daughter Connie, the bride of Carlo, have a dance. During the dance, the main theme that was heard in the title credits (and throughout the film) is played diegetically by the live band.[441] Similarly in *The Godfather: Part III* there is a party that follows the ceremony of Michael Corleone awarded Commander of the Order of St. Sebastian. Following a photo of the Corleone family at the party where Michael invites Vincent, his nephew, to join the family photo (thereby introducing Vincent into the family empire), the live band performs the main theme used in the film trilogy.[442] The main theme plays as soon as the shot of the photo flash on Michael's face cuts to a wide shot of the family dispersing back into the crowd of the party. During this audio-visual cue, Michael is referred to as the "Godfather", given a cake which he slices, then dances with his daughter, similar to the example from *The Godfather*. These instances in the *Godfather* series raise the question: how do the characters know this tune? Again, the audio-visual cue has narrative resemblance that ties with its nondiegetic usage in the trilogy, but the question remains of its existence within the film world. Again, the live band in each example could have written the music, or the characters *already know* the music, not heard within the film/s but elsewhere.

---

[441] Bruce Isaacs mentions that the main theme plays "nondiegetically as accompaniment to the visual image." However, the spectator can assume due to the live band playing music throughout the wedding ceremony that the music is indeed diegetic, it can be heard. The live band is not seen during the long shot of Vito and Connie dancing (the band is just offscreen to the left), but with the fusion of diegetic ambient sound and the main theme, while there is a live band within the diegetic space to which Vito and Connie dance, it can safely be assumed that the main theme is being performed and heard within the diegesis. See Bruce Isaacs, *The Art of Pure Cinema: Hitchcock and His Imitators* (New York, USA: Oxford University Press, 2020), 178-79.

[442] The main theme in *The Godfather: Part III* (as well as the other two *Godfather* films) is used in each film's title credits and throughout the films, similar to the *Harry Potter* franchise.

I want to end the chapter on a phenomenon that is not common to film production and subsequently film experience but is intriguing, nonetheless. In an interview with the director of *Guardians of the Galaxy* (Gunn, 2013), James Gunn, he describes using the "unusual" technique of playing music "on the set while the actors are performing."[443] Gunn learnt this technique from director Sergio Leone, who often incorporated this technique in collaboration with composer Ennio Morricone. This technique makes the actors aware of the music that the spectator is unaware the characters can hear. Although this technique in film production is not widely known by cinema spectators, it provides an interesting layer of meaning for those who do know and experience these films. Gunn defends the technique, saying: "music informs the way the cameras move, the performance, the cuts, everything…when the actors hear the music play…they get a much better idea of the tone of the scene."[444] This creates a different tone to the final product of a film when compared to a film where this technique is not performed, but it also changes film experience knowing this occurred. The spectator at face value generates a concept of the image perceived through film experience, and so would include the music within the image of interpretation. With another layer of knowing that the actors hear the same music the audience hears, the meaning of the music-image changes. Knowing that the music heard is played inside the film world, that the actors can hear the music, generates a nonbinary musical source (neither strictly diegetic or nondiegetic) that slips and slides between diegetic realms. Leone once commented: “I had some of the music played on set. It created the atmosphere of the scene. The performances were definitely influenced by it.”[445] In Leone’s final film, *Once Upon a Time in America* (Leone, 1983), if the music indeed was played during filming, the film’s main theme can be heard by the characters. In the scene set in 1918 where Noodles, Patsy, Max, Cockeye, and Dominic, as kids, commit petty crimes, Cockeye plays the film’s main theme on the pan flute. With the knowledge that the music is played on set, Cockeye playing this tune is not as jarring as the music could be deemed part of the structure of the film world, the diegesis. The kids run into local boss Bugsy, for whom they used to work, and try to flee his gunshots. All the diegetic sound evaporates when Dominic screams Bugsy’s name to grab their attention. As they run away in slow motion, Cockeye’s theme plays. It is a melody (known as Cockeye’s theme) played on the pan flute, but never played diegetically during the film. The connection between the audio-visual timbre and Cockeye’s pan flute playing, gives the source of the music an ambiguously nonbinary diegetic ontology.

---

[443] Jon Burlingame, "No Quiet on the Set," *Variety* 336, no. 16 (2017): 117.
[444] Ibid.

Conclusion

In this chapter I have explored the shift towards a nonbinary film music diegesis. In *The Conversation* there is a breakdown in the music-image allowing diegetic and nondiegetic music to converge. The breakdown opens the possibility of analysing film music without restricting music to one spatiotemporal realm of film. Breaking down the music-image in a simple thought experiment can question whether music is heard or not inside the diegesis, which could alter the very existence of the image. The ambiguous spatiotemporal nature of film music allows further engagement and comprehension in the film experience and the process of film analysis. When experiencing music as part of the image, the diegetic status of the musical source resides in the semiotic ambiguity that interlocks the music-image. This is not to say that the breakdown of the music-image is a fault in film experience. On the contrary, as I have attempted to show, this 'breakdown' radically opens the possibilities of the music-image and its effect on the spectator. In this sense, the nonbinary film music diegesis and its creative music-images perform a music-image 'in crisis', but this is a crisis that generates radically new possibilities for film music analysis.

---

[445] Christopher Frayling, *Sergio Leone: Something to do with Death* (London: Faber and Faber, 2000), 235.

Conclusion: Beyond the Music-Image

This research project began by questioning the existing academic literature on the role of music in film and ends by proposing an ontology for film music called the music-image. In bridging the methodologies of Musicology and Film Studies, I have attempted to bring music and visual semiotic systems together in an interlocking form to analyse film as one audio-visual image of experience.

There has been a steady trajectory across the chapters in this thesis towards questioning the object and function of what is commonly described as 'diegesis'. In Chapter 6 I established that music located within the narrative space constructs the narrative image, and so the hard line that has traditionally divided diegetic and nondiegetic music is increasingly porous and movable. When asking 'what' the music-image is, I reference nondiegetic music-images. However, after the breakdown of diegesis explored in Chapter 7, the possibilities of experiencing film diegesis are infinite. For example, what if there really is vaporwave music playing in the space of Jesse walking the runway in *The Neon Demon*? What if there is actually vapor floating around in the room producing a soundscape? I interpreted the hallways on the Nostromo in *Alien* as music *filling* the audio-visual space, as if the musical notes are bouncing off the walls. Similarly, in my analysis of *Jaws* in Chapter 4, the waterline acts as the diegetic barrier between the shark being an impossible object, having total control of the audio-visual space, and then becoming mortal (killable) through swimming to the surface. Emilio Audissino makes the excellent point that if Chrissy, in *Jaws*, heard the E-F ostinato she would have swum back to shore and saved her life.[446] How then is the spectator to anticipate this primal instinct? A film has exact audio-visual parameters. That is, it *cannot be changed* through film experience, but it can be *interpreted infinitely*. The inhabitants of the narrative spacetime are part of those exact parameters, and so they cannot rebel against this audio-visual image construction. The study of film diegesis contains complex film music analysis, but also invites interesting and original ways of thinking about film. The point, however, is that all film music analysis belongs under the umbrella of the music-image. The music-image cannot be anticipated and yet the music-image is always present when film music is heard. As soon as the spectator leaves the experience of a film, the music-image ceases to exist, and yet it lasts in memory.[447] The music-image only exists in this way because it requires the spectator, as film requires an audience. The separation of music and the visual image can be likened to the separation of the film image and the spectator, and so the foundation of the music-image is born with film experience. The music-image model that I have outlined in this thesis is not new in that the music-image has always existed between film and spectator, but by articulating the music-image model, I have sought to make it visible.

---

[446] Audissino, *Film/Music Analysis: A Film Studies Approach*, 33.

What if, however, the spectator cannot articulate what they have experienced? I have selected and analysed a wide range of films in this thesis, based on my subjective engagement through a viewing experience. My interpretation of *Jaws*, *The Thing*, or *The Hours* might differ radically from the interpretation of another viewer. But I also have to acknowledge that my interpretation of the ending of the Iranian film, *The Color of Paradise*, could be erroneous because of what Anahid Kassabian calls ‘competence’.[448] Kassabian makes the point that film music ‘works’ through encodings and decodings of culture, language, music knowledge and the fluency of such processes.[449] Competence is a “culturally acquired skill possessed to varying degrees in varying genres by all hearing people in a given culture,” Kassabian argues.[450] I agree with Kassabian that for film music to work, to engage the spectator, there needs to be cultural competence and to this I would add knowledge of film form, its operation, and its function. The music-image does not cohere within a fixed system or theoretical framework. It also provides room for the competence of the spectator, culturally, intellectually, and in a myriad of other ways. My analysis of *Jaws* in Chapter 4 is founded upon a knowledge structure that has been learnt rather than automatically acquired. Kassabian adds:

---

[447] See the “When music-images repeat” subsection in Chapter 5 for how the music-image can contain memory.
[448] See Chapter 1: *How Film Music* Works, in Kassabian, *Hearing Film: Tracking Identifications in Contemporary Hollywood Film Music*.
[449] Ibid., 20.

> "Competence is based on decipherable codes learned through experience. As with language and visual image, we learn through exposure what a given tempo, series of notes, key, time signature, rhythm, volume, and orchestration are meant to signify. But the acquisition and modus operandi of competence are rarely questioned or examined."[451]

In my analysis of *The Conversation* in which I describe the convergence of diegetic and nondiegetic music, I echo Kassabian's sentiments: how or why does anyone know that this is so?[452] The radical subjectivity of the music-image allows for varying degrees of competence. I bridge the methodologies of Musicology and Film Studies to analyse film music as one whole audio-visual image not to make film music analysis exclusive, but to include and celebrate varying methodologies and competencies.

With this competence in mind, I further acknowledge that I have not explored all aspects of film music. In my analysis, I felt that the exploration of documentary film music and pre-existing music in film were not needed to lay the foundations of the form and function of the music-image. These are nonetheless valuable topics for further exploration. Such topics could

---

[450] Ibid.

[451] Ibid., 23.

[452] Ibid.

then include the music-image as ‘memory motif’, which we see operating in Kubrick’s *2001: A Space Odyssey* (a film I briefly mention in Chapter 7). The film’s composer Alex North was given various pieces of classical music by director Stanley Kubrick as temp tracks for North’s original score. Kubrick subsequently rejected North’s score and instead used the temp tracks, such as Johann Strauss II’s *The Blue Danube* and Richard Strauss’ *Also sprach Zarathustra*. Classical music remains widely used in film, with notable examples *Fantasia* (1940) and *Fantasia 2000* (1999) choreographing animated sequences to reimagined pieces of well-known classical music. In the latter example, *Fantasia 2000*, the opening segment is humpback whales flying into the sky, to the music of Ottorino Respighi’s *Pines of Rome*, a 4-movement tone poem about Rome, Italy. This is an instance of music with its own pre-existing context recontextualised for film.[453] Similarly, the use of pre-existing songs in film has a rich history in both film history and film analysis. The point is that although there are many different components and examples of film music, these different components can be explored in further analysis.[454]

I would like to end this thesis in the way in which Kristin Thompson ends each chapter in her seminal *Storytelling in the New Hollywood* – with a “final lesson.”[455] The lesson I take from the thesis is that our relationship to the music-image cannot be static, and cannot end with a single body of work.[456] In order to prevent the separation of music from image, or image from music, the music-image provides the tools to encourage analysis of new forms and new ideas. Gilles Deleuze, whose thinking and methodologies were vital in the development of my thinking, is a philosopher of *difference*. However, difference for Deleuze is not a rebirth; there is no original thought, as "difference, is in-itself already repetition."[457] I would encourage further exploration of existing film music principles and new principles that will inevitably emerge to conceptualise and use the concept of the music-image to strive for the Deleuzean 'virtual', or ideal scenario: that is, analysing film music as a single, indivisible audio-visual image. As long as film music exists in repetition *ad infinitum*, there will be different music-images to analyse.

---

[453] In my analysis of *Amadeus* in Chapter 7, the music has pre-existing context outside the film, but the point of my analysis was to unpack the transdiegetic music-image, where film music exists both diegetically and nondiegetically.

[454] Some of these examples were in my original vision for this research project, but ultimately sidelined as the discussions of what, when, and where is the music-image are foundational. An engaging starting point for the exploration of pre-existing music in film, for example, is the anthology of essays in the book *Changing Tunes: The Use of Pre-existing Music in Film*. See, *Changing Tunes: The Use of Pre-existing Music in Film*, (London; New York: Routledge, 2006).

[455] Kristin Thompson outlines ways in which “storytelling in the New Hollywood might be *improved*.” See, Thompson, *Storytelling in the New Hollywood: Understanding Classical Narrative Technique*, 49.

[456] Ibid.

---

[457] Deleuze, *Difference and Repetition*, 129.

Bibliography

Addison, Heather. "Cinema's Darkest Vision: Looking into the Void in John Carpenter's The Thing (1982)." *Journal of Popular Film and Television* 41, no. 3 (2013): 54-166.

Afra, Kia. "'Vertical Montage' and Synaesthesia: Movement, Inner Synchronicity and Music-Image Correlation in Alexander Nevsky (1938)." *Music, Sound, and the Moving Image* 9, no. 1 (2015): 33-61.

Albouy, Philippe, Lucas Benjamin, Benjamin Morillon, and Robert J Zatorre. "Distinct Sensitivity to Spectrotemporal Modulation Supports Brain Asymmetry for Speech and Melody." *Science (American Association for the Advancement of Science)* 367, no. 6481 (2020): 1043–47.

Altman, Rick. "24-Track Narrative? Robert Altman’s Nashville." *Cinémas* 1, no. 3 (1991): 102–25.

———. *The American Film Musical.* Bloomington; Indianapolis: Indiana University Press, 1987.

———. "Introduction." *Yale French Studies* 60, no. Cinema/Sound (1980): 3-15.

Armstrong, Tom. "Response: Music, image and the sublime." *Textual Practice* 22, no. 1 (2008): 71-83.

Audissino, Emilio. *Film/Music Analysis: A Film Studies Approach*.Palgrave Studies in Audio-Visual Culture. Edited by K. J. Donnelly. Cham, Switzerland: Palgrave Macmillan, 2017.

———. "Jaws: Williams's Neoclassicism Floats Up to the Surface." In *John Williams's Film Music: Jaws, Star Wars, Raiders of the Lost Ark, and the Return of the Classical Hollywood Music Style*. Wisconsin Film Studies. Madison, Wisconsin: University of Wisconsin Press, 2014.

———. "John Williams and Contemporary Film Music." In *Contemporary Film Music: Investigating Cinema Narratives and Composition*, edited by Lindsay Colemen and Joakim Tillman, 221-36. London, UK: Palgrave Macmillan, 2017.

Bailie, Joanna. "Film Theory and Synchronization." *CR: The New Centennial Review* 18, no. 2 (2018): 69-74.

Barham, Jeremy. "Music and the moving image." In *Aesthetics of Music: Musicological Perspectives*, edited by Stephen Downes, 224-38. New York: Routledge, 2014.

Bazin, André. *What is Cinema? Volume 1.*Translated by Hugh Gray. Berkeley; Los Angeles: University of California Press, 1967.

Berndt, Axel. "Diegetic Music: New Interactive Experiences." In *Game Sound Technology and Player Interaction: Concepts and Developments*, edited by Mark Grimshaw, 60-76. Hershey, PA: IGI Global, 2011.

Biancorosso, Giorgio. "The Harpist in the Closet: Film Music as Epistemological Joke." *Music and the Moving Image* 2, no. 3 (2009): 11-33.

———. "The Shark in the Music." *Music Analysis* 29, no. 1-3 (2010): 306-33.

Blackburn, Simon. *Oxford Dictionary of Philosophy.* Third ed. "Indexicality." Oxford, UK: Oxford Univeristy Press, 2016.

———. *Oxford Dictionary of Philosophy.* Third ed. "Post-Structuralism." Oxford, UK: Oxford Univeristy Press, 2016.

———. *Oxford Dictionary of Philosophy.* Third ed. "Structuralism." Oxford, UK: Oxford Univeristy Press, 2016.

Bordwell, David. "Contemporary Film Studies and the Vicissitudes of Grand Theory." In *Post-Theory: Reconstructing Film Studies*, edited by David Bordwell and Noël Carroll, 3-36. Madison, Wisconsin: The University of Wisconsin Press, 1996.

———. "An Excessively Obvious Cinema." In *The Classical Hollywood Cinema: Film Style & Mode of Production to 1960*, edited by David Bordwell, Janet Staiger and Kristin Thompson, 2-10. London: Routledge, 1985.

———. *Poetics of Cinema*. New York; London: Routledge, 2008.

———. "Principles of Narration." In *Narration in the Fiction Film*, 48-62. London: Routledge, 1986.

———. "Time in the Classical Film." In *The Classical Hollywood Cinema: Film Style & Mode of Production to 1960*, edited by David Bordwell, Janet Staiger and Kristin Thompson, 42-49. London: Routledge, 1985.

Bordwell, David, and Kristin Thompson. "Fundamental Aesthetics of Sound in the Cinema." In *Film Sound: Theory and Practice*, edited by Elisabeth Weis, and John Belton. New York: Columbia University Press, 1985.

Bordwell, David, and Noël Carroll. "Introduction." In *Post-Theory: Reconstructing Film Studies*, edited by David Bordwell and Noël Carroll, xiii-xvii. Madison, Wisconsin: The University of Wisconsin Press, 1996.

Bordwell, David, Kristin Thompson, and Jeff Smith. "The Shot: Cinematography." In *Film Art: An Introduction*, 159-215. New York: McGraw-Hill Eduction, 2019.

Braester, Yomi. "Cinephiliac Engagement and the Disengaged Gaze in In the Mood for Love." In *A Companion to Wong Kar-wai*, edited by Martha P. Nochimson, 467-84. West Sussex, UK: Wiley Blackwell, 2016.

Brigden, Charlie. "Revisiting John Williams' Score for Jaws, 45 Years Later." 2020, RogerEbert.com, <https://www.rogerebert.com/features/revisiting-john-williams-score-for-jaws-45-years-later>.

Brown, Royal S. "How not to think film music." *Music and the Moving Image* 1, no. 1 (2008): 2-18.

———. *Overtones and Undertones: Reading FIlm Music*. Berkeley; Los Angeles; London: University of California Press, 1994.

Buckland, Warren. "“Duel With A Shark”: Jaws (1975)." In *Directed by Steven Spielberg: Poetics of the Contemporary Hollywood Blockbuster*, 86-110. New York; London: Continuum, 2006.

Buhler, James. "Review of Hollywood Theory, Non-Hollywood Practice: Cinema Soundtracks in the 1980s and 1990s, by Annette Davison." *Twentieth-Century Music* 3, no. 1 (2007): 145-49.

———. "Star Wars, Music, and Myth." In *Music and Cinema*, edited by James Buhler, Caryl Flinn, and David Neumeyer, 33-57. Hanover, New Hampshire: University Press of New England, 2000.

Buhler, James, and David Neumeyer. "Music and the Ontology of the Sound Film: The Classical Hollywood System." In *The Oxford Handbook of Film Music Studies*, edited by David Neumeyer, 17-43. Oxford; New York: Oxford Univeristy Press, 2014.

Buhler, James, Neumeyer, David, and Rob Deemer. "Music in Main-Title and End-Credit Sequences." In *Hearing the Movies: Music and Sound in Film History*, 165-80. New York; Oxford: Oxford University Press, 2010.

Burlingame, Jon. "The music of Jaws: an interview with John Williams." 2012, Limelight, <https://limelightmagazine.com.au/features/the-music-of-jaws-an-interview-with-john-williams/>.

———. "No Quiet on the Set." *Variety* 336, no. 16 (2017): 117.

Burt, George. *The Art of Film Music*. Boston: Northeastern University Press, 1994.

Butler, David. "The Days Do Not End: Film Music, Time and Bernard Herrmann." *Film Studies*, no. 9 (2006): 51-63.

Carroll, Noël. *Philosophy as Art: A Contemporary Introduction*. London; New York: Routledge, 1999.

Cavell, Stanley. *The World Viewed: Reflections on the Ontology of Film*. Cambridge, USA: Harvard University Press, 1979.

Cecchi, Alessandro. "Creative Titles: Audiovisual Experimentation and Self-Reflexivity in Italian Industrial Films of the Economic Miracle and After." *Music, Sound, and the Moving Image* 8, no. 2 (2014): 179-94.

Changing Tunes: The Use of Pre-existing Music in Film. Edited by Phil Powrie and Robynn Stilwell. London; New York: Routledge, 2006.

Chattah, Juan. "Background and Analysis of the Score for The Conversation." In *David Shire's The Conversation: A Film Score Guide*, 121-61. Lanham; Boulder; New York; London: Rowman & Littlefield, 2015.

Chazelle, Damien, *La La Land*. Blu-ray Featurettes. Australia: Summit Entertainment, 2017.

Chion, Michel. *Audio-Vision: Sound on Screen*. Edited by Claudia Gorbman.Translated by Claudia Gorbman. New York: Columbia University Press, 1994.

———. "Classicism to Modernism: 1935-1975." Translated by Claudia Gorbman. In *Music in Cinema*, edited by Claudia Gorbman, 87-151. New York: Columbia Univeristy Press, 2021.

———. *Film, a Sound Art*.Translated by Claudia Gorbman. New York: Columbia University Press, 2009.

Citron, Marcia J. "Operatic Style in Coppola's Godfather trilogy." In *When Opera Meets Film*, 19-57. Cambridge: Cambridge University Press, 2010.

Cohen, Annabel. "Film Music: Perspectives from Cognitive Psychology." In *Music and Cinema*, edited by James Buhler, Caryl Flinn, and David Neumeyer, 360-78. Hanover, New Hampshire: University Press of New England, 2000.

Cole, Ross. "Vaporwave Aesthetics: Internet Nostalgia and the Utopian Impulse." *ASAP/Journal* 5, no. 2 (2020): 297-326.

Conway, Kelley. "Cultivating the New Wave Spectator: Cleo from 5 to 7." In *Agnes Varda*, 35-55. Urbana; Chicago; Springfield: University of Illinois Press, 2015.

Cook, Nicholas. *Analysing Musical Multimedia*. Oxford: Clarendon Press, 1998.

———. *A Guide to Musical Analysis*. London; Melbourne: J. M. Dent & Sons Ltd, 1987.

Cubitt, Sean. *The Cinema Effect*. Cambridge; London: The MIT Press, 2004.

Custer, Arthur. "Comparing Concert with Film Music." *Back Stage (Archive: 1960-2000)* 20, 16, April 20, 1979, 34.

d'Escrivan, Julio. "Sound Art (?) on/in Film." *Organised Sound* 14, no. 1 (2009): 65-73.

Deleuze, Gilles. *Cinema 1: The Movement-Image.*Translated by Hugh Tomlinson and Barbara Habberjam. London: The Athlone Press, 1992.

———. *Cinema 2: The Time Image.*Translated by Hugh Tomlinson and Robert Galeta. Minneapolis: University of Minnesota Press, 1989.

———. *Difference and Repetition.*Translated by Paul Patton. London: The Athlone Press, 1994.

———. *Francis Bacon: The Logic of Sensation.*Translated by Daniel W. Smith. London; New York: Continuum, 2003.

———. "On the Movement-Image." Translated by Martin Youghin. In *Negotiations 1972-1990*, 46-56. New York: Columbia University Press, 1995.

Deleuze, Gilles, Guattari, Felix. *A Thousand Plateaus: Capitalism and Schizophrenia.*Translated by Brian Massumi. Minneapolis; London: University of Minnesota Press, 1987.

———. *What is Philosophy?*Translated by Hugh Tomlinson and Graham Burchell. New York: Columbia University Press, 1994.

Donnelly, K. J. "Soundtracks without Films." In *The Spectre of Sound: Music in Film and Television*, 150-71. London: British Film Institute, 2005.

Donnelly, K.J. "Synchronisation." In *Occult Aesthetics*. Oxford: Oxford University Press, 2014.

Eco, Umberto. *Semiotics and the Philosophy of Language*. Bloomington: Indiana University Press, 1984.

Eisenstein, Sergei. *The Film Sense*. Edited by Jay Leyda.Translated by Jay Leyda. New York: Meridan Books, 1957.

Eisler, Hanns. *Composing for the Films*. London: Dennis Dobson LTD, 1947.

Elsaesser, Thomas. "(Un)Representing the Real: Seeing Sounds and Hearing Images." In *Psychoanalysis and the Unrepresentable: From Culture to the Clinic*, edited by Agnieszka Piotrowska and Ben Tyrer, 53-73. London: Routledge, 2016.

Elsaesser, Thomas, and Malte Hagener, ed. *Film Theory: An Introduction Through the Senses*. Florence: Taylor & Francis Group, 2009.

Erkelenz, Ralph. "Ben-Hur: A Tale of the Score." *Pro Musica Sana* 5, no. 1 (2005): 1-39.

Fagandini, Hugo. "Film Score Friday #3: "The Thing" (1982)." 2016, Son et Lumière, <https://sonetlumiereblog.wordpress.com/2016/10/14/film-score-friday-the-thing-1982/>.

Flinn, Caryl. *Strains of Utopia: Gender, Nostalgia, and Hollywood Film Music*. Princeton, New Jersey: Princeton University Press, 1992.

Frayling, Christopher. *Sergio Leone: Something to do with Death*. London: Faber and Faber, 2000.

Genette, Gérard. *Paratexts: Thresholds of Interpretation*.Translated by Jane E. Lewin. Cambridge: Cambridge University Press, 1997.

Goddard, Scott. "Review of Music for the Films, by Leonid Sabaneev." *Music & Letters* 17, no. 2 (1936): 162-63.

Gorbman, Claudia. "Narrative Film Music." *Yale French Studies* 60 (1980): 183-203.

———. *Unheard Melodies: Narrative Film Music*. Bloomington; Indianapolis: Indiana University Press, 1987.

Heldt, Guido. *Music and Levels of Narration in Film: Steps Across the Border*. Bristol, UK; Chicago, USA: Intellect, 2013.

Hickman, Roger. *Miklós Rózsa's Ben-Hur: A Film Score Guide*. Lanham, Mayland; Toronto; Plymoth, UK: The Scarecrow Press, Inc., 2011.

Hillman, Roger, and Deborah Crisp. "Chiming the Hours: A Philip Glass Soundtrack." In *Virginia Woolf and Music*, edited by Adriana Varga, 288-310. Bloomington; Indianapolis Indiana University Press, 2014.

Holbrook, Morris B. *Music, Movies, Meanings, and Markets: Cinemajazzamatazz*. New York; London: Routledge, 2011.

Ireland, David. *Identifying and Interpreting Incongruent Film Music*. London: Palgrave Macmillan, 2018.

Isaacs, Bruce. *The Art of Pure Cinema: Hitchcock and His Imitators*. New York, USA: Oxford University Press, 2020.

———. "Cinema’s Autonomous Image in Michelangelo Antonioni and Francis Ford Coppola." *Sydney Studies in English* 38 (2012): 19-49.

Jacobs, Lea. *Film Rhythm after Sound: Technology, Music, and Performance*. Oakland, California: University of California Press, 2015.

Johnson, Timothy A. "Minimalism: Aesthetic, Style, or Technique?". *The Musical Quarterly* 78, no. 4 (1994): 742-73.

Kalinak, Kathryn. *Settling the Score: Music and the Classical Hollywood Film.*Wisconsin Studies in Film. Edited by Donald Crafton David Bordwell, and Vance Kepley Jnr. Madison, Wisconsin: The University of Wisconsin Press, 1992.

Karlin, Fred, and Rayburn Wright. *On the Track: A Guide to Contemporary Film Scoring*. New York; London: Routledge, 2004.

Kassabian, Anahid. *Hearing Film: Tracking Identifications in Contemporary Hollywood Film Music*. New York: Routledge, 2001.

Keathley, Christian. "Trapped in the Affection-Image: Hollywood's Post-Traumatic Cycle (1970-1976)." In *The Last Great American Picture Show: New Hollywood Cinema in the 1970s*, edited by Alexander Horwath and Noel King Thomas Elsaesser, 293-308. Amsterdam: Amsterdam University Press, 2004.

Keller, Hans. "Hans Keller: Essays on Film Music." *Film Studies*, no. 5 (2004): 106-31.

Kennedy, Michael, and Joyce Bourne Kennedy. *Oxford Dictionary of Music.* 6th ed. "tritone." Oxford: Oxford University Press, 2013.

Kermode, Mark. "Kermode Uncut: Unknown Instrument." 2018, YouTube video, <https://www.youtube.com/watch?v=p1t4EUXjVa4>.

Kielian-Gilbert, Marianne. "Listening in Film Music: Music/Film Temporality, Materiality, and Memory." In *The Oxford Handbook of Film Music Studies*, edited by David Neumeyer, 500-25. Oxford; New York: Oxford University Press, 2014.

Knight-Hill, Andrew. "Sonic Diegesis: Reality and the Expressive Potential of Sound in Narrative Film." *Quarterly Review of Film and Video* 36, no. 8 (2019): 643-65.

Koffka, Kurt. *Principles of Gestalt Psychology*. Oxford: Routledge, 1999.

Langkjær, Birger. "Hearing things in music for films: Music, Fiction and Engagement." *SoundEffects* 3, no. 1-2 (2013): 65-77.

Lehman, Frank. "Tonal Practices." In *Hollywood Harmony: Musical Wonder and the Sound of Cinema*, 15-48. New York: Oxford University Press, 2018.

London, Kurt. *Film Music: A Summary of the Characteristic Features of Its History, Aesthetics, Technique; and Possible Developments*. London: Faber & Faber, 1936.

"Lufthansa New York." 2009, Springer and Jacoby. YouTube video, <https://www.youtube.com/watch?v=VPjmJOzsP28>.

Marks, Martin. "Film Music: The Material, Literature, and Present State of Research." *MLA Notes* 36, no. 2 (1979): 282-325.

———. "Music, Drama, Warner Brothers: The Cases of Casablanca and The Maltese Falcon." In *Music and Cinema*, edited by James Buhler, Caryl Flinn, and David Neumeyer, 161-86. Hannover; London: Wesleyan University Press, 2000.

Marshall, Eugene. *The Spiritual Automaton: Spinoza's Science of the Mind*. Oxford: Oxford University Press, 2014.

Mayne, Judith. *Cinema and Spectatorship*. London; New York: Routedge, 1993.

Mazey, Paul. "Restrained Airs: The Diegetic Surface and Nondiegetic Depth of British Film Music." *Journal of British Cinema and Television* 16, no. 4 (2019): 429-43.

McFarlane, Brian. "'Kokoda': Lost and Found on the Trail: To Australians of a Certain Age, and Even to Some Younger, the 'Kokoda Trail' Has Acquired the Sheen of Legendary Significance." *Australian Screen Education*, no. 42 (2006): 8-13.

McGowen, Todd. "Steven Spielberg's Search for the Father." In *The Real Gaze: Film Theory after Lacan*. New York: State University of New York Press, 2007.

Metz, Christian. "The Cinema: Language or Language System?" Translated by Michael Taylor. In *Film Language: A Semiotics of the Cinema*, 31-91. Chicago: The Univeristy of Chicago Press, 1974.

———. *Film Language: a Semiotics of the Cinema*.Translated by Michael Taylor. Chicago: The University of Chicago Press, 1991.

Meyer, Leonard B. *Emotion and Meaning in Music*. Chicago; London: The University of Chicago Press, 1956.

Meyer, Stephen C. ""Leitmotif": On the Application of a Word to Film Music." *Journal of Film Music* 5, no. 1-2 (2012): 101-08.

Moormann, Peter. "Composing with Types and Flexible Modules: John Williams' Two-Note Ostinato for Jaws and its Use in FIlm-Music History." *Journal of Film Music* 5, no. 1-2 (2012): 165-68.

Morris, Nigel. *The Cinema of Steven Spielberg: Empire of Light*. London; New York: Columbia University Press, 2007.

———. "*Jaws*: Searching the Depths." In *The Cinema of Steven Spielberg: Empire of Light*. London; New York: Wallflower Press, 2007.

Mottram, Ron. "American Sound Films, 1926–1930." In *Film Sound: Theory and Practice*, edited by Elisabeth Weis, and John Belton. New York: Columbia University Press, 1985.

Nagari, Benjamin. *Music as Image: Analytical Psychology and Music in Film*.Research in Analytical Psychology and Jungian Studies Series. London; New York: Routledge, 2016.

Nasta, Dominique. *Meaning in Film: Relevant Structures in Soundtrack and Narrative*. Berne; Berlin: Peter Lang, 1991.

Neumeyer, David. "Overview." In *The Oxford Handbook of Film Music Studies*, edited by David Neumeyer, 1-14. Oxford; New York: Oxford University Press, 2014.

Neumeyer, David, James Buhler, and Caryl Flinn. "Introduction." In *Music and Cinema*, edited by James Buhler, Caryl Flinn, and David Neumeyer, 1-29. Hannover; London: Wesleyan University Press, 2000.

Ondaatje, Michael. "Third Conversation." In *The Conversations: Walter Murch and the Art of Editing*, 151-99. New York: Alfred A. Knopf, 2008.

Online, OED. "arpeggio, n.", 2021, Oxford University Press, <https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/10945?rskey=9zE17s&result=1&isAdvanced=false>.

———. "cadence, n.", 2021, Oxford University Press, <https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/25947?rskey=kISb2f&result=1>.

———. "cue, n.2." 2021, Oxford University Press, <https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/45574?rskey=72N3xh&result=2>.

———. "diegesis, n.", 2021, Oxford University Press, <https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/52402?redirectedFrom=diegetic>.

———. "leitmotiv, n.", 2022, Oxford University Press, <https://www.oed.com/view/Entry/107180?redirectedFrom=leitmotiv#eid>.

———. "non-diegetic, adj.", 2021, Oxford University Press, <https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/69156318>.

———. "overture, n., 7a." 2022, Oxford University Press, <https://www.oed.com/view/Entry/135281?rskey=X3ZfSK&result=1>.

———. "romantic, adj., 7." 2022, Oxford University Press, <https://www.oed.com/view/Entry/167122?rskey=m9T4MX&result=1>.

———. "sonata form, n.", 2022, Oxford University Press, <https://www.oed.com/view/Entry/184563?redirectedFrom=sonata+form#eid1212072090>.

———. "timbre, n.3." 2021, Oxford University Press, <https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/202089?rskey=Kpaswo&result=3>.

———. "tone, n.", 2021, Oxford University Press, <https://www-oed-com.ezproxy.library.sydney.edu.au/view/Entry/203150?rskey=J3vEQI&result=1>.

Oukaderova, Lida. "I Am Cuba and the Space of Revolution." *Film & History* 44, no. 2 (2014): 4-21.

The Oxford Handbook of Film Music Studies. Edited by David Neumeyer. Oxford; New York: Oxford University Press, 2014.

Penner, Nina. "Rethinking the Diegetic/Nondiegetic Distinction in the Film Musical." *Music and the Moving Image* 10, no. 3 (2017): 3-20.

Petermann, Emily. "The Film Musical as a Subject for Word and Music Studies." In *Silence and Absence in Literature and Music*, edited by Werner Wolf and Walter Bernhart. Word and Music Studies, 223-36. Leiden; Boston: Brill, 2016.

Powrie, Phil, and Guido Heldt. "Introduction: Trailers, Titles, and End Credits." *Music, Sound, and the Moving Image* 8, no. 2 (2014): 111-20.

Prendergast, Roy M. *A Neglected Art: A Critical Study of Music in Films*. New York: New York University Press, 1977.

Pudovkin, V. I. "Asynchronism as a Principle of Sound Film." In *Film Sound: Theory and Practice*, edited by Elisabeth Weis and John Belton, 86-91. New York: Columbia University Press, 1985.

Redner, Gregg. *Deleuze and Film Music: Building a Methodological Bridge Between Film Theory and Music*. Bristol; Chicago: Intellect, 2011.

Reinhard, Carrielynn D., and Christopher J. Olson, ed. *Making Sense of Cinema: Empirical Studies into Film Spectators and Spectatorship*. New York; London; Oxford; New Delhi; Sydney: Bloomsbury, 2016.

Romero-Fresco, Pablo. *Accessible Filmmaking: Integrating translation and accessibility into the filmmaking process*. London: Routledge, 2019.

Rosar, William H. "Film Studies in Musicology: Disciplinarity vs. Interdisciplinarity." *The Journal of Film Music* 2, no. 2-4 (2009): 99-125.

Roust, Colin. "'Say it with Georges Auric': Film Music and the esprit nouveau." *Twentieth-Century Music* 6, no. 2 (2009): 133-53.

Sabaneev, Leonid. *Music for the Films: A Handbook for Composers and Conductors*.Translated by S. W. Pring. London: Pitman, 1935.

Schmidt, Gary. "Between Authors and Agents: Gender and Affirmative Culture in Das Leben der Anderen." *The German Quarterly* 82, no. 2 (2009): 231-49.

Scott, D. Travers. "Intertextuality as 'Resonance': Masculinity and Anticapitalism in Pet Shop Boys' Score for Battleship Potemkin." *Music, Sound, and the Moving Image* 7, no. 1 (2013): 53-83.

Scruton, Roger. *Music as an Art*. London; New York; Oxford; New Delhi; Sydney: Bloomsbury Continuum, 2018.

Sim, Gerald. "The Other Person in the Bathroom: Mixed Emotions about Cognitivist Film Music Theory." *Quarterly Review of Film and Video* 30, no. 4 (2013): 309-22.

Smith, Jeff. "Bridging the Gap: Reconsidering the Border between Diegetic and Nondiegetic Music." *Music and the Moving Image* 2, no. 1 (2009): 1-25.

———. "Unheard Melodies? A Critique of Psychoanalytic Theories of Film Music." In *Post-Theory : Reconstructing Film Studies*, edited by David Bordwell and Noel Carroll, 230-47. Madison, Wisconsin: University of Wisconsin Press, 1996.

Spinoza, Benedict de. *Tractatus de Intellectus Emendatione*.Translated by W. Hale White. London: T. Fisher Unwin, 1895.

Spottiswoode, Raymond. *A Grammar of the Film*. Berkeley: University of California Press, 1950.

Stanitzek, Georg. "Reading the Title Sequence." *Cinema Journal* 4, no. 48 (2009): 44-58.

Starrs, D. Bruno. "Film as Text [1931-2007]: "Picnic at Hanging Rock" and the Puzzle of the Art Film." *Screen Education* 48 (2008): 139-44.

Stein, Jack M. *Richard Wagner and the Synthesis of the Arts*. Detroit: Wayne State University Press, 1960.

Stilwell, Robynn J. "The Fantastical Gap between Diegetic and Nondiegetic." In *Beyond the Soundtrack: Representing Music in Cinema*, edited by Daniel Goldmark, Lawrence Kramer, and Richard Leppart, 184-202. Berkeley; Los Angeles; London: University of California Press, 2007.

Taylor, Henry M. "Discourses on Diegesis: The Success Story of a Misnomer." *Offscreen* 11, no. 8-9 (2007): 1-5.

Thompson, Kristin. *Breaking the Glass Armor: Neoformalist Film Analysis*. Princeton, New Jersey: Princeton University Press, 1988.

———. *Storytelling in the New Hollywood: Understanding Classical Narrative Technique*. Cambridge, USA; London, UK: Harvard University Press, 1999.

Uhlin, Graig. "Film Rhythm and the Aporetics of Temporality." *Quarterly Review of Film and Video* 36, no. 2 (2019): 104-24.

Vande Moortele, Steven. "Introduction: Formenlehre, Genre, and the Romantic Turn." In *The Romantic Overture and Musical Form from Rossini to Wagner*, 1-14. Cambridge: Cambridge University Press, 2017.

Vitouch, Oliver. "When Your Ear Sets the Stage: Musical Context Effects in Film Perception." *Psychology of Music* 29 (2001): 70-83.

Wagner, Richard. "Opera and Drama." *The Musical World* 34, no. 9 (1856): 132-33.

———. *Opera and Drama*. Lincoln: University of Nebraska Press, 1995.

Wierzbicki, James. "Anahid Kassabian: Hearing Film: Tracking Identifications in Contemporary Hollywood Film Music." *Journal of Film Music* 1, no. 4 (June 19, 2001): 461-63.

———. *Film Music: A History*. New York; London: Routledge, 2008.

Williams, James. *Gilles Deleuze's Difference and Repetition: A Critical Introduction and Guide*. Edinburgh: Edinburgh University Press, 2003.

Williams, Linda Ruth. "The shock of The Red Shoes." *New Review of Film and Television Studies* 15, no. 1 (2017): 9-23.

Wingstedt, Johnny, Sture Brändström, and Jan Berg. "Narrative Music, Visuals and Meaning in Film." *Visual Communication* 9, no. 2 (2010): 193-210.

Winters, Ben. "Musical Wallpaper?: Towards an Appreciation of Non-narrating Music in Film." *Music, Sound, and the Moving Image* 6, no. 1 (2012): 39-54.

———. "The Non-Diegetic Fallacy: Film, Music, and Narrative Space." *Music & Letters* 91, no. 2 (2010): 224-44.

Yacavone, Daniel. "Spaces, Gaps, and Levels: From the Diegetic to the Aesthetic in Film Theory." *Music, Sound, and the Moving Image* 6, no. 1 (2012): 21-37.

Filmography

Allen, Woody, *Bananas*. USA: United Artists, 1971.

———, *Midnight in Paris*. USA: Sony Pictures Classics, 2011.

Armstrong, Samuel, James Algar, Bill Roberts, Paul Satterfield, Ben Sharpsteen, David D. Hand, Hamilton Luske, Jim Handley, Ford Beebe, T. Hee, Norman Ferguson, Wilfred Jackson, *Fantasia*. USA: RKO Radio Pictures, 1940.

Avildsen, John G., *Rocky*. USA: United Artists, 1976.

Beaumont, Harry, *The Broadway Melody*. USA: Metro-Goldwyn-Mayer, 1929.

Bergman, Ingmar, *Persona*. Sweden: AB Svensk Filmindustri, 1966.

———, *Wild Strawberries*. Sweden: AB Svensk Filmindustri, 1957.

Bont, Jan de, *Twister*. USA: Warner Bros. Pictures, 1996.

Brooks, James L., *Terms of Endearment*. USA: Paramount Pictures Studios, 1983.

Brooks, Mel, *Blazing Saddles*. USA: Warner Bros. Pictures, 1974.

Burton, Mark, and Richard Starzak, *Shaun the Sheep Movie* UK: StudioCanal, 2015.

Buzzell, Edward, *Go West*. USA: Metro-Goldwyn-Mayer, 1940.

Cameron, James, *The Terminator*. USA: Orion Pictures, 1984.

———, *Terminator 2: Judgment Day*. USA: Tri-Star Pictures, 1991.

Carney, John, *Sing Street*. Ireland: Lionsgate, 2016.

Carpenter, John, *Halloween*. USA: Compass International Pictures, 1978.

———, *The Thing*. USA: Universal Pictures, 1982.

Chaffey, Don, *Jason and the Argonauts*. USA: Columbia Pictures, 1963.

Chazelle, Damien, *La La Land*. USA: Lionsgate, 2016.

Clair, René, *À Nous la Liberté*. France: Films Sonores Tobis, 1931.

Columbus, Chris, *Harry Potter and the Philosopher's Stone*. UK: Warner Bros. Pictures, 2001.

Cooper, Bradley, *A Star is Born*. USA: Warner Bros. Pictures, 2018.

Cooper, Merian C., and Ernest B. Schoedsack, *King Kong*. USA: RKO Radio Pictures, 1933.

Cooper, Scott, *Hostiles*. USA: Entertainment Studios Motion Pictures, 2017.

Coppola, Francis Ford, *Apocalypse Now*. USA: United Artists, 1979.

———, *The Conversation*. USA: Paramount Pictures, 1974.

———, *The Godfather*. USA: Paramount Pictures, 1972.

———, *The Godfather Part III*. USA: Paramount Pictures Studios, 1990.

Coppola, Sofia, *Marie Antoinette*. USA: Sony Pictures Releasing, 2006.

Crosland, Alan, *The Jazz Singer*. USA: Warner Bros. Pictures, 1927.

Curtiz, Michael, *Captain Blood*. USA: Warner Bros. Pictures, 1935.

———, *The Sea Hawk*. USA: Warner Bros. Pictures, 1940.

———, *Yankee Doodle Dandy*. USA: Warner Bros. Pictures, 1942.

Daldry, Stephen, *The Hours*. USA: Paramount Pictures, 2002.

DeMille, Cecil B., *The Ten Commandments*. USA: Paramount Pictures, 1956.

Demy, Jacques, *The Umbrellas of Cherbourg*. France: 20th Century Fox, 1964.

Disney, Walt, and Ub Iwerks, *Steamboat Willie*. USA: Celebrity Productions, 1928.

Donner, Richard, *Superman*. USA: Warner Bros. Pictures, 1978.

Ducournau, Julia, *Raw*. France: Wild Bunch, 2016.

Eisenstein, Sergei, *Alexander Nevsky*. USSR: Mosfilm, 1938.

———, *Battleship Potemkin*. USSR: Goskino, 1925.

Fincher, David, *The Social Network*. USA: Sony Pictures Releasing, 2010.

Fleming, Victor, *Gone with the Wind*. USA: Metro-Goldwyn-Mayer, 1939.

———, *The Wizard of Oz*. USA: Metro-Goldwyn-Mayer, 1939.

Forman, Miloš, *Amadeus*. USA: Orion Pictures, 1984.

Gray, F. Gary, *Straight Outta Compton*. USA: Universal Pictures, 2015.

Green, David Gordon, *Halloween*. USA: Universal Pictures, 2018.

Grierson, Alister, *Kokoda*. Australia: Palace Films, 2006.

Gunn, James, *Guardians of the Galaxy*. USA: Walt Disney Studios Motion Pictures, 2014.

Hahn, Don, Pixote Hunt, Hendel Butoy, Eric Goldberg, James Algar, Francis Glebas, Paul and Gaëtan Brizzi, *Fantasia 2000*. USA: Buena Vista Pictures, 1999.

Hawks, Howard, *Gentlemen Prefer Blondes*. USA: 20th Century Fox, 1953.

Hitchcock, Alfred, *Lifeboat*. USA: 20th Century Fox, 1944.

———, *Mr. & Mrs. Smith*. USA: RKO Radio Pictures, 1941.

———, *North by Northwest*. USA: Metro-Goldwyn-Mayer, 1959.

———, *Psycho*. USA: Paramount Pictures, 1960.

———, *Rear Window*. USA: Paramount Pictures, 1954.

———, *Spellbound*. USA: United Artists, 1945.

———, *Vertigo*. USA: Paramount Pictures, 1958.

Hooper, Tom, *Les Misérables*. UK: Universal Pictures, 2012.

Huston, John, *The Maltese Falcon*. USA: Warner Bros., 1941.

Jackson, Peter, *The Lord of the Rings: The Fellowship of the Ring*. USA: New Line Cinema, 2001.

Joffé, Roland, *The Mission*. UK: Columbia-Cannon-Warner Distributors, 1986.

Joon-ho, Bong, *Parasite*. South Korea: CJ Entertainment, 2019.

Juran, Nathan, *The 7th Voyage of Sinbad*. USA: Columbia Pictures, 1958.

Kalatozov, Mikhail, *I am Cuba*. Cuba: Mosfilm, 1964.

Kar-wai, Wong, *In the Mood for Love*. China: Block 2 Pictures, 2000.

Kelly, Gene, and Stanley Donen, *Singin' in the Rain*. USA: Metro-Goldwyn-Mayer 1952.

Kieślowski, Krzysztof, *Three Colours: Blue*. France: MK2 Diffusion, 1994.

Kozintsev, Grigori, *Hamlet*. USSR: Lenfilm, 1964.

Kramer, Stanley, *It’s a Mad, Mad, Mad, Mad World*. USA: United Artists, 1963.

Kubrick, Stanley, *2001: A Space Odyssey*. USA: Metro-Goldwyn-Mayer, 1968.

———, *Eyes Wide Shut*. USA: Warner Bros. Pictures, 1999.

———, *Lolita*. USA: Metro-Goldwyn-Mayer, 1962.

Lanfield, Sidney, *You'll Never Get Rich*. USA: Columbia Pictures, 1941.

Leone, Sergio, *Once Upon a Time in America*. USA: Warner Bros. Pictures, 1984.

Lucas, George, *Star Wars*. USA: 20th Century Fox, 1977.

Luhrmann, Baz, *The Great Gatsby*. USA: Warner Bros. Pictures, 2013.

Majidi, Majid, *The Color of Paradise*. Iran: Varahonar Company, 1999.

Malick, Terrence, *A Hidden Life*. USA: Searchlight Pictures, 2019.

Marshall, George, Henry Hathaway, and John Ford, *How the West was Won*. USA: Metro-Goldwyn-Mayer, 1962.

Mayo, Archie, *A Night in Casablanca*. USA: United Artists, 1946.

McLeod, Norman Z., *Horse Feathers*. USA: Paramount Pictures Studios, 1932.

McTiernan, John, *Predator*. USA: 20th Century Fox, 1987.

Mendes, Sam, *1917*. UK: Universal Pictures, 2019.

———, *American Beauty*. USA: DreamWorks Pictures, 1999.

Noé, Gaspar, *Enter the Void*. France: Wild Bunch Distribution, 2009.

Olivier, Laurence, *Hamlet*. UK: Paramount Pictures Studios, 1948.

Pal, George, *The Time Machine*. USA: Metro-Goldwyn-Mayer, 1960.

Palma, Brian De, *Carrie*. USA: United Artists, 1976.

Petersen, Wolfgang, *Troy*. USA: Warner Bros. Pictures, 2004.

Pollack, Sydney, *Out of Africa*. USA: Universal Pictures, 1985.

Pressburger, Emeric, and Michael Powell, *The Red Shoes*. UK: General Film Distributors, 1948.

Rapper, Irving, *Rhapsody in Blue*. USA: Warner Bros. Pictures, 1945.

Ray, Nicholas, *Rebel Without a Cause*. USA: Warner Bros. Pictures, 1955.

Reed, Carol, *The Third Man*. UK: British Lion Film Corporation, 1949.

Refn, Nicolas Winding, *The Neon Demon*. USA: Amazon Studios; Broad Green Pictures, 2016.

Riefenstahl, Leni, *Olympia*. Germany: Tobis, 1938.

Scorsese, Martin, *Taxi Driver*. USA: Columbia Pictures, 1976.

Scott, Ridley, *Alien*. USA: 20th Century Fox, 1979.

———, *Blade Runner*. USA: Warner Bros. Pictures, 1982.

———, *Gladiator*. USA: DreamWorks Pictures, 2000.

Spielberg, Steven, *Catch Me If You Can*. USA: DreamWorks Pictures, 2002.

———, *Close Enocunters of the Third Kind*. USA: Columbia Pictures, 1977.

———, *Jaws*. USA: Universal Pictures, 1975.

———, *Jurassic Park*. USA: Universal Pictures, 1993.

———, *Saving Private Ryan*. USA: DreamWorks Pictures, 1998.

———, *Schindler's List*. USA: Universal Pictures, 1993.

———, *West Side Story*. USA: 20th Century Studios, 2021.

Stevens, George, *Swing Time*. USA: RKO Pictures, 1936.

Stone, Oliver, *JFK*. USA: Warner Bros. Pictures, 1991.

Tarantino, Quentin, *Reservoir Dogs*. USA: Miramax Films, 1992.

Tarr, Béla, and Ágnes Hranitzky, *The Turin Horse*. Hungary: The Cinema Guild, 2011.

Thornton, Warwick, *Samson and Delilah*. Australia: Madman Entertainment, 2009.

Tykwer, Tom, *Run Lola Run*. Germany: Prokino Filmverleih, 1998.

Varda, Agnès, *Cleo from 5 to 7*. France: Athos Films; Ciné-Tamaris, 1962.

Vidor, Charles, *Gilda*. USA: Columbia Pictures, 1946.

Vigo, Jean, *L'Atalante*. France: Gaumont, 1934.

Watts, Jon, *Spider-man: Far From Home*. USA: Sony Pictures, 2019.

Weir, Peter, *Picnic at Hanging Rock*. Australia: B.E.F. Film Distributors, 1975.

Welles, Orson, *Citizen Kane*. USA: RKO Radio Pictures, 1941.

Wise, Robert, *The Sound of Music*. USA: 20th Century Fox, 1965.

Wood, Sam, *King's Row*. USA: Warner Bros. Pictures, 1942.

———, *A Night at the Opera*. USA: Metro-Goldwyn-Mayer, 1935.

Wyler, William, *Ben-Hur*. USA: Metro-Goldwyn-Mayer, 1959.

———, *The Best Years of Our Lives*. USA: RKO Radio Pictures, 1946.