Narrativity, Worldmaking, and Recorded Popular Song

by

Alexander C. Harden

Submitted for the Degree of Doctor of Philosophy

Department of Music and Media

Faculty of Arts and Social Sciences

Supervisors: Professor Allan F. Moore, Dr. Nicholas Reyland,
Dr. Jeremy Barham, Dr. Timothy Hughes
Statement of Originality

This thesis and the work to which it refers are the results of my own efforts. Any ideas, data, images or text resulting from the work of others (whether published or unpublished) are fully identified as such within the work and attributed to their originator in the text, bibliography or in footnotes. This thesis has not been submitted in whole or in part for any other academic degree or professional qualification. I agree that the University has the right to submit my work to the plagiarism detection service TurnitinUK for originality checks. Whether or not drafts have been so-assessed, the University reserves the right to require an electronic version of the final document (as submitted) for assessment as above.

Alexander C. Harden
January 2018
Summary

This thesis develops a novel approach to hermeneutical popular song analysis through the application of cognitive narrative theory. Throughout, I develop a reader-oriented model of musical interpretation in which a listener may interpret a song as a story situated within a mental model of a possible world, or *narrative world*. As narrative worlds are principally discussed in relation to non-musical texts, this project synthesises and expands existing work to develop new strategies of analysing verisimilitude, diegetic framing, and perspectivity in recorded popular song. Thereafter, I integrate these positions in a narrative reading of The Who’s *Quadrophenia*.

A theoretical basis is proposed in Chapters 1 and 2. Chapter 1 begins by exploring existing applications of narrative theory to Western art music. Thereafter, narrative is redefined as a cognitive construction resulting from the listener’s *narrativisation* of a work. It is argued that narrativisation is based on the phonographic representations of the singer and other sound-sources in recorded popular song. Using the theory of affordance from ecological perception theory, Chapter 2 proposes a model in which recordings offer listeners ways of constructing narrative worlds using information from given tracks and supplementary external schemata.

Analytical approaches to properties of narrative worlds are discussed in Chapters 3-5. Chapter 3 addresses verisimilitude resulting from the suggested logic of the narrative world, representation of sound-sources within the phonographic space, and details of the performance. Chapter 4 explores temporal aspects of narrative worlds, and the ontology of sound-events. Chapter 5 investigates matters of perspectivity in relation to the narrator/protagonist and the implied position of the
listener in relation to the narrative world. Finally, Chapter 6 provides an extended example of these analytical strategies in practice in a case study of *Quadrophenia*. 
Acknowledgements

I would like to begin by thanking Professor Allan Moore for his insight throughout this project, and to Dr. Tim Hughes for his support at its early stages. I am also grateful to Dr. Jeremy Barham and Dr. Nicholas Reyland for their kind advice and feedback as this thesis took shape. I would further like to highlight the advice, enthusiasm, and encouragement from Dawn, Sam, and Carol from the Researcher Development Programme, as well as the rest of the University of Surrey Doctoral College. Finally, thank you to my friends and family, without whom I could not have finished this thesis.

This project was funded with the support of the University of Surrey FAHS Studentship Competition.
Contents

0. General Introduction ........................................................................................................ 1
  0.1 Narratives and Music .................................................................................................... 1
  0.2 Definition of Terms ...................................................................................................... 2
    0.2.1 Narrative .............................................................................................................. 2
    0.2.2 Narrativity ............................................................................................................ 3
    0.2.3 (Cognitive) Narratology ....................................................................................... 3
    0.2.4 Narrative Worldmaking ......................................................................................... 3
  0.3 Hypotheses and Key Issues .......................................................................................... 4
    0.3.1 Recorded Popular Songs and Narrativity ............................................................ 5
    0.3.2 Recorded Popular Songs and Narrative Worldmaking ......................................... 7
    0.3.3 Narrative Worlds and Accommodation ................................................................ 9
  0.4 Approach ..................................................................................................................... 9
  0.5 Content ....................................................................................................................... 11
    0.5.1 Chapter 1: Narrativity and Narrativisation ......................................................... 12
    0.5.2 Chapter 2: Narrative Worlds ................................................................................. 12
    0.5.3 Chapter 3: Verisimilitude and Departure .............................................................. 13
    0.5.4 Chapter 4: Diegetic Framing and Temporal Sequencing ....................................... 13
    0.5.5 Chapter 5: Perspectivity ....................................................................................... 14
    0.5.6 Chapter 6: Interpretation of The Who’s Quadrophenia ......................................... 15

1. Narrativity and Narrativisation ...................................................................................... 16
  1.1 Narrative and Drama in Western Art Music ................................................................. 17
    1.1.1 Mimesis and Diegesis .......................................................................................... 21
    1.1.2 Narrativisation .................................................................................................... 27
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2 Recorded Popular Song and Narrativity</td>
<td>30</td>
</tr>
<tr>
<td>1.2.1 Recorded Music and Narrativity</td>
<td>36</td>
</tr>
<tr>
<td>1.2.2 Plurimediality</td>
<td>40</td>
</tr>
<tr>
<td>1.3 Case Study: Ke$ha - “Stephen”</td>
<td>42</td>
</tr>
<tr>
<td>1.3.1 Narrativising “Stephen”</td>
<td>44</td>
</tr>
<tr>
<td>1.4 Conclusion</td>
<td>48</td>
</tr>
<tr>
<td>2. Narrative Worlds</td>
<td>50</td>
</tr>
<tr>
<td>2.1 Narrative Agents</td>
<td>51</td>
</tr>
<tr>
<td>2.1.1 Narrative Agents in Song</td>
<td>53</td>
</tr>
<tr>
<td>2.2 A World Model</td>
<td>59</td>
</tr>
<tr>
<td>2.2.1 Reference Worldview</td>
<td>62</td>
</tr>
<tr>
<td>2.2.2 Phonographic Environment</td>
<td>64</td>
</tr>
<tr>
<td>2.2.3 Narrative World</td>
<td>64</td>
</tr>
<tr>
<td>2.3 Narrative Worldmaking and Affordance</td>
<td>65</td>
</tr>
<tr>
<td>2.3.1 Subject Position</td>
<td>66</td>
</tr>
<tr>
<td>2.3.2 Affordance</td>
<td>67</td>
</tr>
<tr>
<td>2.4 Worldmaking in Radiophonic Drama</td>
<td>70</td>
</tr>
<tr>
<td>2.4.1 Analytical Concepts</td>
<td>72</td>
</tr>
<tr>
<td>2.5 Accommodation</td>
<td>76</td>
</tr>
<tr>
<td>2.6 Case Study: The Kinks - “Waterloo Sunset”</td>
<td>80</td>
</tr>
<tr>
<td>2.6.1 Setting</td>
<td>83</td>
</tr>
<tr>
<td>2.6.2 Characterisation</td>
<td>83</td>
</tr>
<tr>
<td>2.6.3 Summary</td>
<td>84</td>
</tr>
<tr>
<td>2.7 Conclusion</td>
<td>86</td>
</tr>
<tr>
<td>3. Verisimilitude and Departure</td>
<td>88</td>
</tr>
</tbody>
</table>
3.1 Mimesis and Minimal Departure .......................................................... 90
3.1.1 External Schemata and Minimal Departure ........................................ 92
3.2 Oneiric Aesthetics and Departure ......................................................... 93
3.3 Approaches to Musical Verisimilitude ................................................... 99
3.3.1 Beyond Typologies of Verisimilitude ................................................. 102
3.4 Four Aspects of Verisimilitude ............................................................ 105
3.4.1 Persona and Narrative Logic ............................................................. 106
3.4.2 Landscape and Source-Bonding ....................................................... 110
3.4.3 Correspondence to a Live Performance ............................................. 113
3.4.4 Acoustical Space and Staging ........................................................... 116
3.4.5 Blended Spaces ............................................................................. 118
3.5 Case Study: Madonna - “Hollywood” ................................................ 121
3.5.1 Acoustical Space and Staging ......................................................... 125
3.5.2 Landscape and Source-Bonding ....................................................... 127
3.5.3 Performative Agency ..................................................................... 128
3.5.4 Persona ....................................................................................... 129
3.6 Conclusion ......................................................................................... 130

4. Diegetic Framing and Temporal Sequencing ........................................ 133
4.1 Temporality ....................................................................................... 134
4.1.1 Time and Discourse ...................................................................... 136
4.1.2 Time and Story ............................................................................ 138
4.1.3 Narrator Time ............................................................................. 140
4.2 Song Scenes ....................................................................................... 142
4.3 Diegetic Functions of Music ............................................................... 146
4.3.1 Sonic Diegeses ........................................................................... 148
4.4 Three Frames in Recorded Popular Song ............................................. 153
4.4.1 Diegetic Paramusical Frame ........................................................... 155
6.1 A Narrative Interpretation of Quadrophenia ........................................ 207
  6.1.1 Opening ......................................................................................... 208
  6.1.2 Leaving Home ............................................................................... 214
  6.1.3 A New Life ..................................................................................... 222
  6.1.4 Visiting The Seafront ................................................................. 227
  6.1.5 Seeking Refuge ............................................................................. 232

6.2 The World of Quadrophenia ............................................................. 237
  6.2.1 Narrative Summary ................................................................. 240
  6.2.2 The Prose Narrative ................................................................. 241

6.3 Conclusion .......................................................................................... 242

7. General Conclusion ............................................................................. 247
  7.1 Summary ......................................................................................... 247
  7.1.1 So What? ...................................................................................... 251

  7.2 Future Opportunities in Popular Musicological Narratology .......... 253

References ............................................................................................. 257

Discography ............................................................................................ 279

Appendix: Glossary of Key Terms ....................................................... 286
0. General Introduction

0.1 Narratives and Music

As Roland Barthes argued in his *Introduction to the Structural Analysis of Narratives*, narrative — that is, the representation of real or fictional events — “is present at all times, in all places, in all societies [...] there is not, there has never been anywhere, any people without narrative; all classes, all human groups, have their stories, and very often those stories are enjoyed by men of different and even opposite cultural backgrounds [...] Like life itself, it is there, international, transhistorical, transcultural” (1975: 237). The ubiquity and necessity of narrative has been argued elsewhere: Abbott (2008), for instance, contends that narrative is central to our understanding of life and individual biography, whilst Herman (2009a) develops the thesis, albeit cautiously, that human consciousness is structured through narrative (expanded in Herman 2013).

Barthes later comes to provide a substantial list of media which he contends can tell stories, including literary forms (such as tales, short stories, and epics), oral storytelling (such as myth, legend, and conversation), and several forms of visual art (movies, painting, and stained glass) (*ibid.:* 238). Whilst literary and visual arts are well-represented in Barthes’ list, both music and song are conspicuous in their absence. Yet, the medium of song would appear to be a promising candidate, for instances often involve the telling of a character’s experiences; the telling of a story from the point of view of a possible individual. Despite the relevance of narrative theory to the
interpretations of songs as a form of storytelling, narratological concepts feature somewhat infrequently in popular musicological scholarship.¹

To address the gap in research identified above, this project develops a cognitive narratological approach to hermeneutical popular song analysis. In doing so, I demonstrate that songs can be interpreted in narrative ways and explore various aspects of the possible worlds which such narrative interpretations are situated within. The hypotheses and key issues involved in this project are outlined below, following introductory definitions of the key concepts which we will encounter. A glossary containing concise explanations of key terms used throughout this thesis is also included in the Appendix.

0.2 Definition of Terms

0.2.1 Narrative

As stated, narrative — broadly — refers to a representation of a sequence of events. Typically, this takes the form of written text and, accordingly, a dominant understanding of the term is the text itself. This understanding of narrative, though, presupposes narrative meaning as an encoded property of a given text. In this context, the text can be understood as a representation of some intended meaning, yet this neglects the role of the interpreter and cannot easily resolve discrepancies between

¹ Throughout this project, we will draw from hermeneutical writing (such as Moore, 2012b; Tagg, 2013) which does not extensively employ narratological concepts per se but nevertheless supports the discussion of popular music in relation to narrative. This project also employs other contributions that address aspects of narrativity more explicitly (e.g. Lacasse, 2006; Liu-Rosenbaum, 2012; Negus, 2012a, 2012b; Nicholls, 2004, 2007), which are introduced in Chapter 1.
readings by different individuals. Instead, I take an alternative position that narrative is the subjective representation of events in consciousness, informed by a text.

0.2.2 Narrativity

If narrative refers to the mental representation of events, narrativity refers to the property of evoking a narrative. How one understands narrative thus impacts on the particular understanding of narrativity: if narrative meaning is modelled as a linguistic telling of events then music cannot have narrativity, as it lacks the same semantic ability of language. The understanding of narrative as a representation of events in consciousness, however, allows us to think of music, song, and a number of other non-linguistic media as having narrativity.

0.2.3 (Cognitive) Narratology

Narratology, also referred to as ‘narrative theory,’ is the study of narratives. We will also encounter various work in the field of cognitive narratology, which is a branch of narrative theory that investigates aspects of individuals’ comprehension of stories, often from reception-oriented theoretical positions.

0.2.4 Narrative Worldmaking

An important contribution of cognitive narratology is the recognition that narrative interpretation requires the interpreter to form a mental model of a possible world inhabited by the agents of a narrative. This model provides a framework in which characters’ experiences can be understood.² Cognitive narratologists use the terms

---

² As Herman puts it, narratives act as “blueprints for a specific mode of world-creation” (2009a: 105). This is revised by Marie-Laure Ryan and Jan-Noël Thon in terms of “world
storyworld or narrative world to refer to this mental model, which may reflect (or depart from) an actual world to varying degrees. The interpreter’s formation and deformation of a narrative world whilst encountering a narrative is hereafter referred to as narrative worldmaking.3

0.3 Hypotheses and Key Issues

This project addresses narrativity and worldmaking in relation to recorded popular song. I construe the phrase ‘popular song’ broadly to include a range of pop, rock, and electronic styles that are conceived through electronic music technology. I will draw from a range of popular songs, which involve performed lyrics alongside instrumental musical elements, both in recorded format. I refer to ‘recorded popular songs’ (or ‘tracks’) to emphasise the phonographic nature of these texts for two reasons. Pragmatically, the fixity of recorded music permits readers to form their own views from the same recorded artefact. Secondly, the model of narrativity which I develop in Chapter 2 argues that the composition and development of the phonographic environment is an important aspect of the ways in which a track might prompt the imagining of a narrative world by the listener and distinguishes it from live popular song. Our investigations here will principally concern narrative interpretation on the level of individual tracks, with the exception of Chapter 6, which develops an extended case study across The Who’s Quadrophenia album.

3 This should not be confused with Covach’s (1994) idea of musical worlding.
0.3.1 Recorded Popular Songs and Narrativity

The overarching hypothesis of this project is that recorded popular songs may be interpreted in a narrative way. It is crucial here to capture the importance of both listener and text. Beginning with the latter, Ryan and Thon observe that “stories and their worlds are crucially shaped by the affordances [i.e. possibilities to evoke meaning] and limitations of the media in which they are realized” (2014a: 2). The narrative analysis of recorded popular song should, therefore, reflect the affordances of the recorded voice, other sound-sources, and the ways in which their presentation is informed by the recording studio. This is supported by a growing body of work which explores the affective or hermeneutic implications of record production. However, the study of phonographic aspects relating to narrativity has received

---

4 Elsewhere, Ryan has called for “a comprehensive and widely accepted theory of the importance of the medium as material support for the form and content of message” (2004: 22), which is also reflected in Herman’s (2009a) idea of situatedness to describe the effects of communicative context on narrativity.

5 Lacasse (2006) emphasises in particular that narrative studies of popular song must account for the dominant means of its conception and distribution, i.e. as a phonographic medium.

6 Simon Zagorski-Thomas (2007) has proposed an outline of what he calls ‘the musicology of record production,’ which he expands considerably in his 2014 monograph of the same name. As Zagorski-Thomas notes, the original article is prompted in part by a more cautionary conference paper (published later by Moore), which argues that it is essential to integrate aspects of record production with musical details; that “we cannot afford a musicology of production any more than we can afford a separately located musicology of the voice [...] timbre or harmony, or a verbology [...] Our interpretative apparatus has to be all-inclusive” (Moore 2012a: 111).

With respect to the role of the listener, I would like to acknowledge Kennett’s (2003) and Negus’ (2012b) assured arguments for the role of the listener and listening context in musical interpretation. Beginning from a critique of the arbitrariness of semiotic approaches and their lack of attention to the listening situation or listener demographic, Kennett pursues what he calls a ‘cultural-acoustic model of functional music analysis’ (ibid.: 207). Kennett provides a hypothetical scenario involving characters at a wine store who exhibit different listening abilities, stylistic familiarity, and attitudes towards the music. Their responses to the music vary as, arguably, each listening constitutes a different ‘text’. Negus, on the other hand, develops a model in which musical meaning is constructed by listeners through negotiation of texts and their contexts. He supports this with a report of a student group’s unfolding interpretation of Steely Dan’s “Kid Charlemagne” that emphasises the way in which interpretations developed through interactions between members of the group, supported by additional contextual details, as we might also observe in online forums dedicated to song interpretation. Whereas both models seek to highlight the role of listeners, Kennett’s relativistic approach is problematic, as it does not resolve the arbitrariness which he critiques and would require considerable profiling of listeners.

7 Alan Williams (2010) has also emphasised the general subordination of producers or recordists to the authority often applied to artists, particularly in relation to rock music, though the issue of record production and musical meaning has received some notable discussion beyond the remit of Williams’ article (to name a few: Zak 2001; Moorefield 2005; Bennett 2015; Warner 2003; Zagorski-Thomas 2014; Frith and Zagorski-Thomas 2012). The Association for the Art of Record Production was also formally established in 2009 and has maintained a regular series of conferences and a dedicated journal.
Meanwhile, Negus’ approach diverges further from the idea of a recording as text and focusses on song meanings in relation to society or groups of individuals rather than the experiences relayed in a song.

With Negus’ and Kennett’s work in mind, I would like to propose that narrative meanings are not encoded in a text, but rather emerge from our engagement with texts and their content, which may or may not be designed to prompt narrative interpretations. For one, this does not presuppose correspondence between authorial intentions — which are not possible to confirm — and encoded meaning. Secondly, this recognises the interpretative authority of listeners and the variety of readings or personal significances which songs may evoke for them. Using principles derived from cognitive narratology, my theoretical basis is intended to address the need for a model sufficiently malleable to accommodate the different subjective positions which come from different listeners, in order to allow for the possibility — and indeed, likelihood — of different readings by different listeners. Furthermore, this provides a model which regards the listener as central to the act of interpretation without invalidating the significance of the text itself. We will explore this line of enquiry in Chapters 1 and 2.

0.3.2 Recorded Popular Songs and Narrative Worldmaking

The second hypothesis which guides this project is that recorded popular song affords to the listener ways of imagining a narrative world informed by the perspective(s) of narrative agents and supplemented by external schemata, i.e.

---

8 Although heavily influenced by semiotic approaches himself, Nattiez (1990a: 16-19), who critiques the notion of music being a narrative, also argues that musical meaning is created through a negotiation between interpreter and text.
learned representations of objects and concepts external to song. This hypothesis is
drawn most explicitly from accounts by Jean-Jacques Nattiez and Paul Ricœur.
Within his discussion of signification in music, Nattiez makes the following claim
regarding how meaning is found in an object:

an object of any kind takes on meaning for an individual apprehending that object, as
soon as the individual places the object in relation to areas of his [or her] lived
experience — that is, in relation to a collection of other objects that belong to his or her
experience of the world (1990b: 9).

Nattiez purposefully writes in a general way here and we might apply this thinking to
any form of meaning, musical or otherwise. His claim that meaning (and hence, the
act of interpretation) is contingent upon an individual’s experience of a world,
however, emphasises the importance of external schemata in the art of interpretation.

Complementing the basic claim made by Nattiez, above, Ricœur articulates
another view which instead addresses the world evoked by a text. He writes:

What is indeed to be understood – and consequently appropriated – in a text? Not the
intention of the author, which is supposed to be hidden behind the text; not the
historical situation common to the author and his original readers; not the expectations
or feelings of these original readers; not even their understanding of themselves and
historical and cultural phenomena. [...] What has to be appropriated is nothing other
than the power of disclosing a world that constitutes the reference of the text. In this way
we are as far as possible from the Romanticist ideal of coinciding with a foreign psyche.
If we may be said to coincide with anything, it is not the inner life of another ego, but the
disclosure of a possible way of looking at things, which is the genuine referential power
of the text (Ricœur 1976: 92).
With this, Ricœur points to what he suggests is a crucial interpretative trajectory: the disclosure of a worldview articulated by the text. Ricœur’s claim therefore supports the hypothesis that texts prompt listeners to imagine a narrative world.

0.3.3 Narrative Worlds and Accommodation

The third hypothesis which guides this project is that narrative worlds can afford to the listener a — temporary — sense of immersion. This may prompt an empathetic engagement with narrative characters, and/or a sense of being relocated from one’s actual situational context. This sensation of immersion is variously described by cognitive narratologists as “transportation” (Gerrig 1993), “accommodation” (Herman 2009a), “recentering” (Wolf 2004), and “deictic shift” (Zubin and Hewitt 1995). Various phrases which are commonly used to describe emotional engagement with songs suggest that they also engender a sense of relocation. Expressions such as “it moved me”, “it takes me back”, or “lose yourself to music”, all incorporate some notion of movement or dislocation from the actual situational context in which listening takes place. Depending on aspects such as the lyrics of a song, we could perhaps hear the singer as directly addressing us and imagine ourselves as his/her addressee. Alternatively, a song might capture experience or emotion in a way which affords the sensation of imagining ourselves in the position of the protagonist. Both these cases involve the construction of a narrative world which accommodates us.

0.4 Approach

In this project, my attention will be directed towards three generic aspects of narrative worlds that can be revealed through analytically informed discussion. The first, verisimilitude, involves mimetic strategies which imitate actual experiences. Basic examples of this might include the simulation of actual spaces or media (such
as a telephone call) and the use of non-instrumental sounds with recognisable, everyday sources. These strategies can be found in a variety of recorded popular songs, as well as several other narrative forms and genres, to present experiences as though they were real and could, or did, happen. Accordingly, we might commonsensically imagine narrative worlds after our own social experiences external to the work. In many cases, though, other details of the narrative or the discourse from which it is interpreted might point to greater degrees of departure. We will explore this in relation to performance, production, and experiences depicted in a track throughout Chapter 3.

A second aspect we will explore is the temporal and diegetic architecture of recorded popular songs by discussing the role of time in event sequencing on the level of the narrative world, as distinguished from time as it relates to the narrator or the discourse. By 'diegetic architecture,' I refer to the way in which sound may appear to occupy or move between different degrees of separation from the narrative world. We will also discuss the subsequent issues this raises in relation to the locations of the mimetically projected singer and musicians throughout Chapter 4.

The final aspect we will consider can be coarsely described as ‘perspectivity’. In narrative fiction, we encounter narrators with varying access to information relating to the world at hand and its inhabitants (whether they are omniscient, or report from the perceptions of themselves/others) (Genette 1980; Bal 2009). As the character enacted by the singer provides an important entrance into narrative worldmaking in relation to songs, the equivalent perspective which they report from is crucial in imagining the world they appear to inhabit. In recorded songs, a second form of perspective is constructed by the configuration of sound-sources in relation to the listener's implied point of listening (as we might observe visually in a film, for instance). This introduces the possibility for the discourse to provide not only ways
of worldmaking, but also to afford the listener's imagining of an implicit location in relation to the narrative world.

This investigation is based on a reception-oriented model in which narrativity is regarded as an individual's response to the text. In this respect, narrativity is informed by — but not encoded entirely within — a text. This approach avoids a problem of transmission-oriented alternatives, which cannot satisfactorily account for how songs are interpreted in practice: whilst interpretations often cohere to some extent between listeners, there is the possibility of considerable variety here, particularly when encountering tracks under different personal circumstances. Indeed, listeners' interpretations may change over time. For this to be the case, musical meaning cannot be wholly encoded within and transmitted by a recording in the way in which transmission-oriented approaches presuppose. Rather, we should acknowledge the listener as an active agent in meaning-making.

0.5 Content

This thesis is organised in three main parts. In Chapters 1 and 2, I discuss issues of narrativity in pursuit of a general model that provides the basis for the chapters to follow. In Chapters 3-5, I develop analytical strategies which allow us to discuss ways in which recorded popular song informs particular ways of narrativisation and narrative worldmaking in relation to three aspects of narrative worlds. Finally, in

9 The prominent semiotic method of popular song analysis developed by Phillip Tagg (2013), as one example, involves parameters of codal (in)competency and codal interference to account for listeners' misconstrued meaning. In doing so, this approach compromises the interpretative agency of the listener and suggests there are a finite number of 'correct' interpretations.
Chapter 6, I provide an extended case study of The Who’s _Quadrophenia_ to provide a practical demonstration of the different analytical strategies discussed.

### 0.5.1 Chapter 1: Narrativity and Narrativisation

In the field of cognitive narratology, several approaches locate narrative meaning in relation to a reader’s engagement with a text. In Chapter 1, I explore existing work in relation to narrative in literary and musical contexts, before adopting a cognitive position after Monika Fludernik’s concept of _narrativisation_ (2010) and Marie-Laure Ryan’s (2004) related writing on texts “having narrativity.” In short, both ideas propose that narrativity can be thought of as a matter of the reader’s experience when interpreting the text, rather than being an encoded property of the text. With the case study of “Stephen” by Ke$ha, I discuss how the track can be heard in a narrative way — that is, narrativised — with attention to the performed lyrics and role of recorded instrumental material.

### 0.5.2 Chapter 2: Narrative Worlds

Having developed the position in Chapter 1 that listeners may narrativise recorded popular song, Chapter 2 develops a world-based model of narrative. Using the case of the singer and listener, six agents involved in narrativisation are identified: the _real singer_ and _real listener_; _persona_ and _implied listener_; and, _protagonist/narrator_ and _addressee_. These pairings are positioned in relation to three worlds: a _reference worldview_, _phonographic environment_ (or ‘soundworld’), and _narrative world_, respectively. In Chapter 2, I also introduce the concept of affordance to help describe the role of the text in informing narrative worldmaking. Alongside the semantic role of lyrics in recorded popular song, I use the example of the radio play “The Revenge” to demonstrate the importance of phonographic material in affording ways of
imagining a world. The chapter concludes with a case study of The Kinks’ “Waterloo Sunset” to examine how the track evokes a romanticised narrative world in relation to the actual setting of Waterloo during the 1960s.

0.5.3 Chapter 3: Verisimilitude and Departure

A principal condition of a narrative world is the extent to which it is perceived to correspond to an actual world, for this determines the ways in which interpreters may negotiate the narrative world and the reliability of inferences that one can make based upon it. With the example of Aqua’s “Halloween”, I propose that recorded popular song often incorporates oneiric or ‘dream like’ elements through irrealist approaches to aspects such as acoustical space or combinations of instruments. Using the model of a recording as a form of virtual reality, this chapter discusses devices relating to four aspects of tracks which may encourage a listener to nevertheless construct narrative worlds with the resemblance of truth: the persona and situation/events reported; acoustical space; the selection and combination of sound-sources; and, correspondence to an actual performance. These domains are explored together in a closing case study of “Hollywood” which investigates how the track is constructed in a way that evokes a high-pressured, artificial social climate within a narrative world.

0.5.4 Chapter 4: Diegetic Framing and Temporal Sequencing

Recognising the oneiric qualities of recorded popular song explored in Chapter 3, it is argued in Chapter 4 that recorded popular songs often mediate space and time in a somewhat fluid manner, though these provide important boundaries to the narrative world. Accordingly, this chapter discusses several musical examples which
distinguish between three possible temporal relationships in recorded popular song: duration of a passage, the temporal setting of the story, and the locus of the narrator. These three temporal relationships correspond with the three worlds identified in Chapter 2. A key distinction arises here between the temporal setting of the story and that of the narrator, which is developed further by considering the ontology of sound-events in relation to the narrative world and how this may change over the course of a track. In a closing case study, I discuss how the selection and staging of sound-sources in Kate Bush’s “Waking the Witch” establishes multiple narrative frames which afford new understandings of the relationship between the protagonist and her witch trial.

0.5.5 Chapter 5: Perspectivity

Having discussed issues of ontology in the previous chapter, in the sense of the relationship between sound-events and the narrative world at the time of the story, Chapter 5 addresses a related issue of perspectivity. This is approached by applying, and adapting where necessary, the theory of focalization, which, in narrative theory, describes a restriction of narrative information available to the narrator. Using various examples, I develop an adaptation of focalization theory to describe the different forms of personae that we may encounter in recorded popular song. Thereafter, I turn to the way in which the construction of a recording also affords to the listener a perspective upon the narrative world, outlining different relationships which the accompaniment may have with the position of the narrating agent. The final case study in this chapter, Eminem’s “Guilty Conscience” featuring Dr. Dre, expands on the focalization of three personae (an omniscient narrator and two factions of the protagonist’s conscience) and the phonographic environment as the implied perspective of the listener shifts throughout the track.
0.5.6 Chapter 6: Interpretation of The Who’s *Quadrophenia*

To provide a practical example of musical analysis which draws on each of the key issues raised in this thesis, my final chapter provides an extended case study of The Who’s *Quadrophenia* album. In this case study, I explore how the album’s tracks interconnect to evoke a large-scale narrative structure. This is then positioned in relation to a prose narrative provided in the accompanying liner notes to explore points of coherence and departure.
1. Narrativity and Narrativisation

Whilst prominent figures such as Barthes have argued for the ubiquity of narrative, attempts to apply narrative theories to the study of recorded popular songs are limited in both number and scope. Over several decades, narrative in relation to Western art music has been variously discussed, although this is a surprisingly young area of investigation in light of the longstanding associations between music and storytelling. As Fred Everett Maus, who was among the first to bring musical narrativity under scrutiny, reports in relation to the experience of music: “professional music theory and analysis have tended to avoid certain large questions” (1988: 56). Since the time of Maus’ publication, the narratological possibilities of musical expression have been widely debated, albeit with mixed results which often suggest that music is more compatible with dramaturgy (that is, the enacting of events) than narration (the telling of events).

In this chapter, we will review literature surrounding dramatic and narrative models of music before considering a reception-oriented approach suitable for recorded popular song. Accordingly, we will explore the nature of the media concerned and what it means to ‘be’ a narrative or, alternatively, to ‘have narrativity’ (Ryan 2004). Although discussions of narrativity in Western art music provide a useful starting point, differences of media necessitate adaptation for a popular musicological context: much existing work concerning narrativity of Western art music focuses on musical notation or performance of instrumental music, whereas recorded popular song incorporates lyrics performed by a persona and is generally conceived and distributed phonographically. As Nicholls (2007) argues, there is potential for further study by considering the interaction between different aspects
As we will see in the following section, work surrounding musical narrativity often either subscribes to a model of narrative as a sequence of events intentionally represented in music or, more commonly, as the result of ascribing story-like developments to a piece. In the position which I set out later in Section 1.1.2, I draw on a cognitive approach aligned closest with the latter. More specifically, I argue that narrativity is instantiated by the listener interpreting music in a narrative way — that is, he/she *narrativises* its content. I would like to begin by exploring the central question “what is narrative?”

### 1.1 Narrative and Drama in Western Art Music

The association of music with narrativity has a lengthy history spanning centuries, particularly in relation to programme music. As Newcomb (1987), Maus (1991), and Meelberg (2006) observe, the temporal nature of music lends itself to being conceived as successions of events and subsequently understood as an unfolding tale. The sustained study of narrativity in music, however, has a relatively short history, which began during a period of interest in the 1970s and 1980s.\(^\text{10}\) Since then, the area has received significant attention by scholars of Western art music, yet this scrutiny has also raised challenges to the notion of musical narrativity. As we will see in this section, this uncertainty reflects several possible understandings of narrative and narrative media.

\(^{10}\) Despite not addressing issues of narrative directly, this discussion was preceded by debates over the meaningfulness of music, particularly in response to Hanslick’s (1891) views on so-called ‘absolute music’.
In its simplest form, the prominent narratologist Gérard Genette regards narrative to be “the representation of a real or fictitious event or series of events by language, and more specifically by written language” (1976: 1). In his other writing, he distinguishes between three common uses of the term: 1) narrative as discourse or text in the form of writing or speech; 2) narrative as the events (real or fictional) which are the subject of the discourse; and, 3) narrative as the act of recounting or telling (1980: 25-27). In order to address these three understandings of the term separately, he subsequently labels these narrative, story, and narrating respectively.\(^\text{11}\)

Clearly similar to Genette’s view of narrative as a written representation of events, Mieke Bal describes a narrative text as “a text in which an agent or subject conveys to an addressee (‘tells’ the reader) a story in a particular medium, such as language, imagery, sound, buildings, or a combination thereof” (2009: 5). Here, Bal’s understanding of narrative is predicated on Genette’s narrating and language is mentioned once again, although not exclusively. Bal also explicitly distinguishes between the fabula, “a series of logically and chronologically related events that are caused or experienced by actors”, and story, the “particular manifestation, inflection, and ‘colouring’ of a fabula” (ibid.: 5).

Further comparison between Bal’s and Genette’s descriptions highlights other noteworthy commonalities. Firstly, they each discuss narrative in relation to events (whether real or fictional), i.e. transitions between states (Bal ibid.). Secondly, the pair describes narrative as a representation of these events in a particular discourse, as seen in Genette’s distinction between narrating and story, and Bal’s distinction between

\(^{11}\) Elsewhere, these terms may be referred to using Genette’s original French: récit, histoire, and narration, respectively.
fabula and story. Finally, the two both emphasise narration or the act of telling to a reader.

In *New Sounds, New Stories*, Vincent Meelberg adapts Bal’s narrative theory and defines narrative as a “representation of a temporal development” (2006: 39). Here, Meelberg retains the emphasis on representation and event sequencing shared by Bal’s and Genette’s accounts, but his definition does not highlight narration.\(^\text{12}\) Although his definition is particularly concise and defended at length, its emphasis on the representation of an event or sequence of events by a text does not address the role of the listener. Negus’ and Kennett’s comments noted in the General Introduction concerning popular music and subjectivity (see Section 0.3.1), therefore, remain unresolved. Meelberg’s position is also vulnerable to some criticism when applied to musical scores, as the particular rendering will vary with different performances.\(^\text{13}\)

In pursuit of a model which acknowledges the role of both the text and the interpreter, this project approaches narrative as a cognitive construction formed in the mind of the interpreter in response to a text. Much of my discussion is based on David Herman’s four ‘basic elements’ of narrative, which, he argues, are theoretically

---

\(^{12}\) Whilst we can read Meelberg’s definition as a terse reduction of Bal’s *story* only, he later explicitly endorses her *fabula/story/narrative* trichotomy (*ibid.*: 43).

\(^{13}\) It should be noted that Meelberg’s sources also include recordings. He addresses the transformative effects of performance with reference to the concept of focalization, which I discuss throughout Chapter 5.
possible in a variety of media, including: literature, oral storytelling, and graphic novels (see 2009a). These four principles are:

1) **situatedness**: narratives are set within a particular discursive context which frames its communication and interpretation. This encompasses issues of both narration (i.e. the communication of narrative information) and the broader context in which narratives are consumed;

2) **event sequencing**: a narrative text must represent temporal developments, from which an interpreter can structure a narrative reading of a chronological series of events;

3) **worldmaking/world disruption**: the text prompts the interpreter to imagine a mental representation of a possible world to encompass characters and their situational or temporal/spatial contexts, which is disrupted or reformed as additional information is provided by the text;

4) **what it's like**: engagement with narratives give an impression of immersion within the narrative world or prompt empathy with its characters.

As with Genette and Bal (and indeed, Meelberg), Herman describes narrative in terms of the representation of events. However, his inclusion of situatedness, worldmaking/world disruption, and what it's like indicates an important development from classical literary theory towards a cognitive understanding of narrative — that

---

14 In her introduction to the edited collection *Narrative Across Media*, Ryan (2004) similarly advocates the consideration of narrativity in relation to non-literary media and questions how media informs the possibilities and understanding of narrative.
is, narrative as the cognitive construction built by a subjective interpreter from his/her negotiation with a text.\textsuperscript{15}

Herman’s focus on general properties of narratives from a cognitive perspective allows us to broaden our understanding of narrative media and interrogate the differences between them. Herman’s ‘basic elements’, for instance, are shared by the media which Barthes’ identifies (cited in the General Introduction), many of which are not relayed through language which Genette regards as typical of narratives. In moving away from this received view that narratives involve a text that 	extit{tells} a story, we may include other media that instead appear to 	extit{enact} a story — as I discuss below.

\subsection*{1.1.1 Mimesis and Diegesis}

As we have seen from Genette and Bal, narration (or more broadly, communication) is a defining feature of engagement with narratives. The etymology of this thinking can be seen as far back as ancient Greece, wherein Plato’s third book of \textit{The Republic} also used the presence of the narrator to distinguish narrative from the related practice of drama. In particular, he described three kinds of storytelling: \textit{diegetic}, in which a poet speaks in his/her own voice to relay a story; \textit{mimetic}, which employs an actor who adopts the character and voice of another to enact a story; and, a third, intermediate position which combines or mixes the two.\textsuperscript{16} Most commonly, the first

\footnotetext{15}{Indeed, Herman’s \textit{situatedness} could be linked to the concerns raised by Kennett and Negus for the listening or social (inter)context to musical encounters.}

\footnotetext{16}{A second understanding of the term “diegetic” is used to refer to elements which relate to the base level of a story (the ‘diegesis’). I discuss this further in Chapter 4 (see Section 4.3).}
two categories are cited as a binary that is used to distinguish diegetic narrativity from mimetic dramaturgy.

Due to the fashionable distinction between narrative and drama, much discussion around the possibilities of music to relay some form of story to the listener can be classified in terms of either mimetic or diegetic hypotheses (Reyland 2014). This distinction introduces significantly different interpretative approaches: a mimetic model involves the ascription of characteristics to instruments, as if representing agents in a play, and drawing meanings from their interaction; meanwhile, a diegetic approach considers the music as a narrator instead *telling* a story.¹⁷

Much discourse developing the diegetic hypothesis of music as narrative followed Edward T. Cone’s influential book *The Composer’s Voice*, in which Cone begins with a provocation: “music is a language” and, hence, “if music is a language, then who is speaking?” (Cone 1974: 1). In response, Cone argues that composed musical expression is symbolic and constructs a figurative ‘voice’ which communicates to the listener. However, Cone explicitly models this as a persona constructed through instrumental expression. In an epilogue, he acknowledges that his study principally addresses the musical manifestation of the composer’s persona rather than further hermeneutical issues. He argues that musical expression is deictic in the sense that it is context sensitive: “if music is a language at all, it is a language of gesture” (*ibid.*: 164).¹⁸ Given that Bal and Genette both suggest narrative is primarily

---

¹⁷ Reyland (2014) also describes an intermediate position, in which music is modelled as expressing narrativity in terms of an experiencing consciousness, which he suggests is open to the critiques of both diegetic and mimetic approaches.

¹⁸ Cone draws a parallel in this respect to utterance and gesture in speech, including the communicative ability of context-sensitive utterances such as “oh” and “ah.”
the representation of events by language, Cone’s likening of music and language would seem to support the possibility of music to be a form of narrative, if this comparison is valid. However, Cone’s deictic argument undermines the diegetic hypothesis as he suggests that any content of a musical utterance is debatable without context, which undermines the capacity of his composer’s persona to narrate.

The diegetic hypothesis is explored more explicitly by Maus (1991), who illustrates his discussion with a commentary on Beethoven’s Sonata, Op. 14, No. 1, in which he contends successive musical events lend themselves to being anthropomorphised as a series of actions which together form a story relayed to the listener. Maus’ writing is particularly instructive when, in a parallel to Genette’s narrative/story/narrating differentiation, he addresses the ambiguity of the term ‘narrative’ and its use to describe both story and narration. This observation leads to an extended discussion of how listeners may interpret different temporal orderings in music and so distinguish a musical story from musical narrating.19

Although there is optimistic support for musical narrativity from Maus and his contemporaries (see also Meelberg 2006; Almén 2008), Cone’s comparison between music and language — and diegetic models with it — is contested, most assuredly by Jean-Jacques Nattiez (1990a). As Nattiez correctly observes, instrumental musical expression cannot intrinsically offer the semantic detail of language, which Genette’s understanding of narrative requires. Without either the support of additional

---

19 Maus cautiously suggests that such a distinction could be upheld in cases of breaks in and between movements, which can suggest a decoupling of the temporal setting of the story and the linear flow of the discourse. He further suggests that “the gestural nature of [musical] events can imply orderings that differ from the order in the piece” (ibid.: 28) or in other words, moments of apparent closure may suggest that the flow of a story may not necessary be heard in the same linearity of the musical events which express it.
knowledge about a piece’s supposed meaning, or a texted programme, Nattiez’ argument suggests musical references to narrative aspects (settings, events, motivations, and so on) are entirely deictic and left open to a variety of potentially competing interpretations. In his conclusion, Nattiez forecasts a discouraging outlook for musical narratology: “in itself, and as opposed to a great many linguistic utterances, music is not a narrative [...] any description of its formal structures in terms of narrativity is nothing but superfluous metaphor” (ibid.: 257, italics in original).

In an article from the following year, Lawrence Kramer expresses agreement with Nattiez, and writes “in the strictest sense, there can be no musical narratology” (1991: 143). However, he later becomes more encouraging. Kramer argues that approaches to musical narrativity have begun from an unhelpful view of narrative as giving structure, which he suggests narrative instead resists. Refuting the possibility of musical narrative expression in a literal sense, he makes a case that music often acts as a supplement to narrative in a broader sense. In several examples, Kramer later proposes that the music ‘voices’ agents, which suggests a return to a narrative model. Accordingly, he argues that musical narrativity may arise through forces such as cultural or social meanings, rather than from musical structure.

Further challenge to the diegetic hypothesis is brought by Carolyn Abbate, who argues that music cannot project a “narrating survivor” (1989: 230), which she regards as fundamental to narratives. Secondly, Abbate writes that “like any form of theatre, any temporal art, [music] traps the listener in present experience and the beat of passing time” (1991: 53). Here, her objection is to the limitations that she perceives of music to project a narrative voice and, if music does appear to narrate, that the story unfolds only in the present. As Maus (1991: 22) observes, however, live radio coverage of an unfolding situation takes place in the present but still offers narrativity.
Meelberg (2006: 132-134) also argues that Abbate's predicate of ‘narrating survivors’ does not acknowledge other valid forms of narration.20

Abbate further challenges musical narrativity by arguing that music is chiefly mimetic and, accordingly, that it instead exemplifies drama. As she puts it: “music’s distinction is fundamental and terrible; it is not chiefly diegetic but mimetic” (1991: 53). Alongside such criticisms of diegetic models, several scholars have explored mimetic models of musical narrativity. Maus’ account of musical narrative (cited above), for instance, was preceded by a discussion of musical drama in a passage from the opening of the first movement of Beethoven’s String Quartet, opus 95 (Maus 1988). Reflecting on his analysis, Maus notes the frequent exchanges between technical musicological vocabulary and comments which anthropomorphise musical segments. Through his anthropomorphism, he models the passage in dramatic terms which liken musical relationships and gestures to character statements and actions.

The writing of Anthony Newcomb also supports a dramatic model of musical narrativity. In a discussion of what he refers to as ‘narrative strategies' in Western art music of the late 18th century, he perceives a shared basis between literary narrative event-sequencing and formal aspects of music in that “both can be thought of as a series of functional events in a prescribed order [and] are critically or theoretically derived in the same fashion” (1987: 165). Unlike stricter arguments cited above, however, Newcomb pursues an understanding of what narrative could mean in musical terms, rather than attempting to apply linguistic narrative theory more or less directly to music.21 As such, he continues by arguing that music may enact a plot.

20 Indeed, non-autobiographical narrators are common throughout various narrative forms.
21 A similar philosophy underlines Meelberg’s (2006) larger work on narrative and music, which is founded upon literary narrative theories but negotiates several developments
With Abbate’s comments in mind, the lack of a narrating agent and the necessity of the present tense in enacting a musical plot signals a dramatic approach. Yet, in his later work, Newcomb’s position is concretised when he equates “musically expressive” with “clear elements of narration and agency” (1997: 133). Here, Newcomb conflates narrative and dramatic aspects which signals an important departure from the often-held binary between diegetic and mimetic models, which he develops further through his study of the second movement of Mahler’s Ninth Symphony (ibid.).

In a recent expansion upon Maus’ and Newcomb’s studies, Jerrold Levinson supports the distinction between narrative and drama by arguing against the plausibility of musical narrativity. In a parallel with Genette, Bal, and Herman, Levinson (2004: 431) sees narrativity as predicated upon representation — in contrast to expression or suggestion — which, he argues, requires intentionality. Hence, narrative under Levinson’s understanding must communicate meaning intended by a composer. With Nattiez’ criticism in mind, the limited ability for music to communicate intended meaning explicitly is open to doubt. However, Levinson later makes an important observation that opens an alternative possibility for discussions of musical narrativity. Like Kramer and Maus, he suggests that music can be thought of evoking narrativity in a broader sense “in which narration is effected by means of drama” (2004: 437). With this, Levinson departs from stricter narrative communication models which require an identifiable narrator, whilst remaining compatible with Bal’s observation cited earlier of audio/visual media or buildings as

towards a more functional musical narratology. Whilst Newcomb’s approach challenges a direct application of narrative theory to music, he couches his discussion in relation to literary theories and has also explored the validity of plot paradigms in Mahler’s music after the work of Vladimir Propp on plot archetypes in folk stories.
modes of narration. Crucially, narrative in Levinson’s second sense acknowledges the role of the listener in narrative interpretation and therefore affords the consideration of a cognitive approach to musical narrativity which I shall introduce in the following subsection.

1.1.2 Narrativisation

Many of the models we have seen so far inherit an understanding from the tradition of literary narrative theory that narrativity is dependent on plot, a term used to describe a scheme of events inscribed within a text. In others, writers have resorted to relativistic understandings of narrativity in which a plot is inferred from close attention to musical drama. Both these positions are somewhat problematic: taking plot as inscribed risks overlooking the influence of the interpreter, whilst relativism risks overlooking the text itself. With this in mind, I would like to argue in support of Levinson’s second definition of narrative using an understanding of Herman’s event-sequencing that also incorporates events inferred by an interpreter in response to a text.22

As Reyland identifies, the interpretative process in music is highly dependent on the listener and his/her cultural positioning:

musical narratives have an intriguing quality: any particular telling or representation (discours) of a plot of musical events (i.e. any live, recorded or imagined performance) can tell a markedly different story (histoir) to different interpreters; collections of those different stories, though, tend to form productively fuzzy networks of interpretative coherence (2014: 204-205, italics in original).

22 In Chapters 2 and 3, we will explore worldmaking and how the construction of a narrative world provides an interpretative structure in which interpreters may make such inferences.
Reyland suggests here that narrative interpretations of music are influenced by factors beyond emplotment, particularly by the listener. Unlike literature, then, in which plot is often taken as encoded in the text, Reyland suggests that conceptions of musical plot arise from the more subjective interaction between the text and reader. More explicitly, Donald McWhinnie writes of radio drama that the listener must “translate the sound-pattern he [or she] hears into his [or her] own mental language; he [or she] must apply his [or her] imagination to it and transform it” (1959: 25). As with Reyland, McWhinnie highlights the role of subjectivity and emphasises a cognitive understanding of narrativity.

With Reyland and McWhinnie’s thoughts in mind, it is helpful to concretise the two understandings of narrative to which Levinson alludes. In cognitive narratology, Ryan reports “[o]n one hand, narrative is a textual act of representation – a text that encodes a particular type of meaning” (2004: 9). Here, Ryan summarises a conservative understanding similar to Bal and Genette. She continues, though, with an alternative understanding that is more consistent with Levinson’s proposition of narrativity as effected rather than encoded: narrative as “a mental image – a cognitive construct – built by the interpreter as a response to the text” (ibid.).

Ryan’s distinction between textual and cognitive understandings of narrative is helpful for, like Reyland, emphasising a subjective dimension of narrative, though she goes on to make a further important differentiation between ‘being a narrative,’ and ‘having narrativity’. “The property of ‘being’ a narrative,” she writes, “can be predicated on any semiotic object produced with the intent of evoking a narrative script in the mind of the audience. ‘Having narrativity,’ on the other hand, means being able to evoke such a script” (ibid.). Hence, although music may not necessarily be produced with narrative intentions, this does not disqualify it from the listener interpreting it in terms of a narrative structure, or ‘having narrativity’ as Ryan puts it.
This cognitive understanding is also supported by Monika Fludernik’s (2010) earlier writing (originally published in 1996) and her concept of *narrativisation*. Fludernik’s work problematises the conservative understanding (such as narrative in Ryan’s first sense), and observes that experiences of narratives are mediated by subjective consciousnesses. She uses the term *narrativisation* to name the process in which a reader interprets a text in terms of a narrative structure.23

By arguing for a model in which narrative represents our experience of a text rather than textual representation of events, Fludernik’s thesis leads to her assertion — preempting Ryan’s two understandings — that a text cannot itself be a narrative. Rather, aspects of a text must be *narrativised* by its reader, that is, organised by the reader in terms of a narrative structure. As she writes, “readers narrativise texts as they read [...] narrativity is not something that is simply present in or absent from texts but rather something that is recognized by readers or sometimes projected onto the text by them” (2009: 109). As Keith Oatley similarly argues, “the reader becomes the writer of his or her own version of the story” (2002: 43).

In parallel with Reyland’s (2014) argument for the subjectivity of musical narrative interpretation, Fludernik, Ryan, and Oatley each emphasise an understanding in which narrative interpretation occurs on the level of the subjective interpreter. This is not to reject some texts’ predisposition towards narrative readings, as the understanding of narrative as a cognitive construct remains contingent upon the negotiation of a text. A similar argument is also made by Rick Busselle and Helena Bilandzic (2008), who propose a model in which narrative is regarded as a subjective

23 Fludernik’s *narrativisation* is an extension of Culler’s (2002 [first published in 1975]) earlier concept of *naturalisation* which, in brief, describes a process in which readers interpret texts in a way which resolves or accommodates inconsistencies in textual information to provide a coherent reading.
mental model that is constructed by the reader from information in the text and supplemented by schemata learnt from embodied experience. Given the same text and broadly similar external experience, therefore, we might expect some resemblance between interpreters’ narrativisations, explaining Reyland’s “fuzzy networks of coherence”.

Fludernik’s concept of narrativisation offers a middle-ground between relativism and prescription, for both the text and issues of biography will inform one’s interpretation. This makes it particularly well suited to discussions of musical narrativity, in which a plot is not readily observable or, as some have argued, even present. In music, the concept allows us to model narrativity as the listener’s structuring of musical discourse in terms of a narrative structure. This returns us to Jerold Levinson’s argument, which we could reformulate thus: narrativity is instantiated by the listener’s narrativisation of musical drama.

1.2 Recorded Popular Song and Narrativity

The critiques of musical narrativity which I have highlighted so far in the context of untexted (that is, instrumental) Western art music suggest that music cannot, strictly speaking, be a narrative. Whilst I hold that a listener may narrativise a piece and so project a narrative world upon his/her interpretation and develop a sense of what it’s like, one could rightly argue that the deictic nature of music (after Cone) compromises worldmaking and what it’s like because of the limited specificity offered when compared to language. This therefore challenges a narrative understanding of music under Herman’s basic elements. To my mind, this should not preclude the possibility of narrativising untexted music, but does invite consideration of narrativity in music with lyrics or a written programme, which works cited above do not, for the most part, engage with.
Moving our attention now to recorded popular song, the presence of performed lyrics would appear to make this a more straightforward candidate for having narrativity, as it incorporates the greater semantic capacity of language. David Nicholls proposes a model in which he approaches records as what we might call ‘plurimedial’, in that they concern multiple connected media. Nicholls explicitly refers to lyrics, music, packaging, and accompanying prose and notes that “the customary presence in a vocal work of one or more texts immediately provides an entry point for narrative analysis” (2007: 299). However, he prefaces this with a note of caution:

narrativity and popular music are not the most obvious of bedfellows: narrativity is theoretically a feature common to all activities involving the representation of events in time, but it is almost invariably encountered in the context of storytelling. Popular music, on the other hand, tends to manifest itself in three- to four-minute songs, often intended as dance accompaniments, which describe essentially static—rather than kinetic—cameos, vignettes, or states of mind (ibid.: 297).

With this, Nicholls casts an uncertain view of narrativity within popular song due to an apparent general lack of event sequencing. He later proposes “music can become part of a narrative discourse, either in those instances where it is ascribed extra-musical meaning through association with an object or a concept [...] or where it interacts with one or more other media” (ibid.: 300-301, italics in original). This argument indicates some applicability of narrativisation to records because he

\[24\] See also Cook’s (1998b) earlier writing, which considers the importance of cover art in the reception of Western art music. In the interest of delineation and due to the popularity of intangible forms of music reception, I limit my own investigations to what can be perceived aurally.
identifies that listeners may interpret pop songs in a narrative manner by ascribing extra-musical meaning.

As Nicholls continues, he appears to develop an uneasy hierarchy, as he contends that musical narrativity may be instantiated by textual information but not music. This point becomes clearer as he sets out what he calls five ‘levels’ of narrativity. These begin with a ‘control level’ of which Nicholls writes that there is no story in the lyrics and hence no narrativity within the musical discourse. Further levels introduce additional media or more sophisticated relationships between these constitutive elements. At the most complex level, Nicholls argues “[a] complex narrative discourse is rendered through multiple media, including lyrics, music, prose, and art work” (ibid.: 301).

Nicholls’ emphasis on lyrics as the basis of narrativity marks a departure from his previous comments which support the possibility that listeners ascribe narrative meaning to music. Furthermore, his relationship between music and lyrics is one-directional as lyrics, he argues, might lend narrativity to music but not the reverse. This therefore denies the possibilities of narrativising music, or using music to narrativise text. Such approaches are also problematised by Byron Almén (2008), who argues that they reflect a conservative position that conflates the text with narrative. Instead, Almén argues, musical narrativity should be treated as a “sibling,” rather than “descendent,” of textual narrativity (lyrics in this case).

Keith Negus (2012b) offers a further critique of Nicholls’ model. Specifically, Negus challenges Nicholls’ location of meaning within a song and instead claims that songs participate in broader social narratives and so their narrativity should not be wholly located in the lyrics, nor the music itself. Instead, he contends, narrative interpretations of songs generally begin as an interaction between listeners, listening
contexts, and aspects of the musical artefact itself. Negus writes of Nicholls’ ‘control level’:

[w]hat I find contentious is the dismissal of three minutes of pop romance [...] on the assumption that meaning resides in and arises out of a composite of tangible, easily identifiable, and directly related “texts”—song lyrics, musical structures, liner notes, art work, and the self-consciously conceptual artistic statements made by musicians. My challenge [...] is that there are other equally complex narrative meanings that are emergent in and articulated to many single pop songs, due to their embedding in a broader social and cultural context (2012b: 370).

In support of this claim, Negus notes that Nicholls’ fleeting use of The Beatles’ “I Want to Hold your Hand” to exemplify his control level does not address the social role of the song, nor the band, in the context of girl group music. Negus also sees narrative aspects emerge from what he characterises as a dramatic musical environment due to aspects such as: the melodic leaps; harmonic patterning complementing the lyrical material of the verses; the participatory connotations of hand claps; and, the rising dynamic and melodic contour of “I can’t hide”. One could

25 The intercontextuality of musical interpretation which Negus identifies is also supported by John Covach’s (1994) earlier writing on what he refers to as musical worlding. Following Heidegger’s ideas on the notion of ‘being’ in a world and the dualism between subject and object, which he finds unsatisfactory and instead proposes the melding of subject and object as a single ‘world’, Covach argues that music analysis tends towards the subject-object split which Heidegger targets. Yet, he argues, this cannot account for the phenomenology of listening and particularly the relationships between musical works. Accordingly, he understands the network of musical works in which a given work is situated as a ‘musical world’ and hence this prompts a process of worlding, or interpreting a work in relation to others.
further point towards the resemblance between the repeated downward melodic contours in the refrain and a peal of church bells, perhaps foreshadowing a marriage between the protagonist and addressee.26

The narrative capacity which Negus finds in “I Want to Hold your Hand” through the interaction between music and sung lyrics is valuable for demonstrating that songs may be regarded as having narrativity for listeners, whereas Nicholls appears to predicate narrativity upon intentional representation. Indeed, this contrasts Nicholls’ description of the most complex level in his typology as cases, whereby “the authors have intentionally linked the various tracks as a narrative cycle, and where the album packaging has been designed to emphasize or complement the narrative theme or themes of its constituent songs” (ibid.: 308, my emphasis). Furthermore, Negus also notes that interpretations may change over time (see also von Appen 2017).

In addition to the elevation of the status of lyrics, which I find to be problematic, Nicholls does not address how lyrics are articulated by the singer. This is an important point, as the singer enacts a character and communicates more than semantic meaning alone (see Section 2.1). This point is also argued by Simon Frith, who, like Nicholls, explores songs as plurimedial narratives (here from a more sociological perspective), but expands on the mediation of the singer:

[t]oday’s commercial pop musics are [...] song forms, constructing vocal personalities, using voices to speak directly to us. From this perspective it becomes possible to look at pop songs as narratives [...] In discussing the narrative devices of contemporary pop in particular, we are not just talking about music but also the whole process of packaging.

The image of pop performers is constructed by press and television [...] by the routines

26 My thanks to Nicholas Reyland for sharing this possible hearing of the refrain.
of photo-calls and journalists’ interviews, and through gesture and performance [...] pop singers are rarely heard ‘plain’ (without mediation). Their vocals already contain physical connotations, associated images, echoes of other sounds (2007: 270).

Frith makes clear the importance of ‘vocal personalities’ within narrative structures of popular song. Frith’s argument also complements Nicholls’ by considering ‘packaging,’ which appears to mean the singer’s social identity portrayed through media publicity and so on.

A non-hierarchical model of popular song, of the sort advocated by Almén, is offered by Allan Moore’s (2005; 2012b) persona-environment paradigm. Like Frith’s ‘vocal personalities’, Moore’s concept of the persona refers to the mimetic projection of the singer in recorded song. The persona, he explains, is constructed from the lyrics, aspects of articulation (pace, timbre, and so on), and melodic contour. He also introduces the ‘personic environment,’ a term which he uses to refer to the instrumental music of a track to avoid the hierarchical implications of the term ‘accompaniment’.

Moore outlines several possible interactions between the persona and environment. Although the model is not explicitly intended for narrative analysis, these interactions consider both increasing participation of the persona/environment in meaning-making and the degree to which they support one another, whereas Nicholls’ levels principally consider discursive sophistication. Moore outlines five possible relationships, beginning with an ‘inert’ level, in which the environment contributes no apparent dramatic support. ‘Quiescent’ and ‘active’ environments enrich the persona’s expression through suggesting certain attitudes or moods, and actively supporting the lyrics, respectively. Moore’s ‘interventionist’ and ‘oppositional’ levels are more participatory still by adding or amplifying details of the
lyric, or contradicting them, respectively. Throughout a track, the relationship between persona and environment may also change, an implicit temporal development which may support a reading of the track as a narrative.

1.2.1 Recorded Music and Narrativity

A further crucial consideration concerns the medium used as the object of analysis. Typically, popular music is disseminated in a phonographic format, which introduces additional questions of how record production informs narrativity. However, neither the persona-environment paradigm, nor Nicholls’ levels, account for the phonographic nature of most popular music. Meanwhile, the scholars cited above in musicology of the Western art music tradition each address the musical score but do not investigate the ways in which it is transformed in the performances put to audiences. In this section, I would like to consider in greater depth the recorded format of recorded popular song and its potential contributions to narrativity.

To describe the virtual acoustical space constructed in recorded popular music, Moore (2012b) presents a heuristic model which he calls the soundbox: a four-dimensional representation of lateral, vertical, and proximate positioning of phonographic elements, and their development over time. An alternative approach to the soundbox, which, like Moore’s persona-environment paradigm and Nicholls’

---

Moore also points to Cook’s (1998a) study of musical multimedia, which develops a model of intermedia relationships, principally in audio-visual formats. Cook’s results are determined through a test of similarity, followed by a test of difference if the former results false (ibid.: 98-129). This provides three possible outcomes: conformance, in which the multimedia pass the similarity test and are determined to be consistent with one another; complementation, where the media are coherent and contrary, failing the similarity test and passing the difference test; and, contest, whereby the media are coherent but contradictory.
multiple media, considers the interactive relationships between aspects of a recording, is proposed by Lelio Camilleri (2010). Camilleri conceives of recordings in terms of three ‘space frames’: localised space, the arrangement of sounds in horizontal and front-back space; spectral space, the distribution of energy throughout the frequency spectrum; and, morphological space, concerning temporal structures, from single gestures to larger metric or hypermetric units.

Several of the parameters highlighted by Moore and Camilleri are explored by Samantha Bennett (2015) in the context of what she calls a tech-processual analysis of The Sex Pistols’ Never Mind The Bollocks, Here’s the Sex Pistols. Although, again, narrativity is not her prime concern, she offers several pertinent insights, in particular by addressing timbre, texture, and gesture in “Sub Mission”. Speaking of the production generally, she identifies the particular importance of texture in evoking fear and anger through the aggressive and claustrophobic connotations of the multiple layered distorted guitars. When moving to the specific details of “Sub Mission”, though, she explores the use of gesture in order to ‘submerge’ the voice beneath the rest of the phonographic environment, realising a narrative detail through the phonographic representation of musical material.

The concerns of Bennett, Camilleri, and Moore outlined above are placed within a narratological context by Aaron Liu-Rosenbaum’s (2012) discussion of what he calls sonic narrative. Liu-Rosenbaum argues for a consideration of how production and phonographic elements may evoke narrativity, with a supporting reading of Led Zeppelin’s “When the Levee Breaks”. His reading begins with a sketch of a lyrical plot: continuous rainfall leads to a flood, prompting the protagonist to attempt evacuation. However, he suggests that sonic transformations of the voice provide narrative

---

28 Bennett describes gesture here as being audible modulations of spatial, volume, and spectral parameters evocative of somatic movement.
information not supplied in the lyrics: he interprets the spatial movement and use of a phaser effect, for instance, as the protagonist struggling to remain afloat amid the flooding.

Liu-Rosenbaum takes a mimetic approach in which he identifies groups of instruments that he associates with the protagonist and the antagonistic force of the river. “The protagonist tracks”, he writes, “are those whose collective role is to serve as storytellers, whether via words or melodic gestures”, in contrast to the antagonists: “those tracks that may be heard as background tracks representing the river” (2012: ‘The Sonic Narrative’ par. 1). Although he does not set out his criteria, they appear to include spatial, timbral, and musical characteristics of the instrumental agents he identifies.29

Two publications by Serge Lacasse develop an alternative approach to phonographic narrativity. In one, he briefly explores the effect of staging in conjunction with the musical setting and performed lyrics in two songs by Peter Gabriel (Lacasse 2005). Unlike Liu-Rosenbaum, he does not explore phonography as a form of discrete dramaturgy, rather he adopts a position in which phonography is positioned as an enhancement of lyrical narratives. In a conference paper from the following year, Lacasse (2006) discusses recorded popular music in relation to what he calls récit phonographique (“phonographic narrative”). Lacasse particularly examines the ways in which phonography contributes to narrativity by indicating temporal change, establishing setting or situational context, and incorporating sounds which appear to emanate from the fictional world inhabited by the

29 It should be noted that although the protagonist-antagonist duality between foreground and background appears reasonable in the case of the this track’s narrative, it is not necessarily a general configuration.
protagonist. Lacasse makes various comparisons between the phonographic environment and a form of theatrical stage but his reading of “Stan” does not involve mimetic representations of characters by instruments in the way that Liu-Rosenbaum identifies. In this sense, Lacasse’s findings of the phonographic environment establish a contextual structure to a narrative interpretation, whereas Liu-Rosenbaum provides some insight into reading of characters and actions.

Lacasse and Liu-Rosenbaum share a position that phonographic aspects of tracks can effect narrativity, though both propose a degree of segregation in the sense that they classify these as “sonic” or “phonographic” narratives. Yet, their readings are couched within lyrical analyses. This is perhaps unsurprising, for there is significant difficulty in conveying events phonographically within the context of popular music styles and, as Nattiez has criticised of music, this is largely deictic. It is premature, therefore, to describe the recorded music of “When the Levee Breaks” or “Stan” as narrative texts in their own right. Rather, our discussions should acknowledge how the constituent aspects of recorded popular song collectively inform narrativisation.31

30 See also Mads Walther-Hansen’s (2015) discussion of sonic diegeses, which I discuss in Chapter 4 (see Section 4.3), regarding the ontology of sound-events in relation to the base level of a story.

31 Burns, Woods, and Lafrance (2016) also explore what they call ‘sonic narrative’ in relation to three tracks by Kanye West, in which their understanding of narrative appears to relate to the communication of social meaning or social commentary rather than the sonification of events as with Lacasse or Liu-Rosenbaum. Nevertheless, they highlight the role of the musical environment and the presentation of sound-sources, which they read in support of lyrical themes, in addition to the intertextual affordances of sampling.
David Nicholls (2004) also offers some important discussion of how phonographic elements contribute to narrativity in what he refers to as ‘virtual opera’ in progressive rock. Nicholls provides readings of four albums spanning 1969-1979: The Who’s *Tommy* and *Quadrophenia*, Genesis’ *The Lamb Lies Down on Broadway*, and Frank Zappa’s *Joe’s Garage*. For Nicholls, the extended form of albums opens additional avenues of interrogation and introduces several further issues beyond the work cited above, such as the telling of subsidiary tales and the role of quotation/motifs. There remain several noteworthy comparisons to be drawn with Liu-Rosenbaum’s analysis, such as his identification of narrativity in lyrics and music (in addition to packaging), and the use of particular musical elements in *Quadrophenia* to depict different agents.\(^{32}\) Unlike Liu-Rosenbaum, however, he elects to use the term *virtual opera* to emphasise the dramatic nature of a plot across several movements performed by a singer and instrumentalists.

### 1.2.2 Plurimediality

Nicholls acknowledges the inseparability of the constituent aspects of the recordings: “for me,” he writes, “the key to understanding the genre [of rock opera] lies in its identification as a hybrid, in which the most typical characteristics of its progenitors come together to create a distinctive artistic form whose whole is greater than the sum of its parts” (2004: 104). Although he is speaking specifically about progressive rock, I would like to argue that we should extend similar thinking to recorded popular song.

Herman (2009a) stresses that the form of a narrative discourse is particularly influential in informing how narrative information is rendered, and how interpreters negotiate a text. Using the example of graphic novels, for instance, he emphasises the importance of the different expressive media that contribute to the narrative

---

\(^{32}\) I return to *Quadrophenia* and provide my own analysis of the album in Chapter 6.
discourse. Here, he asserts that narrativity is informed not only by linguistic syntax, but also parameters such as typography, page formatting, and the visual style of speech bubbles. Regarding popular music, Tim Warner makes comparable remarks by arguing that “for the past 70 years or so the sonic and musical characteristics of much popular music have been informed by the processes and technologies of audio recording” (2009: 131). Based on the approaches to recorded popular songs explored so far, we could describe tracks as plurimedial combinations of elements which together prompt and inform the listener’s narrativisation of a track. These media might include: the persona, environment, packaging, production, and accompanying programmatic prose.

Although discussions of sonic narrativity attend to the phonographic nature of recorded popular song, this approach risks neglecting the music itself and — as we have seen — is influenced significantly by the persona. Whilst Nicholls’ model helpfully addresses the plurimediality of popular song, the role of production is not represented and the lyrics are given precedence over each other component. Therefore, I shall employ Moore’s persona-environment paradigm in my discussion of narrativity, as it emphasises the interconnectedness between music and the persona without presuming a hierarchy between the two. Furthermore, it focuses on the level of the song without presuming the presence of packaging or other media.

33 See also Moore (2010a), which similarly stresses in the case of the recorded track that the phonographic medium renders the combination of performed lyrics and music distinctive from its constitutive elements in isolation, and Lacasse (2006).

34 Although my use of the term ‘media’ to describe sonically identifiable aspects of production may seem unusual, Ryan (2004) emphasises that the term itself is uncertain; my use of it is consistent with Webster’s Dictionary’s definition of medium as “[m]aterial or technical means of artistic expression” (qtd. in Ryan 2004: 16).
which the listener may or may not encounter. Throughout this project, I also refer to the *phonographic environment* in order to account for the important aspect of production and the issues raised by Camilleri, Lacasse, and Liu-Rosenbaum. I will introduce this concept more fully in the following chapter, but, in short, I use the term to refer to the phonographic representation of the persona and personic environment, and to also encompass sound-sources which are not compatible with Moore’s conception of the personic environment in terms of tuned or percussive sound-sources.

### 1.3 Case Study: Ke$ha - “Stephen”

In our discussion so far, we have explored the theoretical possibility of narrativising a track alongside several analytical approaches to musical narrativity. To apply the ideas raised above to a worked example, I would like to discuss Ke$ha’s “Stephen” and consider how this particular track can evoke narrativity.

Throughout the track, the singer takes on the role of a protagonist as she fawns over a romantically unavailable addressee, Stephen. The track opens with a solo voice as she calls to him with some discernible conviction, staged without noticeable reverb close to the listener in the soundbox. The lead voice is quickly joined in the second phrase by multiple recordings of the same singer that are spread throughout the stereo field and sung in harmony as she asks why her addressee will not call her. The introduction ends as the phrase is repeated and the harmony vocal parts are vocoded, which leaves the backing vocals less timbrally expressive whilst the lead vocal takes on an exaggerated, servile character. A sweeping filter gesture on the vocoded timbre moving from left to right in the soundbox transitions into the first verse.
The first verse is marked by several textural changes and the first depiction of a situation between the protagonist and antagonist. The protagonist continues to directly address Stephen and recounts drunkenly flirting with him at a social occasion whilst his girlfriend sneered at her from the other side of the room. The voice here is presented with a subtle reverb, joined by electronic percussion and two synthesiser lines. The first synthesiser is a stylophone-like timbre that enters from the start of the verse, panned to the left of the soundbox, which approximates the vocals’ melodic contour. The second synthesiser line enters in the last two bars, when the protagonist acknowledges Stephen’s “ugly girlfriend”. The timbre of this second synthesiser is somewhat more pleasant than the first. It is panned towards the right of the soundbox, symmetrically opposite the stylophone, and similarly impersonates the melodic contour of the voice.

The verse is followed by a 3-bar prechorus, in which the texture changes abruptly as the protagonist declares her fondness for Stephen. Here, an accordion enters, the two synthesisers are replaced with a soft electric piano line which arpeggiates around scale degrees outlined by the bass movement of a synth pad. The somewhat brash electronic percussion is also removed, with the exception of narrow-bandwidth electronic claps. To denote the closure of the section, the electronic piano plays broken chord flourishes on either side of the stereo field, leading to a chorus that contains the material used in the introduction. This is perhaps the most texturally-rich passage of the track: the electronic percussion reenters in full, joined by the bass pad, electronic piano, voice, and vocoded backing vocals. The vocal material here largely pivots around asking why Stephen will not call the protagonist, who explains how inadequate it makes her feel to wait for him.

From 1′05″, we move to the second verse which, as before, consists of a vocal-led melody imitated by the stylophone, electronic percussion, and the second
melodic synth symmetrically opposite the stylophone (this time from the opening of the section). The protagonist here describes all her male admirers and acknowledges her fickleness with men. Leading into the fifth bar, the lead vocal explains how the protagonist need only bat her eyes to break her admirers’ hearts. As she does so, a bass synth is introduced with brighter and harsher timbre than before, whilst the stylophone is taken out. When the protagonist explains how she would flirt with men, a bright synthesiser with a bell-like envelope performs a playful undulating countermelody. Again, the verse is followed by a prechorus, in which the protagonist expresses her bemusement that she is the one chasing after Stephen despite her apparent charm.

Following the second verse, the chorus is repeated, which leads into a middle-8. In desperation, the protagonist asks whether Stephen refuses to call her because he may think she’s mentally unwell or ugly, and whether he simply does not love her. During this section, a second drum part is added in addition to the electronic percussion heard previously, which plays a pattern reminiscent of bossa preset accompaniment rhythms which were popularised by keyboard manufacturers in the early 1990s. After four bars, an electronic organ timbre enters and references Wagner’s Bridal Chorus. At the end of the middle-8, an extended prechorus leads to a playout in which the chorus is repeated with some small lyrical variation.

1.3.1 Narrativising “Stephen”

On the face of it, “Stephen” could be said to exemplify Nicholls’ control level in which there is no significant narrativity as the song is largely situational. However, with greater attention to the environment, production, and temporal progressions implied by the persona, we may pursue a more rewarding narrative reading of the track. We can begin to assemble a rudimentary narrative from the protagonist’s telling of a
previous drunken encounter with Stephen (verse 1), which precipitated an intense infatuation, despite his unavailability and disinterest (choruses, prechoruses, and middle-8), whilst reportedly having various other men fighting for her attention (verse 2). To show how the musical environment also offers the opportunity for narrativisation, I would like to start by considering the middle-8, before highlighting several other moments in which the persona is supported by the environment.

During the middle-8, the environment offers a narrative cue which is not revealed by the persona. Whilst the persona here interrogates Stephen’s possible reasons for rejecting her advances, the quotation of the Bridal Chorus suggests that she is fantasising about marrying her addressee. Furthermore, the extent of vibrato in conjunction with the timbral character of the electronic organ, and the presence of drums playing a pattern we might find in an ‘off the shelf’ keyboard, might prompt us to narrativise an unsophisticated, chintzy event; a point of irony which supports a reading of this ceremony as imagined by the protagonist. The presentation of the voice also supports this reading, for the lead voice is staged using an effect which simulates the timbral characteristics of a telephone call. This could explain the rhetorical way in which the protagonist questions him, as though leaving a voicemail message, but it certainly suggests some form of spatial dislocation from the rest of the environment. Although the protagonist is clearly infatuated with her addressee, she never verbalises an intention to marry him, suggesting that the environment provides some insight to her private, internal psychology.

The presentation of sound-sources also offers some further details when taken in conjunction with the lyrics. The introduction, for instance, can be taken — as with the wedding scene — as a moment in which the phonographic environment offers some insight into the protagonist’s internal psychology. Unlike the majority of the track, the voice is staged without any reverb, suggesting great intimacy as though the
voice acts as a form of internal monologue. We might also read the multitracking of
the singer as a reflection of the many times in which the protagonist has rehearsed
her approach towards Stephen.35

Whilst the way in which the introduction is presented sonically could suggest
an internal monologue, the subtle reverb added during the first verse situates the
voice in some space. Yet, here, the environment can be taken as a depiction of the
situation she recalls when she drunkenly flirted with Stephen across the room from
his girlfriend. This interpretation relies on narrativising the two synthesisers as
depictions of the protagonist and her antagonist, as supported by the timing and
stereo placement of the two instruments. The stylophone, which I suggest represents
the protagonist, enters at the start of the verse as she begins her autobiographical
recollection and is placed on the left of the stereo field. The entry of the second
synthesiser, meanwhile, coincides with the mention of Stephen’s girlfriend’s
sneering, and is placed on the opposite side of the soundbox, depicting the same
reported separation between her and the protagonist. Taking the two synthesisers
here to represent the two narrative agents allows us to infer characteristics from the
two timbres. As McNamee’s (2009) rather unflattering account of the stylophone’s
sound suggests, its timbre is rather ungainly and so, if we are to narrativise it as
representing the protagonist, we might ascribe similar unkept or disorderly
characteristics to the protagonist also. Likewise, the second synth is chorused and a
great deal more consonant, lending a smoother, polished character to Stephen’s
girlfriend.

35 This multitracking is effectively an analogue of Denis Smalley’s (2007) holistic space, i.e. a
description of a soundscape using the spatial characteristics of sound-sources without their
duration or order of appearance, collapsing time in a similar manner to a long-exposure
photograph.
The textural change between the verse and prechorus should also be noted. The accordion, for one, is an interesting addition, which could perhaps be seen to reference some association with passion and romance. Furthermore, the exclusion of most previous instruments brings the focus to the solo voice in the centre of the stereo field, as if to offer a greater sense of sincerity and directness.

Further lyric painting occurs in the second verse, in which the slight distortion of the synth as the protagonist sings about batting her eyes suggests a degree of deviousness, while its location in the stereo field (borrowed from the stylophone, which also was dropped from the mix when the phrase begins) suggests an association with the protagonist. If we take this as the case, we might understand the different instrumentation which occupies the spatial position previously associated with the protagonist as an illustration of various outward appearances as she pursues different possible suitors.

In these ways, we can see that the persona and environment together offer ways of narrativising the track and interpreting it within a narrative structure in which the protagonist continually chases her addressee whilst reporting events from the past and imagining their possible future together. Notably, we see a key distinction in the production of the vocal track between the verses and the other sections. In the verses, we are presented with what we might call a ‘realist’ voice, whereas other sections make greater use of distortion or vocoder effects. This changing degree of apparent realism supports the suggestion that the wedding scene, for instance, is a fantasy, not to be confused with the reality that her addressee will not call.36 Developing a mental model of this situation and the forces involved in it also makes it possible to infer possible outcomes. In the following chapter, I will go on to outline a theoretical approach to how this mental model or ‘world’ is drawn

36 I expand on the topic of realism or, more specifically, verisimilitude throughout Chapter 3.
from a text, and also expand our discussion of narrativisation to account for the implicit agents involved.

1.4 Conclusion

In this chapter, we have explored several positions on the feasibility of musical narrativity. As we have seen, there is some opposition to the application of narrative theory to instrumental music on the grounds that musical expression cannot represent events with the specificity of the written word. Several scholars also propose that music is mimetic rather than diegetic and therefore argue that music is dramatic in nature and not a form of narrative. However, both these criticisms of musical narrativity employ a conservative understanding in which narrative meaning is wholly encoded within a text.

As Ryan explains, we might think of narrative in simple terms either as a series of encoded events, or as a cognitive phenomenon which is instead *effected* by a text. In conjunction with Fludernik’s theory of *narrativisation*, Ryan’s latter understanding can be applied to recorded popular song in the sense that tracks prompt the listener with means of narrativisation. Through this process, I contend, listeners organise and draw inferences from the information offered by a track.

Although my position departs from the older and better established model of narrative as an encoded sequence of events, I contend that narrativised songs are compatible with the four basic elements of narrative which Herman proposes. Regarding *situatedness*, a communicative context is instantiated most explicitly by the presence of a persona, and also in many cases by the musical environment. As “Stephen” illustrates, both these aspects may support the identification of situations or events which may be structured or supplemented through narrativisation. This
also requires the imagining of a possible world inhabited by Stephen and the protagonist to situate the dance, the context of the protagonist’s advances, and hypothetical wedding. Finally, what Herman refers to as *what it's like* may be afforded by several aspects, including: the affective role of the phonographic voice and environment; the construction of a narrative world sympathetic to the protagonist; or, sequences of events with which the listener empathises. In the following chapter, I will support the compatibility of Herman’s four elements with narrativised songs by discussing several elements particularly relevant to *situatedness* and *worldmaking*.

Of course, the ways in which listeners engage with and narrativise songs is also influenced by the form of the recorded artefact itself. I take the stance after Nicholls that recorded popular song is plurimedial and so narrative meaning should be discussed as a function of the multiple media concerned, and the relationships between them. However, unlike Nicholls, I do not believe that music can only be invited into a narrative discourse specified by the lyrics; rather I hold that the persona and environment each introduce various creative possibilities and the possibility of narrativisation. Furthermore, the persona and environment are presented phonographically, which may also extend interpretations. Such a direction could otherwise also be augmented with greater attention to sociological context, as suggested by Frith (2007) and Negus (2012b), though I remain principally interested in the recorded text as the object of analysis.

Although my intentions here are to establish that recorded popular song is capable of having narrativity, it should be noted that the principals I have discussed here, especially narrativisation, can readily be applied to discussions of art music. Furthermore, as I have highlighted, there is a significant lack of writing on narrative in the context of texted art music, which stands to benefit from a plurimedial approach as I have argued for in recorded popular song.
2. Narrative Worlds

Having discussed how recorded songs may prompt narrative readings, this chapter will develop our definition of narrative as the mental representation of events by considering aspects of Herman’s *situatedness* and *worldmaking* (see Section 1.1). As stated earlier, situatedness refers to the communicative contexts of narratives, and particularly the agents that are involved in the telling of a story. To this end, the following section introduces Seymour Chatman’s narrative communication model, which describes six agents (a combination of actual and implicit beings) through which a story is told. In order to adapt this for the purposes of our discussion, I propose six such agents in the narrativisation of recorded popular song.

The second of Herman’s ‘basic elements’ which we will discuss in this chapter, *worldmaking*, refers to the evocation of what I refer to as a narrative world. This follows the hypothesis stated in the General Introduction that narratives are situated within a possible world imagined by the interpreter. Narrative worlds therefore encompass aspects of place/time, characters, motivations, and other governing structures. As narrative worldmaking — the process of building a narrative world — takes place through the narrativisation of a text, this chapter will systematise the process in terms of three types of world (broadly construed) in Section 2.2, resembling Genette’s distinction between narration/narrative/story. These are the interpreter’s understanding of: the actual world in which he/she is objectively located; the ‘sound world’ of the recording; and the narrative world in which a story unfolds. The identification of these three worlds allows us to situate the six agents developed earlier in response to Chatman’s model.

Having described narrative in terms of three worlds, our focus will turn to the narrative world specifically. I shall also introduce the concept of affordance in Section
2.3.2, derived from ecological psychology, to describe the way in which texts prompt interpreters to imagine narrative worlds. In brief, this theory holds that environments offer possible interactions to organisms and, as a model of musical meaning-making, provides a theoretical approach which acknowledges the role of the text whilst accounting for the subjectivity of interpretations. Finally, we will connect the concept of narrative worlds with the sensation of immersion or ‘accommodation’ in Section 2.5, which I develop further in the following chapter.

2.1 Narrative Agents

On the one hand, Herman’s situatedness involves properties of a text which communicate or evoke certain interpretations. This includes, for instance, written language in novels; spoken word and body language in oral storytelling; and, audiovisual cues in film. On the other hand, Herman emphasises, communicative situations occur within broader socio-cultural contexts that might, for instance, allow one to reconstruct motivations for writing or some critical intentions behind the text. Of these two components, I am concerned principally with textual aspects to support analytically informed discussion of narrativisation and worldmaking.

As I have noted in the previous chapter (see Section 1.1), narratives are generally taken to be the telling of a story, and, whilst I hold that the interpreter has a more significant role than this understanding might suggest, the communicative quality of narratives is especially pertinent to songs. With this in mind, we should recognise the various actual and implicit agents whose roles inform narrativisation. A helpful starting point to such an investigation can be found in Seymour Chatman’s narrative communication model, which identifies six actual and implied agents involved in the process:
Chatman’s model commonsensically begins with the real-life author, which he distinguishes from the image of him/her evoked by the text, the implied author (see Booth 2006). As Rimmon-Kenan puts it, this implied author is “often far superior in intelligence and moral standards” (1983: 87) than the real author. Thereafter, Chatman poses that the implied author communicates meaning via a narrator. In contrast to Genette and Bal’s definitions cited in the previous chapter, Chatman places the narrator in parentheses, as he suggests that narrators are common but not ubiquitous to narrative.37

The remaining agents which Chatman describes concern receiving agents that act more or less symmetrically to the real author, implied author, and narrator. The narratee refers to the agent whom the narrator addresses. In other words, the narratee is a character within the same world occupied by the narrator and is the intended recipient of the tale. Moving on, the implied reader describes an idealised, model reader capable of interpreting textual information as intended by the real author. A text might incorporate, for instance, untrustworthy narration that the implied reader should be sufficiently competent to recognise. Importantly, though, Chatman’s model distinguishes the implied reader from the real reader, i.e. the real-life individual who interprets the text (and who may depart from the idealised, implied reader).

---

37 The narratee is also parenthesised for this reason. In her critique, Rimmon-Kenan (1983: 86-89) argues that the narrator and narratee are more intrinsic than Chatman suggests.
2.1.1 Narrative Agents in Song

Given the linearity and unidirectionality of Chatman’s model which suggests that a ‘correct’ meaning is encoded and contingent on the author’s intentions, I do not wish to endorse it as a model of narrativity. Such a model cannot account for the way in which interpretations may change over time, nor different interpreters’ experiences and forms of engagement with a text. However, Chatman helpfully identifies four agents implied by the text and provides the basis of a reformulation in this subsection to derive a set of related agents in the case of recorded song.

With Chatman’s model in mind, I propose below a typology of agents which participate or are implicated in the narrativisation of recorded popular song, with the support of existing scholarship where applicable. I focus here on the singer, for songs are often regarded as “personally expressive” (Frith 2002: 186). Furthermore, the singer as we hear them in a track often provides the most immediate point of entry into a narrative world because personae generally enact a central character. 38 Accordingly, I begin with the real singer, before we come to the persona which he/she enacts in performance, and the protagonist/narrator of a given track. As with Chatman’s narratee, the protagonist/narrator will generally appear to direct themselves to an intended recipient, the addressee. Thereafter, the choice and configuration of phonographic sound-sources construct a particular perspective and denote what I term an implied listener, as distinct from the real listener.

38 I should stress that due to the general nature of the following typology, we might approach instruments in a similar manner, in terms of: 1) actual performers, 2) their phonographic representation, and 3) a narrativised character. This may be used to provide the basis of a mimetic narrativisation, as we have seen in “Stephen” (see Section 1.3).
Real Singer and Persona

Chatman’s model initially distinguishes between the author as a real-life individual, and his/her textually mediated representation. Within both musicology of popular and of art music, similar distinctions have been made between real-life individuals and the personae which they adopt. Scholars have keenly argued, for instance, that Western art music can project the metaphorical ‘voice’ of the composer.39 This is also supported by Roger Watt and Roisin Ash, who, in reporting on the psychology of musical listening, argue that listeners respond to music and assign it attributes as though it were a person. In their words: “loosely speaking, music creates a virtual person” (1998: 49).

In a similar manner to Chatman’s first three agents, Phillip Auslander’s (2009) reformulation of Simon Frith’s (2002) work on the voice in song describes three modes of the singer: the real-life individual; the performance persona, i.e. a personality adopted by the singer in a social or professional setting; and, the narrative character he/she may take on. To place a greater emphasis on the persona’s narrative involvement, Moore’s (2012b: 180) adaptation of Auslander’s categories distinguishes between the real-life singer; persona, as performed both socially and within song; and, the role of this character in relation to the situation or events described (typically a protagonist). In other words, the voice which we hear in a track, as with the words in a book, cannot be taken as the unmediated expression of their originator. In song, we are not presented with an unmediated real singer, but rather, his/her persona.

39 Cone (1974) offers a prime example of investigating music as a reflection of an authorial persona. Although Jean-Jacques Nattiez assuredly contests the possibility of musical narrativity, he also begins with an observation which he calls “an intuition of common sense: through the work, the composer speaks to us” (1990a: 240).
The fictionality of personae is not necessarily immediately obvious, though we can observe some cases in which this is clearer. For one, artists may ‘reinvent’ their public identity, leading Grossberg (1993) to observe the commoditisation of ‘authentic inauthenticity’. Secondly, popular musicians often record and perform under a stage name. This again testifies to a distinction between the real-life individual and the persona which he/she performs. Madonna provides one such example, as her public-facing image and musical identity have radically shifted and embraced various, often controversial, subject matter. In this sense, we can observe several personae performed in different periods throughout her career. Furthermore, we can distinguish between a private, real-life individual originally born Madonna Louise Ciccone, and the personae she enacts musically, collectively, Madonna the entertainer.

**Protagonist/Narrator**

In recorded popular song, the persona is rendered sonically, though it will generally express from the point of view of a character. In many cases, the persona adopts the first-person perspective of a protagonist, an involved agent who speaks principally about his/her own experiences. As Nicholls (2007) and Moore (2012b: 183) both observe, this position is more common than a third-person narrator, but both generally involve the adoption of a character. Occasionally, we may encounter the persona as a third-person narrator who relays the experiences of others. Billy Joel’s “Piano Man” offers an example of this latter position, in which the persona adopts the
role of a bar’s piano player but spends much of the song reporting on the scene and voicing the other characters within the bar.40

After Herman’s worldmaking element, we should acknowledge that the protagonist/narrator must inhabit a world. In many cases, the relatability of experiences reported through song encourage us to think of this as the same world inhabited by the real singer and real listener. The popular hook in Carly Simon’s “You’re so Vain” in which the persona sings to a presumptuous addressee, for instance, has attracted considerable attention over the years because the song is taken to be autobiographical and that the protagonist and anonymous addressee are actual, real-life individuals. Indeed, the persona acknowledges that she is singing a song with this lyric, which acts as a mimetic strategy that obfuscates the distinction between persona and protagonist to suggest the song does not involve a fictional world, but rather inhabits the social realm of the real listener and real singer. However, this world inhabited by the protagonist/narrator is crucially mediated through songwriting and, in many cases, is fictionalised to some extent.41

However clearly the persona enacts a protagonist or narrator, this distinction should be maintained. In David Bowie’s “Space Oddity”, for instance, the persona voices both a ground control operative and the central character, Major Tom. As Major Tom is voiced in the first-person and sings from the perspective of the central,

40 Some styles tend towards different forms of narrator. Although it is common in contemporary mainstream pop for the singer to enact a first-person protagonist, for instance, third-person narrators are often more common in folk ballads, which tend to adopt more conventional narrative elements.

41 Throughout the following chapter, we will discuss this degree of verisimilitude or departure, in relation to the construction of the persona and phonographic environment, in greater detail.
involved character, his role is best described as the *protagonist*, yet the persona also voices the supporting role of a ground control operative. As with “You're so Vain”, other tracks may also demonstrate the mediation of the protagonist by acknowledging the performed or fictional nature of a recorded song. In P!nk’s “Centerfold”, for instance, we hear her asking to increase the volume of the music, as though communicating with the studio technicians, ‘breaking the fourth wall.’ Clearer still, Professor Green’s “Need you Tonight”, featuring Ed Drewett, begins to fadeout at 3′25” as the persona begins speaking a monologue in which he disputes the content of the song to suggest that, as a real-life individual, he would not encounter difficulty attracting women. In short, whereas the persona is reflected in the sonic discourse of a recorded track, the protagonist/narrator(s) that it voices are ontologically separate. Chatman’s progression between *real author - implied author - narrator* can therefore be put alternatively regarding the singer in popular song as *real singer - persona - protagonist/narrator*.

In some cases, we may also see multiple personae take on the role of the same character during the course of a song. One example of this is Nico and Vinz’s “In Your Arms”, which I discuss in greater detail in Chapter 4 (see Section 4.1.3), in which two personae take on the role of the protagonist at two points in his life. Other examples of this can be found in Owl City’s “Tokyo”, in which the personae of Adam Young of Owl City and Fukase of Sekai No Owari both voice the protagonist as he celebrates his love of Tokyo, and SAARA’s “Ur Cool”, which combines the voices of the lead singer and a younger female voice which suggests a sonification of the protagonist’s ‘inner child.’
Addressee

The presence of a protagonist within a song’s narrative also suggests the presence of an addressee or implicit antagonist, and a common aspect of popular song writing sees the protagonist address another party, hereafter labeled the addressee. The addressee, therefore, is analogous to Chatman’s narratee in that it is the intended recipient of the narrator/protagonist and occupies the same narrative world. In solo songs, the addressee is often unqualified and addressed through the word “you.” This invites the listener to either narrativise an addressee and infer its relationship to the protagonist or perhaps even imagine the listener him/herself as the addressee.

Implied Listener

As stated, the protagonist/narrator and addressee inhabit the same world, i.e. the world in which a given narrative unfolds. This is consistent with Chatman’s model, though as we move from narratee to implied reader, we shift from an agent within the narrative world to an implied agent which results from the form of the given discourse. That is to say, the narratee is a character whereas the implied reader is an implicit agent who does not occupy the narrative world. As we move from addressee to what I term the implied listener, we see this same shift from narrative to phonographic worlds.

Although Watt and Ash’s idea of a ‘virtual person’ imagined through listening is applied principally to the projection of a composer, we could also think of music in terms of some form of experiencing consciousness. This is especially applicable to recorded music, in which the phonographic environment specifies a virtual acoustical space in which materials are organised. An analogue to Chatman’s implied reader — the idealised or intended form of reader deducible from the text — is suggested by Stan Link (2001) through his discussion of noise in recorded music. The
presence of noise, he argues, testifies to the existence of an intermediate agent, as though we encounter musical discourse through the perceptions of another. It is this sort of intermediary which I refer to as the implied listener, that is, a perceiving agent implied through the selection and arrangement of sound-sources in a recording.

As a particularly overt example of the implied listener in recorded music, consider Sigur Ros’ () album: the first track begins with a needle dropping upon a phonographic record and after track four there is a lengthy pause during which the listener presumably changes sides. Some vinyl noise and allusions to physical recording media are also present throughout the tracks as an implied listener acts as intermediary and we hear him/her ‘listen’ to the record. The first verse of our previous example of Ke$h'a’s “Stephen” provides a subtler case due to the anthropomorphism of instruments but nevertheless, we can describe an implied listener here as a third-party watching the protagonist and Stephen’s girlfriend from across the room. The implied listener, though, is not necessarily restricted to possible human agents and we will see several further examples when I extend this discussion later in Chapter 5 (see Sections 5.3-5.4).

2.2 A World Model

In my derivation of agents from Chatman’s model, I noted an ontological separation in that whilst the real singer and real listener are actual, real-life individuals, the persona and implied listener are manifested phonographically. The protagonist/narrator and addressee, meanwhile, both inhabit a possible, narrative world. In this section, therefore, I would like to concretise this distinction by proposing a series of three worlds that situates the six agents above and underpins my approach in subsequent chapters. Of particular interest is the narrative world
because, as Herman’s inclusion of worldmaking amongst his basic elements emphasises, a key property of narrative is that it prompts the interpreter to imagine a narrative world inhabited by the characters of a song.

For Marie-Laure Ryan, the “ability to create a world, or more precisely, to inspire the mental representation of a world, is the primary condition for a text to be considered a narrative” (2013: 363-364). Indeed, the narrative world which Ryan is referring to here establishes the context in which we narrativise event-sequences and perceive Herman’s what it’s like. Furthermore, the imagination of a world affords the immersive quality of narratives through the construction and negotiation of this mental representation. Samuel Taylor Coleridge’s (1817: 2) fashionable phrase to describe narrative immersion was “suspension of disbelief”, which charged the interpreter as responsible for accepting a fictional narrative. Contrastingly, however, the novelist Michael Frayn writes: “[w]hat I accept, provisionally, is an alternative world […] The storyteller, in other words, creates by fiat, or web of fiats, an alternative world that I agree to inhabit for a while” (2006: 245). Frayn, therefore, regards the interpreter to be responsible for accepting a narrative but captures the critical importance of engaging with, or even “inhabiting”, the world of the narrative.

Following Genette’s and Bal’s tripartite description of narratives which differentiate between understandings of narrative as telling, text, and event-sequencing, we can position the six agents in terms of three equivalent worlds. The most straightforward of these worlds is the actual world inhabited by the real singer and real listener. These are distinguished from their implied, phonographic representations, i.e. the persona and implied listener. At a further level, the protagonist/narrator and addressee inhabit the fictional world of the narrative itself. These three frames of reference allow us to distinguish between an actual world, a
textual domain, and the world of the narrative itself, as represented by my revision to Chatman’s diagram:\(^{42}\)

![Diagram](image)

Fig. 2.2: six narrative agents in recorded popular song represented in different levels or ‘worlds’.

A dominant aesthetic position of popular song is to write songs that resemble the actual world — that is, the world in which the listener is physically located — though I must place an emphasis here on the distinction between this and the imagined narrative world. As described in the General Introduction, I take narrative worlds to be mental models which readers construct from available textual information, supplemented by their exterior knowledge.\(^{43}\) As Herman defines them:

\[
\text{[s]toryworlds are global mental representations enabling interpreters to frame inferences about the situations, characters, and occurrences either explicitly mentioned}
\]

\(^{42}\) In this representation, the arrows which described a linear communicative trajectory in Chatman’s model have been removed, as I wish to focus on the agents involved in narrativisation and a linear communication model places a problematic emphasis on intended authorial meaning and representation.

\(^{43}\) The concept of narrative worlds is sometimes equivalently labeled *storyworld* or *narrative universe*. Marie-Laure Ryan (1991, 2001) instead uses several categories of possible worlds which are also imagined by the reader using textual information and supplemented by external schemata.
in or implied by a narrative text or discourse [...] mental models of the situations and events being recounted — of who did what to and with whom, when, where, why, and in what manner (2009a: 106–107).44

In the case of music, distinguishing between an actual world/discourse/imagined world leaves us with three discrete worlds, which I shall discuss in turn below:

A. a reference worldview, in relation to which the narrative world is understood;

B. the phonographic environment of a recording, which affords to the listener ways of narrative worldmaking;

C. and, the narrative world which the listener imagines in response to the recording.

2.2.1 Reference Worldview

Whether humans interact directly with a single coherent world or instead navigate subjective mental models of what we take to be an actual world, we should remain mindful that our knowledge is learned from our own experiences external to song.45

44 Although Herman suggests a particular form of character-bound narrator using the term “recounted,” I should stress the relevance of his argument to all forms of narrator.

45 For the sake of simplicity here, I follow the hypothesis of actualism that there is a real, actual world which encompasses everything and everyone that exists. This provides a straightforward means of describing the relationship between fictional or otherwise non-actual worlds and what we might call ‘the real world’. Whilst actualism offers a simpler basis to this project, it is a topic of considerable discussion by philosophers. Putnam (1981), for
Therefore, I take the position that narrative worldmaking is contextualised by a *reference worldview* — that is, one’s subjective view constructed from embodied experience and knowledge of what we might call ‘the real world’. For Ryan, this ‘actual world’ as she calls it is central to her discussion of possible worlds. She defines this actual world as the “center of our system of reality. AW [the actual world] is the world where I am located. Absolutely speaking, there is only one AW” (1991: vii). Due to the importance of subjectivity in the interpretation of recorded popular song, I should emphasise that whilst I support Ryan’s use of an actual world as the centre of her system, each reader will bring to a song different external experiences and knowledge which will create some variety in their narrative worldmaking. The importance of the reference worldview is also emphasised by Rick Busselle and Helena Bilanzic’s (2008) model of narrative comprehension and engagement, in which the pair argue that narrative worldmaking is predicated on the interpreter’s activation of external schemata.

Throughout the following chapter, we will see how the relationship between the reference worldview, phonographic environment, and narrative world may inform narrativisation. As mentioned previously, a dominant aesthetic trend in popular music is to align a narrative world with the reference worldview such that the events or situation expressed through song could (or even have) actually taken place. This overlap between a narrative world and the reference worldview is important in two ways. Firstly, it ensures that narrative events can be understood by a listener within a known frame of reference, encouraging empathy with the reported example, provides a popular example of a disembodied brain which is sustained artificially and provided all necessary stimuli to simulate its embodied existence in another world, emphasising the limits to knowledge and suggesting that individuals interact with a mental model of a world rather than a world directly (see also BonJour 2013).
events. Secondly, by disguising the constructed nature of the narrative world, writers may make it appear to be in closer correspondence with actual or possible events, as though the listener him/herself could inhabit the same world or even experience the same situation/events. Therefore, this provides one means of a narrative situation or series of events appearing more ‘authentic’ or ‘true to life’.

### 2.2.2 Phonographic Environment

Writing on the effects of recording on musical listening, Patmore and Clarke argue that recordings act as “‘gateways’ into a virtual world of instruments and voices [...] what we then imagine that we hear are the sounds of those instruments and voices, rather than the sounds of the record, tape or disc” (2007: 270). This virtual world which Patmore and Clarke describe is also frequently discussed in terms of a ‘soundworld’, although this term is open to somewhat broad usage. For consistency, I shall use the alternative term *phonographic environment* to refer to the phonographic representations of the persona and other audible materials. It is this phonographic environment which affords to the listener ways of imagining a narrative world.\(^\text{46}\)

### 2.2.3 Narrative World

Narrative worlds refer to the mental models formed by an interpreter of a world mediated by a text and make it possible to infer narrative details which may not be

\(^{46}\) I shall use the term *phonographic environment* in relation to the phonographic representation of all sound-sources, including the persona, in terms of aspects such as imagined source, timbre, amplitude, sound processing, and the encapsulating acoustical space, but not semantic or performative aspects which are discussed in relation to the persona and personic environment, respectively.
explicitly stated. We have already seen an example of inference relating to narrative worlds in Kennett’s wine shop thought experiment in the General Introduction (see Section 0.3.1). His is a narrative world employed to infer possible events in a reference world.\textsuperscript{47} Importantly, whilst it is possible for the phonographic environment to be reproduced identically upon each repetition (subject to playback media and technological considerations), interpretations of tracks (and, accordingly, the narrative worlds which we construct from them) may change over time, as Negus (2012b) has argued.

### 2.3 Narrative Worldmaking and Affordance

Having described three worlds in recorded popular song, I would to like to reflect on the process of narrative worldmaking, i.e. the interpreter’s imagining of the narrative world. Goodman’s (1978) \textit{Ways of Worldmaking} provides an authoritative general introduction to the ways in which worlds are made, and is also the basis of Herman’s (2009a) own discussion. For Goodman, worlds are generally understood as symbolic systems such as caricatures, musical scales, and mathematical diagrams, though can also be applied to \textit{versions} of what we might call the ‘real world’.\textsuperscript{48} As he does not focus on specific functions or attributes, Goodman’s theories might be applied readily to

---

\textsuperscript{47} Ralf von Appen (2015) gives a further example of a narrative world as a hypothetical model to extrapolate possible actual events in his fictionalised account of the recording process of Ke$ha’s ‘Tik Tok’, which he constructs by extrapolating from various accounts of the track’s production.

\textsuperscript{48} This recourse to symbolic representation recalls the tradition of semiotics. However, as he explains in his earlier \textit{Languages of Art}, his use of the term is less specific than the more prominent semiotic systems (for instance, of Pierce and Saussure) which have been propagated in musicology: “‘[s]ymbol’ is used here as a very general and colorless term. It covers letters, words, texts, pictures, diagrams, maps, models and more” (1976: xi).
visual or aural arts as well as physical worlds, though his support for the multiplicity of worlds resonates well with Herman’s writing on storyworlds and Ryan’s (1991; 2001) possible world model.

In order to describe worldmaking, Goodman (1978) proposes several processes such as composition and decomposition, ordering, and weighting. These processes, however, are more relevant to Goodman’s symbolic worlds, rather than imagined narrative worlds and so I shall set out an alternative approach in the discussion which follows.\(^{49}\) With this in mind, I discuss two more pertinent approaches in this section: subject position, and affordance.

### 2.3.1 Subject Position

Within cinema studies and musicology, the theory of subject position has been used as a model of finding meaning within a text that accounts for the coherence of interpretations of the same text between different interpreters. The concept was originally developed from cultural theory and subsequently adapted to the study of film. As a hermeneutical paradigm, subject position proposes that interpretations are delineated by a text through the determination of possible ways in which the interpreter engages with discourse. The concept has been applied in the field of musicology of the Western art music tradition to model musical meaning (Cumming 1997; Clarke 1999; Whittall 2000; Kramer 2001). In the field of popular music, Nicola Dibben has also applied the theory in her study of Pulp (2001a) and a later piece exploring the work of Björk, in which she argues that music forms part of the cultural construction of Western emotion and therefore offers a means of experiencing and

\(^{49}\) Goodman himself is clear when commenting that these processes are “ways that worlds are made. I do not say the ways. My classification is not offered as comprehensive or clearcut or mandatory” (ibid.: 16, italics in original).
conceptualising subjective emotion (2006). For Zagorski-Thomas (2010), it also provides a helpful basis to the analysis of phonographic staging in recorded music because it proposes that perceptual information connotes social meaning.

The theory of subject position usefully suggests that the meanings which listeners draw are, to some degree, inscribed in the music. Furthermore, as Dibben (2006) explains, it is consonant with the proposition that listeners may personify an experiencing consciousness in music (as Cone, and Watt and Ash also suggest) and draw meanings in response to it. Indeed, Clarke (2005) later connects this with his ecological approach by arguing that musical meaning is afforded to the listener through his/her interaction with a subject position — or, speaking more loosely, ‘perspective’ — determined by a text. Yet, as Negus (2012b) cautions, subject position places a large emphasis on the text which can undermine the interpretative authority of the listener and lead analysts to make claims of meaning on behalf of diverse and hypothetical or unknown listeners. As subject position accounts for an implicit, textually-determined property of a recording, and not the real listener, applications of subject position can go no further than accounting for an idealised interpreter.

2.3.2 Affordance

Within musicology, several scholars have demonstrated an interest in ecological perception — the study of how an organism perceives and interacts with its environment — to describe musical meaning in a way that highlights the role of both the listener and text in interpretation. Ecological thinking as a model of musical meaning was proposed most comprehensively by Eric Clarke’s Ways of Listening, which draws from James J. Gibson’s influential work in visual perception. Rather

50 We will return to this issue of perspectivity in Chapter 5.
than modelling meaning as inscribed and transmitted, ecological approaches such as Clarke’s regard meaning to be constructed by the listener as a response to what the text affords to him/her.

In Gibson’s work, he proposes that objects within an environment offer a perceiving organism possible forms of interaction. As he wrote:

> [w]hen the constant properties of constant objects are perceived (the shape, size, color, texture, composition, motion, animation, and position relative to other objects), the observer can go on to detect their affordances. I have coined this word as a substitute for values, a term which carries an old burden of philosophical meaning. I mean simply what things furnish, for good or ill. What they afford the observer, after all, depends on their properties (1966: 285).

Hence, affordances are, according to Gibson, dependent on the properties of constant objects, which he terms invariants, as an observer perceives them. This places a particular emphasis on the objects themselves and may suggest that affordances are objective properties of an environment. Clarke (2005: 37), however, explains that his reading of Gibson’s theory views affordances to be determined by the organism in conjunction with the environment. In his subsequent investigation in the applicability of Gibson’s ideas to the design of manufactured objects, Norman (2013), a student of Gibson and specialist in design interaction, adopts the phrase “perceived affordances” to clarify this. 51 Clarke’s derivation of ecological perception is nevertheless compatible with Norman’s understanding of perceived affordances and provides the basis for my position that the persona and phonographic environment

---

51 Other commentary makes Norman’s decision explicit and supports the reading of Gibson’s affordances as objective properties of an environment (Norman 1999).
afford to the listener ways of constructing a narrative world. As such, narrative worlds are subjective, but open to analytically informed discussion.

In his discussion of narrative comprehension and the negotiation of narrative worlds, Herman similarly speaks of affordance in terms of a reader’s possible opportunities:

in arguing that interpreters of stories use textual patterns to frame, to the extent required by their engagement with a given narrative, answers to questions about the WHEN, WHAT, WHERE, WHO, HOW, and WHY aspects of a storyworld, I mean to suggest that recipients of narrative similarly exploit textual cues as affordances for negotiating the storyworld in question (2013: 48, italics in original).

Here, Herman uses the term affordance to refer to the offering of possible interpretative routes and does not explicitly refer to affordance in relation to ecological perception. Nevertheless, although he speaks of textual cues, his notion that they afford possible ways of negotiating a storyworld adds emphasis to the personal nature of narrative experience alongside the active process of worldmaking. This also departs from Goodman’s concern with representation in terms of symbols to instead consider possible action-consequences offered by a text. This leap is mitigated somewhat, though, by Goodman’s broad understanding of symbols and, as he does not insist on the same arbitrariness of several schools of semiotics, there is some compatibility with the (non-arbitrary) concept of affordance. Indeed, the relativist approach which Goodman adopts in his theory of worldmaking could be read from either a deconstructionist or structuralist perspective, yet as this is left unclear, it is difficult to assess the extent to which theories of affordance depart from his understanding of symbolism. I also advocate affordance rather than adhere to Goodman’s understanding in order to preserve the interpretative agency of the listener, for, as Kennett (2003) observes in relation to Tagg’s method, semiotic
schemes more often take a structuralist position that regards meaning to be pre-determined within a text.

2.4 Worldmaking in Radiophonic Drama

As it is straightforward to create phonographic simulations of particular spaces or places, recorded music may rapidly afford to the listener ways of imaging a narrative world. To point to a basic example, the first second of Daft Punk’s “Giorgio by Moroder” consists of various voices speaking simultaneously, spread across the stereo field, accompanied by two chinking sounds. The timbre of the voices suggests comfortable conversation, whilst the chinks resemble two gentle impacts on ceramic. Taken together, these afford to the listener imagining — and accommodating oneself within — a restaurant, which provides a contextual framework to narrativising the track. This example shows that aspects of the phonographic environment afford rapid worldmaking by prompting one to furnish or supplement a narrative world using knowledge of external experiences.

Further discussion of phonography in relation to worldmaking can be found by considering ways in which electroacoustic music and radiophonic drama evoke narrativity in relation to the recorded format. As Wishart (1986; 1996: 129-161) discusses in his theory of landscape, the use of concrete sounds in phonographic art forms allows composers to direct the listener towards (that is, the sounds afford) possible associations which allow him/her to imagine a given space or setting without the visual domain. Richard Hand and Mary Traynor (2011) provide a thorough grounding to the theory and practice of radio drama, suggesting a position

For Kersten Glandien (2000), there lies a close relationship between these fields, as radio provided not only the infrastructure for Schaeffer’s and Stockhausen’s work in electronic music, but also the seeds of their aesthetics which champion the use of acousmatic sound.
reminiscent of Levinson — that radio drama is principally dramatic but can be narrativised by the listener. Indeed, they regard radio’s strength as a dramatic medium to be its capacity to provoke imagination (ibid.: 33) and hence regard the listener as central in the interpretative process. In one case study, the pair compare two dialogue-less radio dramas, Andreas Bick’s “A Pot Calling the Kettle Black” and Andrew Sachs’ “The Revenge”.

First broadcast in June 1968, “The Revenge” was a novel undertaking for the time as it incorporated a number of actors without dialogue. As the play opens, the sound of birds and running water with subtle reverberation afford imagining an outdoor, wooded location as a siren begins in the distance. The sound of breathing is close-miked to simulate the aural perspective of the protagonist, while groups of military personnel run in close synchrony nearby. Having crossed under water, the protagonist acquires a motorbike and breaks into a house and ultimately drowns its occupant. The success of the play as a stand-alone narrative, though, is open to some discussion and indeed, critical responses vary. Hand and Traynor, for instance, suggest the omission of speech restricts the play’s narrative sophistication. Whilst they suggest the play achieves its aims, “the narrative is simple and linear, and the locations are straightforward and heavily signposted, because this is all that is possible without speech” (2011: 67). Similarly, Shingler and Wieringa observe that its use of sonic materials are effective in denoting location and action, but that the play cannot communicate an explanatory background to the story. “By the end of the play,” they write, “we know what has happened, how it happened and where it happened but have no sense of who the characters are, what their relationship is (or was) to each other nor what motivated this act of revenge. In other words, we have only half of a story” (Shingler and Wieringa 1998: 53). The success of “The Revenge” as narrative work also requires particularly engaged, uninterrupted listening of a kind.
that, as Hand and Traynor (2011: 60) note, is uncommon in the sphere of radio. Moreover, the play was broadcast at a time in which average consumers could not simply replay or rewind a broadcast as listeners might today. Hence, the ephemerality of the medium placed further demand for the listener’s attention in order to fully comprehend the play’s event-sequencing.

Today, recordings of “The Revenge” can be rewound and listened back. This different situatedness allows new means of engaging with the text which may appear to make it more promising for narrative interpretation. In either case, whilst “The Revenge” shows some limits and suggests that recorded sound in isolation is incapable of providing the same level of specific details that verbal or written accounts could, it represents all four of Herman’s basic elements of narrative, and offers a particularly rich example of how sounds can afford worldmaking. The majority of sounds used within the play are contextual and afford imagining a setting, such as the gentle flow of water and sound of birdsong which afford an outdoor wooded area near a stream, or afford both event sequencing and worldmaking, such as the sound of footsteps on wet gravel which affords narrativising the protagonist as he runs towards the stream.

2.4.1 Analytical Concepts

With the example of “The Revenge” in mind, I would like to discuss three related analytical concepts which may help describe several of the ways in which the play — and recorded music generally — affords to the listener ways of worldmaking. The first aspect of “The Revenge” which I would like to highlight is the use of Foley or ‘found sound’ which allows one to rapidly identify a connection between a sound and the source which it specifies. Within electroacoustic discourse, this aural trace of a source in a gesture is often referred to using Smalley’s concept of source-bonding,
which he defines as “the natural tendency to relate sounds to supposed sources and causes, and to relate sounds to each other because they appear to have shared or associated origins” (1997: 110). Sounds with clear source-bonding afford furnishing one’s narrative world with the source and associated characteristics.  

The second theoretical concept I should like to introduce is Tagg’s concept of anaphones, which draws from the observation that experience of music is often conceived as multimodal, that is, it engages with multiple senses (Tagg 2013; see also Adlington 2003; Emmerson 1998; Smalley 2007). Tagg describes the anaphone as a “musical sign type bearing iconic resemblance to what it can be heard to represent” (2013: 582). Tagg proposes five categories of anaphones which call upon visual, kinetic, tactile, spatial, or social experience. Consistent with his argument supporting the multimodality of musical listening, he notes that descriptions tend to be synaesthetic-aesthetic, citing adjectives such as “rough,” “bright,” “rich,” or “cold” (ibid.: 305).

Whilst Tagg writes from a semiotic perspective in which the operation of anaphones is attributed to iconicity, we can think of anaphones in terms of affordance such that the specification of an invariant affords comparison to, or recollection of, particular sensations. In this way, we might suggest that Kraftwerk’s use in “Trans Europe Express” of flanged noise, its rhythm, and the application of audio delay affords an impression of the kinetic experience of train travel. An alternative example, of a spatial anaphone, is offered in the opening of David Bowie’s “Space Oddity”, in which the environment consists of a strummed guitar, electric guitar harmonics, bass, and snare. In the soundbox, the strummed guitar is panned entirely right, in symmetry to the bass and snare on the left, whilst the guitar harmonics are

53 A telephone call, for instance, specifies not only the caller if we know them, but also their mood, sex, age, and accent.
located in the centre.54 The sparsity of musical material and the placement of sound-sources within the soundbox emphasises the conspicuous separation of instruments analogous to the spatial separation of the song’s protagonist, Major Tom, from earth amidst the vast emptiness of space. As the song concludes, an instrumental section also exemplifies a kinetic anaphone through undulating unquantised tonal gestures which waver through the stereo space to evoke the destabilisation of Major Tom’s ship.

The final hermeneutic concept which I would like to introduce in relation to phonographic worldmaking affordances is phonographic staging, that is, the manner in which a sound is presented within the phonographic environment. This theory is often attributed to William Moylan’s (2002) model to describe the way in which sounds are phonographically presented. As Serge Lacasse explains, “the phonographic staging model [...] aims to describe the effects following the manipulation of four main categories of sound perception through recording technology: loudness, space, time, and timbre” (2005: 2). Moylan’s original model was intended to address the construction of a spatial image within recorded music, though Lacasse’s (2000) reception testing explored several affective implications of staging. In his tests, seven different sonic effects produced statistically significant responses and eight affective outcomes were identified, ranging from “benevolence”, “happiness” and “stability” to “distance” and “temporality”.55

54 Of course, this organisation of sound-sources is typical of the time, as the limited number of tracks which could be recorded onto tape required recordings to be mixed together on individual tracks and panned in groups, often either side and centrally within the soundbox.
55 Whilst this was facilitated by the recorded format, precursors exist that demonstrate previous interest in the presentation of musical sound for affective purposes. Several
Related research by Nicola Dibben supports the hypothesis that sonic design contributes to listeners’ perception of emotion and acoustic character in addition to source identification. In her work, Dibben contends that the act of musical listening involves hearing sounds “in terms of their sources and cultural specifications” (2001a: 166) and undertakes an experiment in which participants were played three audio clips in succession, of which one pair share an acoustic resemblance, and two share a common specification. Her findings indicate that both musicians and listeners with no musical training recognise specification of source with similar accuracy (although unsurprisingly musicians were more accurate in recognising specifications of genre and musical function). Extending her investigation to explore the range of possible specifications, a second experiment asked participants to describe what they hear (drawing from the same pool of audio clips). In this case, source identification was the dominant specification found, alongside other specifications such as acoustic characteristics, emotion, and social context.

More recently, Simon Zagorski-Thomas (2007; 2010) has set out to expand the concept of phonographic staging with what he calls functional staging, which is concerned with optimising sonic materials for a particular mode of consumption (in this sense, functional staging is closely related to mixing and mastering). Zagorski-Thomas (2009) has also proposed the concept of media-based staging, that is, phonographic staging which emulates particular forms of sonic transmission. This could include the use of distortion, compression, and bandpass equalisation to simulate a telephone conversation, or the addition of crackle to connote age by

romantic-era works reveal the considerable use of what we might call ‘acoustical staging’ in which composers such as Gabrieli, Berlioz, Verdi, and Mahler employ novel instrument placement to control the sonic properties of the performance heard by the audience.
sonically referencing vinyl records. Acknowledging these extensions, therefore, we can observe phonographic staging to have the capacity to afford localisation in relation to the narrative world by evoking the perspective of an implied listener.

In relation to timbre and phonographic staging, the voice has been emphasised as an important reference point. As Zagorski-Thomas notes, “adding a certain pattern of both harmonic and non-harmonic overtones [using overdrive distortion] conveys meaning through relating the guitar sound to the type of emotional human states that we associate with shouting voices – aggression and anger, for example” (2007: 195). Timothy Hughes observes similar associations between Kurt Cobain’s vocal technique and guitar style (2006), though goes further, citing work which describes the equivalence between overdrive audio effects and this vocal technique. Hence, Hughes suggests that Cobain’s acoustic simulation of the effect makes him unintelligible but capable of highly affective non-linguistic expression.

2.5 Accommodation

As highlighted in the General Introduction, narrative worlds afford experiencing individuals the opportunity to recenter their consciousness around the narrative world. For Gerrig (1993), a fundamental property of narratives is that they invite the interpreter to be “transported” to an imagined system of reality in this manner. Hilary Dannenberg similarly describes reader engagement with literary narratives in terms of a journey, suggesting that

the pleasure of the reading experience can [...] be grasped by conceiving of it as a journey of exploration into a new world — a journey whose very attraction resides in the

56 Although Zagorski-Thomas does not employ reception testing, Lacasse’s (2000: 160-164) data indicate that the telephone effect connotes age to his respondents.
exhilaration of jumping across and transgressing ontological boundaries and mentally relocating oneself far away from one’s true spatiotemporal or ontological level (2008: 21).

Dannenberg makes the argument, like Gerrig, that narratives afford the reader a sense of taking up residence in another world. David Herman, who has discussed this issue in relation to other media, calls this process accommodation, after David Lewis’ (1979) discussion of presupposition.\textsuperscript{57}

Whilst several works have explored accommodation in literary media,\textsuperscript{58} and considered narrative worlds in relation to non-literary media,\textsuperscript{59} the topic has not yet been applied widely in musicology. Nevertheless, the immersive capacity of music is widely documented throughout journalistic discourse. In the field of environmental psychology, Barry Blesser also makes a telling point relating to the consumption of

\textsuperscript{57} In brief, Lewis proposed rules of accommodation in relation to presupposition and permissibility which he uses in relation to the interpretation of a language game. In this context, Lewis suggests that “[i]f at time t something is said that requires component s\textsubscript{n}, of conversational score to have a value in the range r if what is said is to be true, or otherwise acceptable; and s\textsubscript{n}, does not have a value in the range r just before t; and if such-and-such further conditions hold; then at t the score-component s\textsubscript{n}, takes some value in the range r” (\textit{ibid.}: 347).

\textsuperscript{58} See, for instance, Gerrig’s (1993) discussion of transportation in relation to literature; Dannenberg’s (2008) discussion of immersion in relation to coincidence and counterfactuals in literary narrative; Ryan’s (1991) formulation of possible worlds theory; and, Segal’s (1996) proposition of ‘deictic shift’ in relation to narrative comprehension.

\textsuperscript{59} See in particular Ryan and Thon’s (2014b) collection \textit{Storyworlds Across Media} or Jenkins (2006) on transmedial narrativity and its relationship to worldmaking. Other texts such as Herman (2009a) explore the possibilities of oral storytelling, literary tales, and comic books; and Perkins (2005) discusses the pertinence of worlds andworlding to film.
music by posing that one attraction to loud musical playback is that “loud music makes a listener functionally deaf to everything but the music [...] In [a] real sense, loud music transports listeners into another aural space, moving them from the social space of people to the musical space of the performers” (2007: par. 7). Blesser is writing here in a general sense which highlights the importance of the (acoustical) phonographic environment, rather than the (cognitive) narrative world. However, he comes strikingly close to the following description from Richard Gerrig on what he also calls transport, here referring to narrative worlds: “[s]omeone (‘the traveller’) is transported [...] by means of transportation [...] as a result of performing certain actions. [...] The traveller goes some distance from his or her world of origin [...] which makes some aspects of the world of origin inaccessible” (1993: 10-11). When he/she returns, Gerrig notes that the traveller is ordinarily “somewhat changed by the journey” (ibid.). It is this two-fold capacity of music to immerse a listener in both an alternative acoustical and narrative space which affords arguably greater extents of transport than literary media.60

As we will come to explore in greater depth, particularly in Chapter 5, our relationship in relation to the narrative world of a text may vary through different narrativalisations. In “The Revenge”, for instance, the centrality of the protagonist within the soundbox, and the apparent movement of the point of listening through several narrative spaces invites the listener to form a narrative world from the perspective of the protagonist. In other words, this point of listening mimetically

60 See also Cox (2011) and references therein which connect musical interpretation with mimesis and embodiment, supporting the notion that music — as distinct from literature or other narrative media — effects immersion differently. We will extend this line of thinking in the following Chapter as we address mimetic aspects of the phonographic environment which may promote accommodation (see Section 3.4).
projects an experiencing consciousness, the implied listener, who here corresponds with the protagonist. In other cases, though, details of the recording make for more voyeuristic narrativisations which may involve the track affording to the listener the perspective of a third party or the perspective of a character who is temporally and/or spatially separate from the protagonist/narrator.

As mentioned above, accommodation can also be afforded through approaches to lyric writing in which the protagonist is understood to be addressing a real-life individual, as this may suggest correspondence with the listener’s reference worldview and hence allow him/her to imagine or assume a position within the narrative world. Two common instances of this are the unspecific use of “you” in a way which appears to address the real listener directly, and the protagonist/narrator’s presumption of knowledge. As Moore (2012b: 117), Schinko, Kiefer, and Huck (2011), and Durant (1984) have investigated, the unspecified use of the word “you” prompts us to qualify the addressee, or to even position ourselves in his/her place. Elton John’s “Your Song”, for instance, is sung from the perspective of a protagonist as he dedicates his song to an addressee. The protagonist here is at pains to explain the significance of his addressee, describing several hypothetical situations in which he would express his love, and his joy that his addressee is alive. Whilst the song offers several details about the protagonist, the addressee is entirely unqualified and referred to exclusively using second-person pronouns. In doing so, the listener is free to narrativise the addressee in terms of a hypothetical individual within the narrative world, an actual real-life individual, or as the real listener him/herself.

Kate Bush’s “Running Up That Hill (A Deal With God)” offers an alternative example in which lyrics afford accommodation and worldmaking through the use of assumed knowledge. Like much of Bush’s work, the song is written somewhat enigmatically, reporting an unqualified feeling, a deal of some kind, and the sense of
an emotional struggle. These are manifested through indeterminate lyrical phrases such as “that hill”, “how it feels?”, and “exchange the experience”. In each case, the lack of precision calls upon the listener to supplement his/her narrative world to fill gaps of indeterminacy using schemata from his/her reference world, hence linking the fictional realm of the storyworld with his/her lived experience and accommodating his/herself within it. Throughout the following chapter, we will extend this line of thinking to consider the role of external schemata to build verisimilitude into a narrative world.

2.6 Case Study: The Kinks - “Waterloo Sunset”

The Kink’s “Waterloo Sunset” offers a helpful example of narrative worldmaking in relation to recorded popular song, as it is centred explicitly around an actual location, yet departs from it to present a highly romanticised account of Waterloo Station as it would have been at the time of the song’s writing in the late 1960s. As Moore comments,

by 1967 [The Kinks] had begun singing songs focused very much on everyday, but ‘altered’ reality (a frequent experience under LSD was that the ‘ordinary’ tended to become ‘marvellous’). ‘Waterloo [S]unset’ can be seen as a forerunner of that hippy aesthetic (2012b: 142).

The track begins with a bassline descending by step in E Ionian from B (\(^5\)) alongside loosely strummed guitars until the descending figure is replaced with an instrumental quotation of the main refrain melody which comes later. After eight bars, the first verse begins to set the scene, consisting of 4 four-bar phrases in an AABA structure. The narrator starts by describing the polluted Thames rolling ever-

---

61 This analysis is based on the stereo mix of the song, which was originally released in mono.
onward. As he continues, the narrator starts to suggest some discomfort in noting the dizzyingly busy people, and the bright light of the local taxis. Nevertheless, he explains that he is comfortable without the company of friends and describes the paradisiacal feeling of watching the sun set over the city.

The sense of contentment relayed by the lyrics can also be linked to the melodic contour of the vocals. In the first, second, and fourth phrase, the vocals take a simple undulating gesture, which is approximately repeated three times, resting on the $^5$, $^2$, and $^6$:

![Fig. 2.3: Melodic contour sketch of the first vocal phrase in the verse of “Waterloo Sunset”](image)

This gradual descent requires less physiological effort to sing than upward progressions. In contrast, the third phrase begins as the vocal melody ascends from $^4$ to $^6$, before descending from a melodic peak of $^3$ back to $^6$. At this point, the verbal space of the vocals is also increased: whereas the first, second, and fourth phrases each last 16 or 17 syllables in a fairly rigid quaver-based pattern, the third phrase consists of only 6 syllables in the same duration. The third phrase also features an added second vocal line symmetrically opposite the first, towards the left of the soundbox.

Verses 1 and 2 are both followed by a ten-bar chorus, which begins with a laissez-faire “sha-la-la”. Here, the narrator continues by explaining how he looks over the world each day from the comfort of his window to avoid the bitter cold of the

---

62 For a full introduction to the concept of verbal space in the context of popular song, see Griffiths (2003).
evenings. He concludes by describing the sunset as being “fine”, but does not disambiguate between two possible understandings here: ‘fine’ meaning sufficient, or instead meaning beautiful.

In the second verse, two further characters, Terry and Julie, are introduced and take the focus as the narrator reports on their regular routine of meeting at the station on the same night each week. In contrast, however, the narrator characterises himself as too idle to leave his home and repeats the refrain from the end of the first verse. The musical material and verbal space of this second verse is virtually identical to the first, although the double-tracked vocals continue this time from the third phrase into the fourth.

So far, the lyrics present a measured but positive overall representation of the setting. In contrast, the final verse begins by comparing the crowds of people around Waterloo station with a swarm of insects, which evokes a hitherto-unseen negative outlook. The narrator reports, however, that upon crossing the Thames, Terry and Julie feel content and outside of any risk. This final verse is again followed by the refrain, although it moves from the first-person singular to the third-person plural, shifting the focus from the narrator to Terry and Julie.

For the most part, the placement of sound-sources within the soundbox is consistent. During each verse, the lead voice is placed at the right-most extreme of the soundbox and, as I have noted, a second voice is placed symmetrically opposite. They are joined by a drum kit and strummed guitar towards the left, bass slightly left of centre, and wordless vocals towards the right. Whilst the lead vocal track is mostly presented with minimal reverberation, additional reverb is added during the first chorus. In both choruses, further vocal lines are introduced in the final phrase, in which the vocal tracks are clustered around the centre of the stereo field. Substantial
reverb is also added from the third phrase of the final verse (2′30″), which comments on Terry and Julie’s self-sufficiency.

2.6.1 Setting

Throughout, the geographical setting of “Waterloo Sunset” is restricted to the local area of the narrator, principally as seen from his window. The construction of the soundbox is particularly telling in light of the vantage point (looking over the world from his window) reported to us in verses 1 and 2, which is also supported by the analogy of swarming flies, as though the narrator is looking upon the station from a height. Likewise, the lead vocals are panned to the right in each verse. Although the track comes at a time in which the normative positioning of instruments within the soundbox was still being established, positioning the vocals in this way enforces the suggestion that the narrator is looking from a distance rather than taking ‘centre stage’ himself.63 Indeed, the increasing use of reverb in the final verse when Terry and Julie are furthest away could be read as denoting the increasing distance between them and the narrator. Both the narrator’s apparent placement and his reports of the local area afford our construction of the setting for a narrative world.

2.6.2 Characterisation

The narrator provides perhaps the most immediate point of entry into the narrative world and consideration of the persona and phonographic environment allows us to

63 See Dockwray and Moore (2010) for a discussion on the development of a normative configuration of the soundbox at the time of the song’s recording which they call the ‘diagonal mix’. Moore explains that this configuration consists of “a lead voice, a snare drum and the harmonic bass (normally a bass guitar), which are situated centrally on a (very) slight diagonal” (2012b: 32).
make several inferences about his character. As discussed, the phonographic staging positions the narrator as an outsider and, indeed, his lyrics support this through his apparent disinterest with social interaction and by locating him within an apartment which he does not leave in the evenings or cold. Furthermore, despite his apparent fondness for the sunset, a profound sense of apathy is suggested by the singer’s vocal delivery, use of descending melodic contours, and unclear word choices.

Whilst the lyrics generally are somewhat terse, only two lines are devoted to Terry and Julie, leaving the listener to infer additional key details. Chiefly, although it is not stated by the persona, we might imagine that the pair are a romantic couple. Inferring from the opening of the final verse, the pair seem to have no fear of the bitter weather or dark evenings and instead find solace in one another’s company, suggesting they are more outgoing, and possibly younger, than the narrator. The location of the narrator looking from his flat also positions him as looking down — literally and, perhaps, metaphorically also — on the couple.64

2.6.3 Summary

Based on the information reported to us by the narrator, in addition to musical details such as the harmonic patterning and instrumentation, the song affords the construction of a marvellous narrative world, primarily from the narrator’s more melancholy perspective looking over Waterloo. The song begins from a basis in an actual, locatable geography, using seemingly ‘objective’ narration in the first verse, in which the narrator comments on the polluted river, reflecting the actual state of the

64 In cinema, for instance, such an elevated point of view looking down may be associated with dominance or judgement. In combination with the narrator’s apparent (deliberate) isolation, this might suggest a critical view of Terry and Julie.
Thereafter, though, the verses cast this setting in a surprisingly positive light. The track offers a rather optimistic description of Waterloo station at the time of the song’s original release in 1967, as it makes no mention of local crime, racial tensions which are now widely reported in online accounts, or pollution of the time. The song offers a flattering view of the station as Ray Davies and the band might have known it. By offering this alternative, possible take on London at the time, the song draws attention to the ability of songwriters to depart from a reference worldview to instead invite the listener to imagine an alternative narrative world. This also exemplifies how narrative worldmaking affordances are informed by perspectivity, a topic which I will return to in Chapter 5.

The narrative world centred around Waterloo Station from the narrator’s perspective in his apartment, some distance above, offers a framework within which we can narrativise the track. As Negus’ (2012b) commentary on the song highlights, the temporal dimension in this narrative world is represented as routine and repetitious due to the ever-rolling river, weekly meetings of the couple, and the narrator’s gazing upon the world outside each day. Indeed, this is also supported by the use of a fade-out at the end of the track, which implies a sense of continuity and ongoingness. If we take the position that the pair are a romantic couple, we can make sense of the track as a passage in which the two meet for a date across the Thames.

This example allows us to draw three general principles that I would like to highlight, and which form the basis of our investigation in the following chapter. Firstly, tracks open gaps in the information provided to the interpreter, whether in terms of event-sequencing or relating to the forces, agents, and objects of a narrative

---

65 Pollution of the river is documented by Wheeler (1969), who reports that some advances had been made to reduce pollution of the River Thames with some notable success, but that further work was ongoing.
world. To address these gaps, narrative worldmaking prompts the incorporation of external schemata from a reference worldview. Secondly, worldmaking is informed by the position of the narrating agent and his/her knowledge or attitudes. Finally, although the narrative world may resemble a reference worldview, the worldmaking affordances of a track may depart from actual states of affairs.

2.7 Conclusion

With Herman’s (2009a) ‘basic elements’ in mind, it holds that narrativity involves the imagining of a narrative world. The process of narrativisation which we discussed in Chapter 1 must, therefore, incorporate some (de)formation of a possible world inhabited by the agents of a narrative. In order to address the sorts of agents involved in this process, I have drawn in this chapter from Chatman’s model of narrative communication to outline six agents in relation to the narrativisation of recorded popular song. Although I have focussed on the role of the singer here, we can equally describe the environment in terms of real musicians, their recorded performance, and narrative agents (of the sort Liu-Rosenbaum finds in “When the Levee Breaks”). Indeed, in the discussion over the following chapters, we will consider how the different representations of the persona and environment contribute to narrative worldmaking.

Following this extension to our understanding of narrativisation, it is helpful to distinguish between three key different worlds which are also reflected in my development of Chatman’s model. Again, this follows the cognitive approach of Chapter 1 by thinking of these worlds as informed by the listener him/herself. Commonsensically, I have begun from a reference worldview, the interpreter’s understanding of a world in which we can say that the listener and musicians are actually located. Secondly, the phonographic environment constitutes our discourse
and affords particular means of engagement and interpretation. Thirdly, the narrative world refers to a mental model of a possible world which is constructed in response to a text. This scheme reflects pairings in Chatman’s model of actual individuals, their mediated representation in relation to a narrative discourse, and the characters they adopt.

To complement my cognitive understanding of narrativity and worlds, my approach draws from Gibson’s theory of affordance (particularly in the sense developed by Clarke’s adaptation to music). This allows us to maintain a position in which aspects of the text can be discussed in relation to personal meanings drawn without compromising the importance of the listener’s subjectivity. Hence, I take the position that the persona and phonographic environment afford to the listener ways of narrativisation and, alongside it, narrative worldmaking. In the following chapter, I will develop this further to emphasise the importance of the reference worldview in offering schemata which can be used to support narrative worldmaking, but for now I should like to emphasise that, as the reference worldview will differ between listeners, so too should we expect particularities of a narrative world to differ.

Recordings’ affordances of ways of imaging a narrative world raise further issues regarding the nature of the narrative world and its implicit relationship with the listener. The following three chapters each explore one of these issues and opportunities for analysis. In Chapter 3, we will address the degree of verisimilitude between the phonographic environment and persona with the reference worldview. This provides the basis for the discussion of temporality and the narrative architecture of narrative worlds in Chapter 4. Finally, this is developed in Chapter 5 to address perspectivity and the ways in which the persona and implied listener are mediated in recorded popular song.
3. Verisimilitude and Departure

Having established a model in which the phonographic environment affords to the listener ways of narrative worldmaking that frame his/her interpretation of a track, this chapter will consider the resemblance of truth which arises most often through a correspondence between a narrative world and the reference worldview. My premise for this chapter is encapsulated by Goodman, who writes that organisation is “not ‘found in the world’ but built into a world” (1978: 14, italics in original). That is to say, I argue in this chapter that narrative worldmaking involves the building of verisimilitude — the resemblance of truth — into a narrative world. In many cases, this verisimilitude is derived from a close resemblance to a reference worldview, though in other cases — including those in which a narrative world departs from a reference worldview — verisimilitude can be found in a sense of internal consistency.

For some, departure from an actual world can be seen as a source of pleasure: Fludernik suggests “a good deal of the charm of texts is to be found in the exotic and escapist scenarios they offer the reader” (2009: 55). On the other hand, Busselle and Bilandzic indicate an important connection between perceived realism and the concept of accommodation, the feeling of momentarily inhabiting a narrative world (see Section 2.5). The authors argue in particular that irrealist fiction compromises engagement with a literary narrative and its power of persuasion (2008), whilst realist film contributes positively to the viewer’s experience of narrative and his/her enjoyment (Bilandzic and Busselle 2011).

Before we begin, then, I would like to propose that we consider the phonographic environment as a mediated form of reality for, as Patmore and Clarke put it, “the creation of a recording is always the creation of a virtual world” (2007:
Using this analogy allows us to consider tracks in relation to the scale between what Thurman and Mattoon refer to as the *physical*, i.e. “correspondence to physical laws”, and the *abstract*, i.e. “novel environments” (1994: 57). As I discuss momentarily, what Thurman and Mattoon describe as the ‘physical’ pole is, strictly speaking, inaccessible to recordings due to the limits of mimesis, but tracks may employ mimetic strategies which afford the building of verisimilitude into a narrative world.

Culler (2002) suggests that the extent of perceived verisimilitude affects the interpreter’s ‘naturalisation’ of a text. Importantly, it is Culler’s *naturalisation* which forms the basis of Fludernik’s subsequent concept of *narrativisation* which we use in this project. As Culler argues regarding the ascription of verisimilitude to a text, “[w]hatever one calls the process, it is one of the basic activities of the mind” (ibid.: 161). He continues by outlining several categories of vraisemblance, of which two are particularly relevant to this chapter: *the real*, wherein a text resembles the natural

---

66 Similarly, as Evan Eisenberg has highlighted: “[t]he word ‘record’ is misleading. Only live recordings record an event; studio recordings, which are the great majority, record noting. Pieced together from bits of actual events, they construct an ideal event. They are like the composite photograph of a minotaur” (1986: 109).

67 Thurman and Mattoon describe this scale in terms of verity, though, as mentioned above, I employ such thinking in terms of verisimilitude. Although the pair derive the term verity from the Latin ‘veritas’ or “true to life” (ibid.: 56), their definition as “the degree that our natural environment is represented in a virtual environment” (ibid., my emphasis) suggests that they use the term in the same way which verisimilitude is understood in this chapter.

68 Culler and others use the French term ‘vraisemblance’ in place of the English verisimilitude, after Genette (2001).
world, and cultural vrai semblance, in which a text adheres to shared cultural knowledge.  

3.1 Mimesis and Minimal Departure

As we encountered it in Chapter 1 (see Section 1.1), mimesis is a form of narrativity that involves enacting a story, though taken more broadly, mimesis is generally used to refer to the simulation of, or resemblance to, reality. To avoid confusion here, I shall discuss this second, broader understanding of mimesis in terms of verisimilitude and refer to the aspects of a recording which evoke it as ‘mimetic strategies’.

Two general forms of mimetic strategies can be derived following Halliwell’s (2000) commentary on Goethe’s Über Wahrheit und Wahrscheinlichkeit der Kunstwerke ("On truth and the appearance of truth in works of art" Halliwell ibid.: 1, his translation), in which he writes on two types of mimesis:

first, the idea of mimesis as committed to depicting and illuminating a world that is (partly) accessible and knowable outside art, and by whose norms art can therefore, within limits, be tested and judged; second, the idea of mimesis as the creator of an independent artistic heterocosm, a world of its own, though one that, as in Goethe’s case, may still purport to contain some kind of “truth” about, or grasp of, reality as a whole (ibid.: 5).

---

69 Culler also describes three further categories: models of a genre, such that a text adheres to genre conventions; the conventionally natural, that is, vrai semblance which explicitly capitalises upon models of a genre; and, parody and irony, in which a text acknowledges the mediated nature of another text, in relation to which it must be understood.
Here, Halliwell appears to distinguish between 1) correspondence between actual and artistic worlds; and, 2) internal consistency within an imagined or represented world.\textsuperscript{70}

Following Halliwell’s distinction in the context of this project, some mimetic strategies prompt the interpreter to incorporate external schemata derived from a reference worldview, as with Thurman and Mattoon’s\textit{physical} environments. Examples of this approach appear to feature minimal audio processing and include sound-events which can be traced to physical sources. Meanwhile, examples of mimetic strategies that function in terms of internal consistency include concept albums, transmedia works, and songs that supplement the narrative world of another text through intertextual referencing.

Examples of both types of mimetic strategies can be found in relation to literature. Fludernik, for instance, exemplifies how the correspondence between a text and an actual world evokes verisimilitude by observing that “realistic novels refer to aspects of reality which are already familiar to readers” (2009: 55).\textsuperscript{71} Meanwhile,

\textsuperscript{70} Busselle and Bilanzic emphasise the equivalence between mimesis and verisimilitude in these two positions by observing a similar distinction between ‘external realism’, the correspondence of mental models with external reality (c.f. mimesis as imitation of another world), and ‘narrative realism’, or “plausibility and coherence within the narrative” (2008: 256) (c.f. mimesis as internal consistency).

\textsuperscript{71} Similarly, Wolf (2004) suggests that we interpret narrative worlds \textit{in terms} of life, but not as \textit{being} actual. Alber (2009) also posits five strategies which attempt to explain or naturalise impossible scenarios: 1) explaining impossibility as the product of an unreliable mental state; 2) taking the impossible to be symbolic of possible other states or events; 3) modelling the narrative as allegorical; 4) conceptual blending of schemata, and; 5) infer parameters from
Dannenberg suggests that internal consistency (in the sense of what she calls ‘intradiegetic connective systems’ below) encourages believability in certain genres: realist texts (and semirealist texts, such as the genre of science fiction) attempt to camouflage the ultimate, extradiegetic causal level of the author (who actually writes the text and thus causally manipulates all events within it) by constructing a narrative world with its own intradiegetic connective systems. If these are convincing, the reader is encouraged to believe in the internal logic and autonomy of the narrative world and thus that it is a ‘re-creation’ as opposed to a fictional ‘creation’ (Sternberg 1985: 99). These connective systems are created by the text’s suggestion to the reader that she should understand the narrative world’s configuration of events and characters using the same basic cognitive operations that she uses to make sense of the real world (Dannenberg 2008: 25).

3.1.1 External Schemata and Minimal Departure

Both types of mimetic strategies identified above rely on schemata from external sources — either learnt from an exterior worldview or derived from another text — as a point of comparison. An important function of mimetic strategies, though, is that they may also afford the incorporation of external schemata to supplement the worldmaking cues of a text. This is especially pertinent given the brevity of recorded popular song, which lends itself to schematic portrayals of a narrative world. Roman Ingarden has similarly addressed “schematised worlds” in literature, and particularly “spots of indeterminacy” which refer to areas in which insufficient information is provided to the reader, prompting him/her to ‘fill the blanks’:

the scenario which, when applied to the storyworld logic, justify and permit unnatural behaviours.
Among the spots of indeterminacy one must distinguish between those which can be removed purely on the basis of textual supplementation and those for which this does not occur in the same sense [...] In the latter [...] the textually established states of affairs are not sufficient for designating a strictly circumscribed manifold of possible completions. In that case, each ‘completion’ or approximate completion actually effected in this way is fully dependent on the reader’s [...] discretion (1973: 253).

In other words, ‘spots of indeterminacy’ prompt interpreters to supplement a text’s worldmaking affordances in some way.

Ryan (1991; 2005) suggests that narrative worldmaking is governed by a ‘rule of minimal departure’ which suggests that, unless a text instructs to the contrary, the schemata used to resolve indeterminacy are derived from what I refer to as the reference worldview. In her own words: “we project upon worlds everything we know about reality [and] make only the adjustments dictated by the text” (1991: 51). This imperative to construct worlds based upon our existing knowledge of external schemata leads us to often build verisimilitude into a narrative world. As Goodman puts it: “worldmaking as we know it always starts from worlds already at hand; the making is remaking” (1978: 6).72

3.2 Oneiric Aesthetics and Departure

Whilst Ryan’s rule of minimal departure helps us to understand verisimilitude in relation to various media, recorded popular song may — and often does — suggest a degree of departure from a reference worldview. Regarding the phonographic environment, for instance, Ragnhild Brøvig-Hanssen and Anne Danielsen (2013) observe that tracks may present listeners with improbable acoustical spaces, which,

72 Segal similarly proposes that there is an imperative to model a narrative world in the first instance after the actual world; that the “default condition is verisimilitude” (1995: 72).
they argue, listeners ‘naturalise’. We may similarly see points of departure in relation to the persona, apparent derivation and combination of sound-events, and the degree to which the recording appears to depict a possible performance. These four aspects may contribute to an oneiric or ‘dream-like’ state of affairs, which I would like to show using the example of Aqua’s “Halloween”.

“Halloween” opens with an introduction section featuring a string pad and “ch” and “ah” vocal sounds, before a phone rings, is picked up, and the protagonist answers. A brief dialogue unfolds between the protagonist and a male voice staged using the telephone effect, which threatens the audibly scared protagonist. Thereafter, musical discourse begins in earnest as the protagonist starts to sing alongside newly added instrumentation. The track continues in a verse-chorus structure, in which both verses are followed by a prechorus that features a further dialogue between the antagonist and protagonist (following the same dialogue both times). A bridge is also added at 2’46” between two final choruses, in which the texture thins to a distorted guitar and half-time percussion as the antagonist returns with a monologue.

The 29” introduction provides several important worldmaking cues which establishes a tongue-in-cheek spooky mood. These include: the sounds of thunder and howling wolves which afford imagining a stormy night and threat of unseen wild animals; the timbre of the vocalist which affords the ascription of fear to the protagonist; the use of a phone ringing and staging that implies a geographical separation of protagonist from unseen caller; and, the caller’s malevolent timbre. The sound-sources also position an implied listener within the protagonist’s house — perhaps her living room — as she receives the call from her addressee.
Whilst the use of Foley appears to implicate us initially as onlookers within the narrative world, the musical discourse which follows it raises several issues. For one, as the music appears to occupy a different acoustical space and the vocals are staged differently, a dislocation is implied between the introduction, which emulates to some extent the use of Foley in radiophonic drama, and the musical discourse, which is not clearly anchored in the manner to a particular narrative setting. Secondly, it would be unusual for an individual to sing in response to a threatening phone call, and her delivery here is controlled and secure, rather than afraid.

We might hear the musical discourse from the first verse onwards as a sonification of the protagonist's internal psychological state as she receives the call. Yet, whilst the introduction is rich in worldmaking affordances through source-bonding and dialogue, the staging of the Foley suggests that the implied position within the room alongside the protagonist is not entirely coherent. During the opening, the antagonist's voice is staged using compression and equalisation to schematise the timbral quality of a phone caller as if we hear through the ears of the protagonist, whose voice is not staged using this effect. The howls and thunder, though, are somewhat clearer, with a particularly noticeable proximity within the soundbox and high frequency content. The sound of picking up a telephone handset specifies a phone suitable for the home, which thus affords worldmaking in an indoor setting, but this contradicts the clarity of the howls and thunder, which would likely observe different sonic characteristics if heard from an interior perspective.

Several aspects of this example, then, illustrate points of departure from what we might expect based on reference worldviews. Firstly, the phonographic staging and landscape of the introductory section is incoherent in that it schematises an interior space on a thundery day but does not conform to the acoustical properties we might expect were we actually in such an environment. Secondly, we begin as
onlookers upon an unfolding situation as the protagonist receives a call, but the musical discourse which follows it does not appear to incorporate a progression of narrative time. Instead, during the verses, the protagonist simply expresses her thoughts from one moment and, in the prechoruses, the antagonist and protagonist repeat the same dialogue. Thirdly, the liberal use of cliché in terms of the Foley, high-pitched dissonant strings, and instrumentation of the introduction trivialises the possible severity of the situation. Finally, the end of the introduction coincides with a conspicuous change in texture which removes the Foley elements, and with it, the suggested point of listening in the room with the protagonist.

Regarding incoherent phonographic staging, Zagorski-Thomas (2014) offers a helpful observation regarding a normative production aesthetic of recorded popular song. He proposes the term *sonic cartoons* to describe the schematised representation of acoustic images in recorded music. Often, this privileges sonic clarity over acoustic accuracy by, for instance, designing an unrealistic reverb response to prevent ‘muddy’ lower frequencies from dominating the mix, or manipulating vocal timbre and dynamic range to enhance intelligibility and penetrate the mix more effectively.

Zagorski-Thomas contends that even schematic acoustical representations of a given scene suffice to afford important interpretative considerations without absolute acoustical accuracy. Returning to our example of ‘Halloween’, whilst we would expect different phonographic staging of the Foley elements and different vocal balancing were we to actually inhabit this setting, the schematisation of acoustical realism allows for greater phonographic clarity whilst still allowing us to draw potentially useful worldmaking affordances from the track. One other obvious example is offered by Billy Joel’s “Piano Man”, in which the persona adopts the position of a piano player at a bar one Saturday night and tells us there are several customers. However, in the phonographic environment, the gathered crowd are
nowhere to be heard. As Zagorski-Thomas describes in relation to sonic cartoons, realism or accuracy are not necessary preconditions for affective production and the same here can be said for worldmaking.

The other, non-phonography-related observations regarding ‘Halloween’ further support a dream-like quality in which the track eschews the logic one might expect based on a reference worldview. Although this has attracted limited attention in relation to popular song, a similar discussion concerning the relationship between narrative worlds and reference worldviews can be found in the writing of Vlada Petric (1989), which addresses films directed by Andrei Tarkovsky. Petric begins by summarising a general belief that the director’s works evoke what Tarkovsky himself called “an oneiric air — a dreamlike impact — which resists the audience’s need to verify the logic, as well as the credibility, of the events presented on the screen” (ibid.: 28, italics in original). Petric claims that this ‘oneiric air’ is reflected in the way in which Tarkovsky’s films eschew realist imagery and conventional event-sequencing. With this in mind, Petric considers cinema in terms of dream-like forms which we nevertheless engage with — at least temporarily — as though they were real, despite their mediated, oneiric appearance. That is, we naturalise points of departure in pursuit of minimal departure.

One of the clearest points supporting the idea of an oneiric aesthetic in recorded popular song is the role of the singer. In ‘Halloween’, why would the protagonist sing in such a straightforward manner and indeed, as Moore asks, “how many of us, in order to communicate with another, will preferentially choose to sing?” (2012b: 188). This topic is discussed in greater detail within discussions of art song. Abbate’s (1991) study of opera, for instance, questions whether narrative agents are aware that they are singing and proposes that they are party to music which is inaudible to others. Accordingly, she suggests that music is not generally a part of the
world inhabited by the characters. Cone (1989) has suggested that this uncertainty between what he calls ‘realistic song’ (in which characters sing as part of the story) and ‘operatic song’ (in which song is used as a substitute for what would ordinarily be speech) is a central component of operatic drama, whilst Penner (2013) has argued against Abbate, again regarding opera, that the act of singing is instead part of the ontology of the fictional world.73

For agents to be unaware that they are singing or for singing to be a standard mode of expression raises similar questions regarding the ontology of the other musical textures: do we take it that the music occupies the world with the protagonist? To draw attention to the musical environment — and the mediated, constructed situation of a musical performance — is uncommon for personae. Take, for instance, P!nk’s “Raise your Glass”: at 2‘28", the persona appears to enter a bar early before the final refrain and curses before repeating herself the following bar and continuing as expected. Similarly MKTO’s “Classic” begins as the persona asks “hey, where’s the drums?” and, in response, a drum fill introduces the musical discourse which continues for the rest of the track. In both these cases, personae demonstrate an awareness of the surrounding musical environment and so draw attention to the fact that they are part of a musical performance. This has ramifications for worldmaking, as it shifts the focus to a performance scenario rather than some form of narrative situation, and questions the ontology of the music and singing in relation to the narrative world.

In relation to the oneiric aspects of recorded popular song, we can also ask questions of temporal sequencing. As Negus (2012a) reminds us, pop songs may creatively manipulate the flow of narrative time or even mediate our own experience

73 I discuss the ontological role of singing and the construction of temporal and diegetic architectures in full in the following chapter.
of linear time. Accordingly, songs can also slip between different temporal locations within a larger narrative structure, and depict anything from a timeless ‘perpetual present’ to selections of an entire lifetime or longer in the course of a few minutes. In ‘Halloween’, for instance, narratological temporal progression appears to be suspended during the verses and choruses, whilst the prechoruses repeat the same material.

3.3 Approaches to Musical Verisimilitude

To recap the key ideas so far, I have proposed that verisimilitude is built into a narrative world, most particularly in response to mimetic strategies that prompt the incorporation of external schemata to supplement the worldmaking cues of a recording. Ryan proposes that interpreters generally construct narrative worlds which minimally depart from their understandings of reality, but as we have just seen, recorded popular song may illustrate a more oneiric position requiring some degree of naturalisation. As Eric Clarke observes, “sounds can specify a virtual domain that both abides by, and stretches or defies, the normal laws of physics” (2005: 70). In this section, we will discuss three typologies of verisimilitude in relation to recorded music that provide the basis of my approach, in order to account for the dominant sorts of mimetic strategies which afford naturalisation.

Moore (2015) builds on a previous discussion of phonographic realism in relation to musical cognition (2012b: 245-258) to address the changing role of electronic music technology in folk music. He begins by distinguishing between two forms of production: a preservationist approach which seeks to document a performance, and the alternative in which the track is conceived using electronic
technology.

Moore goes on to develop four positions which assess the verisimilitude of the phonographic environment: *documentary*, as though one could have been at the space and witnessed the same performance; *enhanced*, in which some details of a recording are altered though one could, broadly speaking, still imagine being present at a live performance from which the recording is derived; *blended*, whereby some details of the track are more clearly constructed through electronic technology and could not be otherwise replicated; and, *fantasy realism*, which Moore uses to describe tracks that are conceived using the recording studio and dependent upon associated technologies.

Moving to the field of electroacoustic music, I would like to revisit Trevor Wishart’s (1986; 1996) concept of *landscape* which describes the imagined sound-sources in a recording, the phonographic acoustical space in which they are situated, and their distribution within this space (see Section 2.4.0). Prefacing his discussion, Wishart briefly writes on different types of Foley or sound effects which could be used in radiophonic drama or art music as ‘contextualising cues’ with which a listener

---

74 A similar approach can also be seen in the writing of Stephen Cottrell (2010), which presents a spectrum in record production between *capturing* musical sound on the one hand, and *creating* it on the other. Both Moore and Cottrell highlight that recording as a means of preservation was often undertaken by ethnomusicologists or individuals aiming to capture an unmediated performance. Aesthetic or creative uses of production technology, conversely, are often intended for a recreational listening audience.

75 Lars Nyre also describes a classical music recording aesthetic in terms of ‘documentary realism’ which “convey[s] the musical performance as vibrantly and realistically as possible, and nothing else” (2008: 12).
could identify a landscape. Wishart identifies three positions, which reveal some similarity with Moore’s approach: actuality, staged, and studio. For Wishart, actuality corresponds with “an actual recording of an event” (1996: 137) in a broadly similar manner to Moore’s documentary category. Staged refers to an acoustical recording of a simulation or reenactment of an actual event which could not otherwise be recorded (for instance, reenactments of historical settings). Studio sounds in Wishart’s terminology are artificial sounds created with sound processing or synthesis to simulate an acoustical sound source. Wishart also proposes a final category, mixed landscapes, which he does not define further.

Source-cause identification is particularly relevant to Ambrose Field’s (2000) later examination of verisimilitude and extramusical signification in electroacoustic music. Field suggests that the phonographic nature of recorded music affords composers the use of verisimilitude or departure as compositional devices. By

\[\text{\textsuperscript{76}}\] Wishart makes it clear that “it is important [...] to differentiate the idea of landscape from that of association as it is frequently used in reference to programmatic music” (ibid.: 130; italics in original). Whilst, he explains, we might associate the final movement of Tchaikovsky’s Manfred symphony with a narrative situation — perhaps the protagonist’s arrival in hell, per the programme notes — the landscape remains musicians playing instruments. Note, however, that Hand and Traynor (2011) argue that an associative listening mode is necessary for sounds to have narrativity.

\[\text{\textsuperscript{77}}\] A similar proposal is put forward by John Young (1996), who discusses what he describes as “virtual” and “surreal” sound worlds in electroacoustic music, particularly in relation to source-identification. Young considers the construction of landscapes through the processing and transformation of sounds as a form of “abstraction” and/or the positioning of sonic objects alongside one another in ways which do not or could not correspond to actual situations. He also briefly mentions the role of space and performance (in the context of mixed media works), though I shall diverge from his implementation below.
judging the coherence of sounds’ (extramusical) contextual implications, he outlines four types of verisimilitude which he holds to be common in electroacoustic music: *hyper-real* in which a simulation of an actual landscape is indistinguishable from the real thing; *real*, in which the phonographic environment has not been manipulated or simulated; *virtual*, in which the environment has clearly been simulated, and; *non-real*, in which the environment seems to bear no correspondence with an actual landscape. Field cautions, though, that extensive departure from recognisable or familiar landscapes often diminishes the strength of extramusical signification.

### 3.3.1 Beyond Typologies of Verisimilitude

Each model above is constructed by accounting for some combination of: the imagined sound-sources, apparent cause, and phonographic acoustical space. However, I would like to offer a sympathetic critique of their applicability in the specific case of recorded popular song as we move toward an approach which I shall use. Firstly, having argued in Chapters 1 and 2 for the importance of the listener in interpretation, we must adopt a position that is sensitive to the subjectivity of perceived verisimilitude. Brøvig-Hanssen and Danielsen (2013) argue, for instance, that perceptions of spatially surreal recorded music change over time, whilst in the field of cinema studies, Sarah Kozloff writes that the vocabulary of resources which appear to connote realism is in continual flux:

> when a text is referred to as ‘realistic,’ one is actually saying that it adheres to a complex code of what a culture at a given time agrees to accept as plausible, everyday, authentic.

> These conventions change through history — what strikes one generation as incredibly realistic may strike another as highly mannered (2000: 47).
We might therefore say that perceptions of realism are subjective, and that perceived verisimilitude is contingent on the presence of a sound within a listener’s vocabulary of known ‘realistic’ objects.

Secondly, recordings cannot be — strictly speaking — *actual*, to use Wishart’s term, and rarely adhere to what Moore calls ‘documentary realism.’ Both these categories describe recordings in which sound is documented without further processing, but true mimesis is impossible because tracks exist outside of the same geographical or temporal constraints of the original recording space, and without a physical presence. Indeed, Max Bruinsma (1990) emphasises that producers working with recorded sound are already engaged with the artificial, as recordings are representations which exist outside of an actual environment which they may seek to document.\(^7\)

Moore’s discussion also introduces an issue concerning listener familiarity. When discussing his *blended* category, he refers to a recording of “The Floo’ers of the Forest” by Dick Gaughan, in which Gaughan takes up the electric guitar during the bridge in place of singing. Moore argues that this exemplifies blended realism as an acoustic guitar line — also played by Gaughan — is also present during the bridge. Were this performance to take place without the assistance of recording technology, Gaughan would be unable to play both simultaneously, and so some internal negotiation is required to make sense of the incoherent phonographic situation.

\(^7\) Furthermore, even in an unedited continuous recording without mastering, needs such as finding a reasonably quiet space, rerecording in the case of interruption or performance error, and consenting to be recorded in the first place all result in a performance which is at least in some respect mediated. As Wishart’s *staged* category emphasises, whilst we could think of some recordings as documentary in a phonographic sense, their authenticity in the sense of spontaneity cannot be known without additional information.
However, unless the listener is sufficiently familiar with Gaughan’s guitar technique, we would be hard pressed to identify that he is multitracked with access only to the recording.

Finally, each of the approaches above provide a typology that attempts to categorise different forms of verisimilitude, which can all be imagined as points along Thurman and Mattoo’s verity spectrum. Yet, whilst the general nature of the typologies above provides an opportunity to apply their categories easily to different tracks, they are only loosely defined.

To illustrate some of these problems, consider Ed Sheeran’s “Kiss Me”. On cursory listening, the track offers an example of what might be thought of as a documentary recording. It begins with Sheeran counting in his ensemble, affording us the position of a listener in an actual performance. The instruments are spread throughout the soundbox as we might expect from a group of performers and occupies a fairly coherent acoustic space. To reinforce our point of listening within the band’s performance space, a faint click is even audible, as though it spills from the singer’s headphones during the performance and included in the recording for the sake of reception-oriented authenticity.

For the most part, then, this section of the track appears to align itself with a position which affords, perhaps, the impression of being in the room as Sheeran sings to us. This scenario is supported by Sheeran’s count-in and the apparent spill from headphones. However, equalisation and balancing are evident, particularly on the voice, which is dynamically compressed and equalised to add a ‘brighter’ high-end, and the knocks on the guitar body, which have been amplified but equalised to muffle the higher end and approximate a bass drum. Later, the track illustrates greater departure due to vocal multitracking, which draws attention to the intervening role
of the recording process. Closer attention provides further clues as to the constructed nature of the track. The click, for instance, would not ordinarily be captured spilling onto a track in a contemporary commercial popular music recording and its placement within the stereo field slightly right of centre suggests that the sound is spilling from Ed’s left ear cup on to one of the microphones during the recording process. If it had indeed spilled onto the vocal track, though, we should hear this panned centrally with the vocals. Instead, the spatialisation here demonstrates that the producers are engaged with the creation of a virtual phonographic space. The beginning of the track when Sheeran counts in — a device which might misleadingly suggest a low degree of mediation by record producers — invites further doubt.

“Kiss Me”, then, may appear at first somewhat realistic, but on attentive listening, offers several points of departure. Nevertheless, it employs mimetic strategies that afford building greater verisimilitude into a narrative world. With this in mind, I would like to consider recordings not in terms of discrete orders of realism, but in terms of the mimetic strategies they employ. We will return later, though, to Moore’s blended realism, which I argue in Section 3.4.5 represents a special case and can be helpfully applied to instances in which multiple acoustical spaces or landscapes are represented simultaneously.

3.4 Four Aspects of Verisimilitude

Having considered several representative approaches to musical verisimilitude in the previous section, this section synthesises the ideas so far in this chapter in relation to four parameters of a recording. In each of the following subsections, we will discuss mimetic strategies that afford the building of verisimilitude into a narrative world, and common points of departure. I will begin with issues relating to the persona, in which we encounter the two types of mimetic strategies noted in Section 3.1, i.e.
strategies which either prompt the incorporation of external schemata or which suggest an internal consistency within the narrative world. Thereafter, we will address phonographic aspects in relation to landscape, staging/acoustical space, and source-bonding/gestural cause. Generally speaking, mimetic strategies in these areas involve the incorporation of external schema.

### 3.4.1 Persona and Narrative Logic

In relation to recorded popular song, we can often see a concern for verisimilitude through the capital of ‘authenticity’, in the sense of songs which are presented as ‘real’ or ‘genuine’ by depicting relatable scenarios which, for all we know, could happen or even have. Indeed, Moore highlights both “realistic persona” and “everyday situation” amongst a list of normative lyrical content which he calls “the ‘bedrock’ position of the persona” (2012b: 183). As he later reports:

> a dominant aesthetic position is to be able to claim that the expression embodied in the performance, in the recorded track, is an integral part of the identity, as a real-life being, of the singer whose words may form the starting-point for the listener’s construction of the persona (2013: 9).

Here, then, we see one of the most prevalent ways in which songs may refer to familiar aspects of a reference worldview. I would like to focus on an understanding of authenticity as the impression of unmediated expression; that the persona is relaying the experiences of a real singer.

The form of authenticity described above introduces the first mimetic strategy relating to the persona: the correspondence between the persona and the real singer. Writing in 1993, Ted Friedman draws attention to the importance of this in relation to authorship. In his brief discussion of Milli Vanilli, a German R&B duo who in 1990
were awarded a Grammy award which was subsequently revoked on the grounds that
the pair did not sing themselves on their releases, Friedman suggests: “while most
viewers recognize the complex division of labor in moviemaking — nobody gets
upset that actors don’t do their own stunts — pop music hangs on to the folk-era
image of the individual artist communicating directly to her or his listeners” (1993).
For Milli Vanilli, the lack of correspondence between the personae in their music and
the real individuals themselves thus led to a distrust in their music and artistry.

For other artists, the correspondence between the real singer and persona is a
crucial aesthetic aspect of their musical practice. We might contrast the case of Milli
Vanilli with that of Adele, who is praised for her apparent sincerity: “she provided, in
2011, a measure of authenticity in the face of her peers, a uniquely identifiable personality
in a din of so much indistinguishable claptrap. That seems to be the consensus
between fans and critics alike” (Messitte 2015, my emphasis). Similarly, for Paul
Williams, “[p]eople see her as a real talent and artist [...] At a time when, too often,
very manufactured artists are being pushed at the public, it feels like a breath of fresh
air” (qtd. in BBC 2011, my emphasis). The authenticity ascribed to Adele here can, I
believe, be attributed to two principal aspects. The first is that her songs are generally
understood to be autobiographical. Indeed, Messitte highlights that Adele’s music
represents something of her personality, whilst Williams emphasises her personal
musical ability and apparent distinction from commercially manufactured artists or
groups. Both accounts suggest that a blurring of persona and real singer implies that
the real singer is relaying his/her own unmediated (that is, actual) experiences to the
listener. In this case, it can be expected that a listener could build verisimilitude into
a narrative world because we are purportedly hearing the actual experiences of the
real singer.
The second characteristic which I would like to highlight, in which Adele again represents a common position, is the attempt to convincingly portray the protagonist/narrator. The voice plays an essential role in suggesting correspondence between the persona and an experiencing individual in the narrative world: Grossberg, for instance, emphasises “[t]he authenticity of rock was measured by its sound and, most commonly, its voice” (1992: 207-208). This involves aspects such as emotive timbre, the pace of the delivery, and the physical effort identifiable in the voice as a consequence of the cavity from which the vocalist sings and the contour/register of the music which they sing.\footnote{He later supports the role of rock’s sound as a marker of authenticity by making the important point that the music was often received through radio or phonographic recordings. Similar concern for the sound of the voice is taken up earlier by Barthes (1990 [1977]), who explores what he calls the ‘grain’ of the voice in an attempt to prompt discussion around the paralinguistic aspects of the voice. This also informs Lacasse’s writing on paralinguistic aspects of the ‘phonographic voice’, in which he emphasises “the aesthetics of recording popular music is first and foremost anchored in performance; more precisely, in individualised performances of feelings and emotions, as expressed by [...] a voice” (2010: 226).}

Whilst Grossberg is concerned with authenticity from an ideological perspective in relation to rock music, this point can also be applied to recorded popular song more generally.\footnote{Indeed, regardless of style, it is often possible to recognise emotional characteristics portrayed by the persona, even in cases of foreign-language songs in which the original meaning cannot be understood.} Here, this memetic strategy affords the incorporation of verisimilitude into a narrative world because the emotional labour of the voice supports a sense of internal consistency, as though hearing the expression of the protagonist/narrator themselves.
The presence of a persona allows us to also consider the plausibility of the situation relayed by the protagonist/narrator. As I have noted previously, popular music often portrays realistic situations which will be familiar or at least accessible to a general audience. An example is offered by the Pet Shop Boys’ ‘Integral,’ which some could say appears clearly fictional but nevertheless possible. The duo’s management suggest that the song is a reaction to a proposal discussed by the UK government to introduce identification cards to increase citizen surveillance (BBC 2006). The song, therefore, begins from actual circumstances from which it departs to afford imagining an exclusive, dystopian world in which citizens are under constant scrutiny. To an extent, then, the song affords the building of verisimilitude into a narrative world due to the link with actual events. Listeners might supplement their vision of this narrative world with the logic of contemporary surveillance culture and invest greater belief as the song extrapolates a possible future from an actual situation. Yet, the coherence of the persona and the quiescent personic environment in affording the envisaging of a menacing futuristic society in the spirit of dystopian fiction such as Orwell’s Nineteen Eighty-Four also acts as an invitation to the listener to draw schemata from our knowledges, experience, and familiarity with other works external to song.

The final mimetic strategy relating to the persona which I would like to propose is the construction of a consistent logic which governs the narrative world. As Segal puts it, “[e]very story world has a set of implicit constraints, rules, or principals which dictate what can or cannot be the case in that world. These constraints serve as a storyworld logic [...] Knowledge of the storyworld logic constrains the interpretation and experience of the story” (1995: 72, italics in original). In the example of “Integral”, the persona acts as spokesman for the dystopian regime, which he celebrates whilst affirming his authority over the addressee. Various aspects
of the persona’s account contribute a coherent narrative logic: it begins by praising
the regime and its supporters, argues against privacy in the chorus with an argument
that the innocent have nothing to be afraid of and celebrates the numbering of all
citizens in the second verse. Alongside the threatening persona, the musical discourse
incorporates various quantised synthesised percussion elements which lend a rigid
metronomic feel and sense of mechanisation to the piece, and the track is mixed with
a strong bass register lending a dark quality to the mix. Rounds of applause are also
added at 2′19″ and 3′21″, suggesting popular support for the dictatorial character
enacted by the persona. Smaller details, though, also support the coherence of the
architecture of this narrative world. One particular detail relates to the central
character’s animosity for the addressee, which is described by the lyrics but
supported also in a tensing of the singer’s jaw at the end of the words “protect you”.
In both prechorusses, the narrator highlights the homogeneity of the possible world,
reflected musically by the bass dwelling on the tonic. Indeed, the voice here is
processed using a vocoder to dehumanise the vocal timbre, using a technology
originally developed for encoding and relaying speech across larger distances at times
of war.

3.4.2 Landscape and Source-Bonding

Whether through sound synthesis or sound recording and processing, recording
technology has enabled the creation and combination of sounds which could not be
otherwise possible. This again will have important consequences for the degree of
mimesis which a track may evoke for the listener: with sounds that resemble (or are
clearly derived from) actual sources, we might expect a degree of mimesis, for they
afford the imagining of known or possible objects. At greater degrees of abstraction,
however, sound-events may suggest a greater extent of departure from the reference
worldview.
In Section 2.4.1, I introduced Denis Smalley’s (1997) concept of source-bonding. This forms the basis of his four-fold taxonomy of *surrogacy*, the separation of gesture from a clear agent or source. Within his system, greater degrees of surrogacy indicate greater abstraction from acoustical sources and, hence, greater departure from a reference worldview. First-order surrogacy describes a gesture that is projected directly into sound without apparent concern for musicality. Thereafter, second-order surrogacy refers to sonic materials which have been transformed as an instrumental musical purpose, which may include both actual instruments and electronic imitations. Third-order surrogacy describes an inferred or imagined gesture because the source or cause is uncertain or unrealistic. Lastly, remote surrogacy describes a gesture of unknown source and cause.

We have already seen the use of source-bonding as a mimetic strategy in “Halloween”, as the track incorporates several first-order surrogates, such as the howling wolves and cracks of lightening. Philip Tagg refers to this category of sound using the term *paramusical*, which he defines as “literally ‘alongside’ the music, i.e. semiotically related to a particular musical discourse without being structurally intrinsic to that discourse” (2013: 596). Commonly, when such paramusical textures are used in recorded popular music, they provide a referential, mimetic character that conventional musical sound-sources could not accomplish in isolation.

---

81 Source-recognition in relation to the perceived verisimilitude of a phonographic environment is also discussed by Ten Hoopen (1994), who explores the contribution of perceived causes of sounds to the characterisation of landscapes. She argues that our cognition of sounds evolved to quickly identify both the source and cause of sounds to respond to potential threats or capture prey. Accordingly, she proposes that the relationships between source-recognition and the apparent causes of sounds inform our perceptions of whether or not sounds originate from human action and actual objects.
Paramusical sounds could fulfil two important functions in relation to verisimilitude. Firstly, they prompt us to incorporate the imagined sources into a narrative world and so promote a greater likeness between the narrative world and the reference worldview familiar to us. A second related factor can be derived from paramusical sounds which do not appear to be used in a narrative capacity per se, but rather originate — or appear to originate — from details of the performance or place in which a recording took place. These details might include, for instance, a singer’s breathing, or guitarist’s fret noise, or spill from a metronome during a recording. This second form of paramusical sound authenticates the performance or place and affords a sense of verisimilitude through the correspondence to an actual performance.

Moore’s discussion of phonographic realism finds performance-related paramusical sounds particularly important in the identification of documentary realism. He refers to the audible tick of a mantle clock in a recording of Pop Maynard singing “William Taylor” (2015: 166) as an indication of documentary realism, for it suggests the recording functioned to document a performance and that the performance situation was not mediated by aesthetic demands. This could also be compared to Stan Link’s (2001) argument that the presence of noise authenticates a recording because it makes the mediated nature of recorded sound explicit.

Having considered the source-bonding of individual sounds, we should also address the perceived verisimilitude of their combination. As Wishart (1986) and Young (1996) note, the use of sampling makes possible novel landscapes which contain sonic materials bonded with sources which would not ordinarily coexist. In other situations, sound-sources themselves may appear abstract if they are not bonded with recognisable sources. The surreal phonographic environments to which Young refers are also commonly effected by multitracking, in which a singer might,
for instance, provide his/her own backing vocals which are played simultaneously. This evokes a landscape which is entirely impossible without the support of recording technology and so suggests a point of departure from the reference worldview.

3.4.3 Correspondence to a Live Performance

Having so far considered mimetic strategies relating to the persona and to the sound-sources heard in a recording, I would like to consider the role of the musical environment in greater detail in relation to verisimilitude. In Sarah Thornton’s discussion of club cultures, she notes that:

[w]hile authenticity is attributed to many different sounds, between the mid-fifties and mid-eighties, its main site was the live gig. In this period, ‘liveness’ dominated notions of authenticity. The essence or truth of music was located in its performance by musicians in front of an audience (2001: 49).

Similarly, Zagorski-Thomas (2010) discusses ‘perceived performance authenticity,’ which may be ascribed to recordings which highlight the performative labour of human performers. In other words, both authors suggest that a believable performance — that is, a performance which appears ‘true to life’ — authenticates the performers. With this connection between verisimilitude and authenticity in mind, we can observe mimetic strategies derived from the discernibility of human agency in the rendering of a performance we hear in a track. We have seen this already, for instance, in the inclusion of a count-in to “Kiss Me”.

I shall use the term performative agency to refer to the extent to which a recording appears to document the action of human performers. Whilst this was initially a prerequisite for recorded popular music, performances implied in
recordings today may represent a collage of different takes and material which is processed or triggered using electronics. Using sequencing technology, it is also possible to automate performances that are not limited by restrictions of, for instance, changes in registration, rhythmic complexity, or duration. In some musical practices, though, performative agency may be reflected through the use of production technology rather than through conventional aspects of performance using musical instruments. To this end, Bennett has explored recordists’ ‘gesture,’ meaning “the intentional and sonically discernible decisions made by the recordist, for example, spatial, volume, or frequency movements occurring during the course of the recording” (2015: 470).82

Joseph Auner (2003) describes departure from performative agency in relation to the voice as ‘posthuman ventriloquism.’ Auner situates this in relation to the work of N. Katherine Hayles, for whom the posthuman is an entity concerned with information and cybernetics unrestricted by the ordinary limits of embodied humans. She reports: “in the posthuman, there are no essential differences or absolute demarcations between bodily existence and computer simulation, cybernetic mechanism and biological organism, robot teleology and human goals” (1999: 3). In the case of posthuman ventriloquism, Auner develops Hayles’ understanding of the posthuman to consider the electronic manipulation or emulation of the human voice through technologies such as speech synthesisers, 

82 Zagorski-Thomas (2010) also makes the point that performance-related authenticity is ascribed to tracks in relation to their cultural and historical context. Hence, mediation and the use of the recording studio to conceive tracks may afford the ascription of authenticity: he argues “[o]ne of the ways in which extensive technological mediation can gain acceptance with an audience [...] is for that mediation to be perceived as part of the artists’ communal or individual creative practice” (ibid.: 261).
vocoders, and autotuning. Although Auner’s attention is directed towards singers and their electronic counterparts, the introduction of drum machines, sequencers, samplers, and beat-editing software has also allowed an analogous possible post-humanisation of any possible agents in popular song performance.

To reflect on the possible forms of performative agency in a recording, we might begin from Smalley’s (1997) second-order surrogate, which describes sounds that are performed in some musical context. This is an analogue to a common position in acoustic music styles in which a performance is effected entirely by a performer with no further identifiable mediation by audio processing or editing. From here, Auner’s posthuman ventriloquism signals a position in which technology is used to conceive performances which begin from an actual performance but could not be reproduced by human performers without the support of audio processing. This might include the use of sampling to achieve otherwise-unachievable drum patterns, the transposition of an instrument beyond its ordinary range, or the use of autotuning and other sound effects which electronically manipulate the acoustic character of a gesture. Finally, performances which are conceived entirely in the recording studio may appear to be programmed and triggered solely through electronic means. This could include, for instance, the quantisation and infinite repetition of drum machines and sequencers. As several scholars have noted, including Thompson (above) and

---

83 Similar forms of mediation form the basis of Loza’s (2001) discussion of the manner in which production technology transforms the image of the female body in electronic dance music as a form of ‘sampling sexuality’.

84 Of course, the divisions between these three forms of instrumental agents may be unclear or subject to change. As Goodwin (1990: 264) discusses in the case of Chris Lowe’s sampled tambourine programming in the music of the Pet Shop Boys, performances conceived using sequencers could instead be programmed in such a way as to instead suggest a posthuman
Frith (1986), greater degrees of technological mediation in the conception of a musical performance is often greeted with suspicion or accused of being ‘unnatural.’ From this, we might also pose a complementary argument that clearly discernible performative agency acts as a mimetic strategy which allows us to think of the recording, and accordingly, the narrative world, as truer to life.

3.4.4 Acoustical Space and Staging

The final aspect of recorded popular song that I will discuss in which we can observe the use of mimetic strategies — or, more often, points of departure — lies in the mixing process. Recordings necessarily involve the creation of a virtual acoustical space through the staging of sound-sources, particularly through the application of reverberation or delay, balancing or other dynamic processing, equalisation, and panning. Although these processes make it possible to construct soundboxes modelled after actual spaces, we often encounter greater departure in recorded popular song as a result of schematised phonographic acoustics or the reconfiguration of staging.

Zagorski-Thomas (2014) emphasises that processes such as level balancing, compression, equalisation, and reverb processing are often applied during mixing to increase intelligibility and audibility, resulting in schematic sonic cartoons (see Section 3.2). This may indicate departure from acoustical norms if the dynamic balance heard in the recording does not reflect the balance we would ordinarily expect. Take the voice as an example: it is straightforward to recognise the difference between a singer whispering or singing quietly, and shouting or singing powerfully. However, by close-miking the vocalist and adjusting the voice’s amplitude in the mix, or even realist performance by programming variation in amplitude or timing (a process known as ‘humanising’).
we can easily decouple timbre from amplitude, which afforded the possibility of ‘crooning’. We can observe here, then, a common point of departure which arises from dynamic balancing.

Zagorski-Thomas (2014) argues that schematic phonographic representations may be as or even more effective than documentary recordings. In his earlier writing, he emphasises that it is unwise for audio engineers to use an entirely realistic reverb response because reverberating low frequencies are likely to cause ‘muddiness’ in the mix (Zagorski-Thomas 2010). Instead, it is common practice to decrease the ratio between reverberation and dry signal for instruments with prominent low frequencies, particularly kick drum and bass. Regarding source-bonding, Wishart (1996: 138) exemplifies a similar case by arguing that to convincingly portray a bonfire without visual stimuli requires multitracking to create an inaccurate representation which nevertheless appears more true than an actual recording.

Brøvig-Hanssen and Danielsen (2013) provide a detailed discussion of space and realism in recorded popular music and emphasise the importance of known spaces as a reference for simulated acoustical spaces in recordings. They set out three categories of spatial configuration: single spatial setting, in which all instruments are placed within a coherent (though not necessarily realistic) acoustical space; multiple spatial settings, in which the dimensions of the soundbox reconfigure during the track; and, spatial simultaneity, in which multiple simultaneous acoustical spaces can be identified within the phonographic environment. The pair’s identification of simultaneous representations of multiple spaces is pertinent to Moore’s (2015) discussion of Kathryn Roberts’ recording of “The Plains of Waterloo”, which notes that the movement of the piano within the soundbox would require significant work

---

85 The inclusion of vinyl crackle, for instance, may connote age without requiring further dynamic or spectral treatment as we would expect of actual playback of a record.
by stagehands (which of course are not heard in the recording) without the support of the recording studio, and that the piano also appears to occupy a different resonant space to the vocal line. These details, Moore suggests, both indicate a point of departure.

### 3.4.5 Blended Spaces

Instances of Brøvig-Hanssen and Danielsen’s spatial simultaneity can be found in various recorded songs. In the middle-8 of Taylor Swift’s “We Are Never Ever Getting Back Together”, for instance, Taylor stops singing and we seemingly hear her on the phone to a confidant(e). In this case passage, we can observe two particular mimetic strategies are used: firstly, Taylor begins speaking rather than singing — that is, the voice becomes a first-order surrogate, where previously it was a second-order surrogate; secondly, the voice is staged within an implied acoustical space of a small tiled room. Hence, we might take this call to be occurring directly within the narrative world.

The musical environment, conversely, evidences several points of departure: the guitar is looped and reverses during the last beat of each bar, suggesting departure in relation to performative agency; Taylor is multitracked during the refrain, suggesting departure in relation to landscape; and, some balancing between instruments is evident, suggesting some departure in relation to acoustical space. As the middle-8 begins, the musical materials also abruptly drop in amplitude as the higher frequencies are aggressively filtered. Over the eight-bar section, the cutoff frequency of this filter, above which higher frequencies are attenuated, gradually rises to reveal the muffled elements, drawing greater attention to aesthetic role of the recording studio. This moment of clear departure concerning the musical environment contrasts the apparent verisimilitude of the telephone call. After
Fauconnier and Turner (2003), and following Moore (2015), we can describe this moment of spacial simultaneity as a *blended space*.

The term *blended space* is derived from cognitive linguistics in which Fauconnier and Turner (2003) propose a theory of ‘conceptual blending’ to describe the process of combining two or more concepts, or mental ‘spaces’ to create a new space that facilitates sense-making in a given circumstance. This theory, the authors suggest, forms the basis of human learning and experience. Under Fauconnier and Turner’s theory, a network is formed when two or more ‘input spaces,’ which share some form of general principals (the ‘generic space’), are combined to form the ‘blended space.’ Fauconnier and Turner explain that blended spaces are subject to frequent processes of compression and decompression, which describe two means of negotiating input spaces. Generally speaking, the act of compression involves some coming together of temporally or spatially desperate input frames. Decompression, on the other hand, refers to the separation of these two spaces. In his introduction to the application of blending theory to narratology, Ralf Schneider (2012) provides a helpful analogy of two projectors, which each show different images. From the two input sources (the original images), a third image is formed through a process analogous to compression - the blended space.\(^8\)

\(^8\) Schneider notes that this analogy is limited by the fact that overlapping projected images are likely to be disorderly and difficult or unhelpful to decipher, whereas conceptual blends are generally more coherent and their input spaces share some common abstract characteristics: a mental blend “is meaningful precisely in its very newness. It is both more than the sum of its inputs and different from them. Blending theory assumes that *cross-space mappings* between input spaces are enabled by analogies they share on a general, abstract level” (*ibid.*: 3, italics in original).
Zagorski-Thomas (2014: 80-81) provides examples of blended spaces in terms of tracks which combine sounds bonded to different acoustical spaces. Through conceptual blending, he suggests that interpreters will often interpret this as if it were a coherent performance and make sense of the phonographic environment as a single space which we might imagine to be inhabited by the implied performers. With this in mind, we may wish to reconsider Moore’s choice of terminology in relation to ‘blended realism’ as a position of greater departure than ‘enhanced realism,’ as each of his categories (with the exception of documentary) could theoretically incorporate some degree of blending in terms of production. Accordingly, I use the concept of blended spaces in relation to recorded music in the specific circumstances that the phonographic environment represents multiple acoustical spaces and/or landscapes simultaneously.

Understanding blended spaces in terms of acoustical space or landscapes is particularly relevant to tracks such as “We Are Never Ever Getting Back Together”, as well as tracks in which two distinct configurations of the phonographic environment appear to overlap. My Chemical Romance’s “Welcome to the Black Parade” exemplifies the latter and offers an intriguing example of shifts between phonographic verisimilitude in successive sections. In an opening sequence, the listener is presented with three principal configurations of the soundbox. The first consists of a piano and solo vocalist, accompanied faintly in the background by a ticking metronome. The presence of the paramusical metronome acts as a mimetic strategy relating to landscape, and indeed, the three sounding objects are placed in a spatial arrangement we would expect from a vocalist sitting at a piano. The attentive listener might also note the creak of the piano stool and sustain pedal which support

---

87 This suggestion that ‘overlapping’ acoustical spaces will be interpreted as a composite is also made by Walther-Hansen (2015).
the verisimilitude of the landscape and performative agency. Shortly after, a marching band enters, following the same melodic contour of the piano, though this plainly could not occur within the same space and one could not expect the vocalist to sing louder than an entire marching band. This hence depicts a blend of two spaces. The third configuration of the soundbox, marked by a drum fill functioning as an episodic marker at $t'47''$ demonstrates greater departure still (the relative amplitudes and frequency character of the instruments have been equalised to aid intelligibility).

3.5 Case Study: Madonna - “Hollywood”

To close this portion of our discussion on verisimilitude and departure, I would like to consider Madonna’s “Hollywood” from her 2003 album American Life to provide an extended example of several aspects relating to our preceding discussion. The release was Madonna’s ninth studio album and was produced by Mirwais Ahmadzaï, with whom she had worked on her previous album Music.

In a brief survey of literature relating to the film industry, Andrew Ali Ibbi (2014) contends that Hollywood signifies an ideology which some take as a quintessential representation of American consumer culture. This makes it a particularly appropriate topic for the American Life album, which is presented as a critique of the ‘American dream’, addressing topics such as war, materialism, and celebrity. Writing in NME, Davis notes some difficulty negotiating the sincerity behind these themes:

88 Indeed, Toby Miller (2005) has coined the phrase ‘global Hollywood’ to emphasise the pronounced international influence of the district.
it's impossible to reconcile Madonna: The Altruistic Philosopher with Madonna: The Bloody Big Superstar - her attempts at English self-deprecation [sic.] (the rap about "yoga and pilates" on 'American Life', the question on 'Hollywood' - "how could it hurt you when it looked so good?") and self-analysis (another rap on 'Mother & Father' - "my father used to go to work/I used to think he was a jerk") just come across as gauche (2005).

In keeping with the album’s theme, the persona takes on the role of a protagonist explaining her ultimately problematic captivation with the city and the pressures placed upon her to conform to the unrealistic expectations of the local culture in “Hollywood”. In conjunction with this theme, the phonographic verisimilitude of the persona and environment changes at several points throughout the track, particularly in terms of vocal staging, landscape, and studio-assisted performative agency. Throughout the track, the persona expands on the culture of Hollywood without reference to specific detail, hence requiring the listener to imbue a schematic narrative world with external schemata relating to the setting of Hollywood and the pressures of the glamorous lifestyle it typically epitomises. Indeed, as a result of the city’s cultural presence, it is difficult not to encounter media coverage equating it with wealth and glamour. We might expect, therefore, a wide variety of listeners to have some preconceptions which they may use as the basis of worldmaking and indeed, the actual basis of the song’s setting encourages belief that the experiences reported to us could happen.

The track is structured in a series of short sections with somewhat economical use of lyrical material. It opens with the sound of ambient bird song before a guitar is soon added towards the right of the stereo field and positioned in close proximity to the listener. The guitar sets out a i-VII-VI-v pattern in B Aeolian over the space of two bars which continues throughout the majority of the track (the same progression is transposed into C# Aeolian from 3′00″). With the exception of the birdsong, which
fades over the course of around 27″, the guitar continues unaccompanied for four bars until the addition of a vocal track marks the beginning of the first verse. Over the eight bars which follow, the protagonist explains the beguiling charm of Hollywood to everyone seeking success there. From the persona’s unenthusiastic expression and assuredness, we could infer that the protagonist is relaying her experience from the perspective of an experienced inhabitant of Hollywood. The lyrical material of the first verse also returns twice later at 56″ and 2′12″.

During the first verse, the melody is provided by a single vocal track, which principally alternates between the ⁴ and ⁵ and is staged in the centre of the soundbox in close proximity to the listener. It is joined by a bass synthesiser which arpeggiates around a ⁴-⁵-⁶-⁴ pattern on a low-pass filtered timbre, slurred with notable portamento without a discernible attack. The bass follows the general contour of the guitar and employs a similar quaver-based rhythm. Meanwhile, electronic percussion provides an explicit beat backing, playing a tightly quantised pattern which repeats every bar. The percussive elements are staged dry, with a particularly short envelope, and bright character.

The first verse reaches a clear conclusion as the drums and guitar rest briefly at the end of the hypermetric group. We then enter the chorus (25″), which uses the same instrumentation, as well as the bass and guitar loops from the first verse. The drum pattern is also almost identical, other than the addition of an open hi-hat added on the first and third offbeats, and an additional kick on the second and fourth offbeats. Meanwhile, the filter raises on the bass synth, which reveals the more ‘biting’ higher harmonics. The chorus is, however, most clearly demarcated by the voice. The melody here is cast in the minor pentatonic and follows a much broader

89 Indeed, the total melodic range of this section is just a minor third, between ⁴ and the occasional ⁶.
melodic contour which pivots around the tonic. A chorus effect and small stereo spread are also added to the vocal track, which is now staged a slight distance from the listener within the soundbox. The chorus itself lasts a brief eight bars, in which the protagonist gives the first suggestion of the pressure upon her and her peers.

A second verse follows the chorus, once again marking a change in the phonographic staging of the voice, which is now multitracked and at its original depth in the soundbox, without the chorus effect. The guitar, bass (with the filter cutoff lower), and drum pattern of the chorus all continue as before, whilst the persona returns to the previous idyllic, aspirational representation of the city. For the first time in the piece, the repeating guitar pattern is resampled at the ends of bars 7 and 8, albeit subtly, to introduce short glitches.

Following a short rest, the vocal track is reintroduced, followed by the guitar. The voice here is a solo singer which repeats the lyrics from the first verse. The melodic contour, though, is far broader, spanning an octave with several large upward leaps from $\wedge 1-\wedge 5$ and $\wedge 5-\wedge 1$. We remain in the minor pentatonic, though occasional $\wedge 6$ and $\wedge 2$ degrees are also added for the first time. Unlike the first verse, the voice also appears less tense and is positioned an intimate distance from the listener with audible sibilance and breathing. Following this, the third verse returns to the phonographic staging of the first and reintroduces the drums alongside an additional guitar, panned hard-left. The chorus which begins at $1'25"$ is also mixed as before with a chorus effect and stereo spread applied to the voice.

After the second hearing of the chorus, the vocal manipulation becomes more explicit. The lyrical material of the first verse is repeated once again, with the bass, guitar, and drums this time, but is followed by a four-bar instrumental section in which the bass begins to alternate octaves each quaver as it descends step-wise from
^1 to ^5 each two bars. Here, an additional synth is added, as well as further distorted electronics on the left of the stereo field, perhaps giving a sense of the freedom and velocity suggested by the lyrics preceding it. The section closes with the vocal declaration of boredom with radio playlists and ethics, during which introduces a small stutter to the end of the word “song”.

A further chorus begins at 2′26″, extended by eight bars as the last “Hollywood” of the chorus is repeated transposed from the F♯ on which it began in earlier choruses to B3, and, finally, to C♯4. In these latter sections of the track, the vocal manipulation also becomes yet more transparent as formant shifting — a technique which transforms the timbre of a voice, for instance to approximate more feminine or youthful characteristics — is increasingly used. A short 6-bar bridge features at 2′58″ in which Madonna speaks with an intimate closeness to the listener as the environment moves abruptly from B Aeolian to C♯Aeolian in preparation of a further chorus. Unlike the previous choruses, however, the voice is not multitracked. A slightly granular quality can also be heard in the vocal line, which might suggest that the voice had been transposed artificially after the line was originally recorded. Indeed, the voice reaches the height of D♭6 during the final chorus, before a two-bar vocal segment is sampled and looped from 3′44″ as gradual formant manipulation apparently transforms the acoustic body from a young, petite, feminine vocalist, to approximating an older, larger body which distorts perceptions of the vocalist’s gender.

3.5.1 Acoustical Space and Staging

Throughout “Hollywood”, the staging of the voice and the acoustical space which it appears to occupy changes at several points. Particularly noticeable is the effect this has on the proximity of the voice to the listener. Moore’s (2012b: 185-188) adaptation
of Edward T. Hall’s investigations of the cultural meanings and significances of space and proximity between individuals holds that the apparent proximity of the voice to us in the soundbox may afford *intimate, personal, social, and public* forms of expression (at greater distance from the listener). A voice located some distance from the listener in the soundbox, for instance, would be best interpreted as making a public declaration, as opposed to a voice placed in close proximity which would ordinarily occur only in intimate forms of communication.

In “Hollywood”, the changing phonographic staging of the voice in different sections has impactful effects upon the apparent proximity between the persona and listener in the soundbox:

Table 1: Multitracking, Formant Shifting, and Proxemic Shifts in “Hollywood”

<table>
<thead>
<tr>
<th>Section</th>
<th>Intimate</th>
<th>Private</th>
<th>Personal</th>
<th>Public</th>
<th>Multitracked</th>
<th>Formant shifted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intro</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verse1</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chorus</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Verse2</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verse3</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chorus</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verse4</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bridge</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verse5</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chorus</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Extension</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Bridge</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chorus</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Extension</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outro</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The rapid changes in proximity and staging themselves reinforce the protagonist’s seeming concern for the mediated, superficial nature of Hollywood, though this is also supported by subtler details. At several moments, the voice is staged either intimately or privately due to the lack of reverberation. However, it is important to note that the vocal track during the verses and outro designated as private above is gated, that is, noise below a threshold is silenced. As a result, the ends of words, as well as the sounds of breathing and lip noise, are cut off, though we would certainly expect to hear these artefacts if an individual was communicating at close proximity to us.90 Hence, the voice has the characteristic of proximity whilst, in a way, detached and mediated. Furthermore, when the voice is staged in a public zone due to the addition of reverb and chorus which causes the voice to appear to recede in the soundbox, the timbre is most explicitly altered as if to suggest a distinction between the way in which the protagonist is represented in private and in public.

3.5.2 Landscape and Source-Bonding

In addition to the changing phonographic staging of the vocals in terms of proximity, the use of formant shifting deforms the image of the singer, whose voice is, for the majority of the track, processed to give the impression of youth by raising vocal formants, but is distorted as formants are shifted downwards which generally connotes older, larger, or male bodies. Similarly, multitracking invokes departure from a possible landscape.

90 A similar effect is used more extensively in Madonna’s earlier “Don’t Tell Me”. In Danielsen and Maasø’s (2009) analysis, they propose that this acts as a way of mediatising the digital format.
Noteworthy details should also be accounted for in the musical environment. One aspect which seemingly contributes to an overall sense of ‘gloss’ or ‘polish’ concerns the sonic design of the percussive elements, which are especially punchy and the abundance of higher frequencies contributes to their ‘bright’ character which could not be produced acoustically in the same way. Here, whilst the percussion functions texturally as we might expect from a drum kit, the sounds used are somewhat distant surrogates and again supports a general theme throughout the track of skin-deep beauty.

One further important aspect relating to landscape and source-bonding is the birdsong which opens the piece. Here, the clear source-bonding acts mimetically and affords accommodation within an environment which may at first feel familiar. In the context of a looping guitar, generally electronically produced instrumentation, and shifting vocal sources, however, the phonographic environment generally operates in a more metaphoric, and indeed, oneiric manner. As the first verse unfolds, the birdsong fades gradually, indicating a degree of mediation, but upon its return in the second bridge, a break in the recording is apparent at 3’02” that signals most overtly that even this naturalistic aspect of the landscape is mediated.

3.5.3 Performative Agency

Departure is also evident in the extent to which production technology informs the rendering of the performance in the track. As I have highlighted, the drums are quantised and looped. Similarly, the guitar is looped and the use of occasional glitch processing also evidences the role of recording technology in the conception of the track. However, we can identify performative agency in some other areas, such as the way in which the filter is modulated on the bass synth (especially clear in the second chorus and second hearing of verse 2). The voice in this case is also an example of
Auner’s post-human ventriloquism, as it reflects Hayles’ concept of transcending the limitations of Madonna’s own performance ability and body.

3.5.4 Persona

Finally, I would like to consider the persona and, more particularly, the lyrics and their articulation. Interestingly, the protagonist provides a flattering yet candid account of Hollywood's apparent charm and pressures, which subscribes to a view of the city that is well-documented in journalism. In the moment of surprising intimacy during the second bridge, the persona invokes the idiom of ‘flying the nest’ which typically describes adolescents’ leaving home and developing self-sufficiency. This detail helps form an image of a protagonist who has experienced and survived the pressures she describes, and supports a coherent logic in the formation of the protagonist in relation to the narrative world.

Despite seeming to cope in the face of adversity, though, the mediation of the persona makes clear the extent to which the protagonist bends to change her image. Indeed, this also extends to the way in which the ‘o’s and ‘d’ in “Hollywood” are understated and, during the chorus and first verse, this is sung in a particularly breathy way, perhaps to appear more appealing.91 Once again, this supplements the logic of the narrative world. Thus, whilst there is considerable departure in terms of the phonographic environment and performance, this supports the development of a narrative world which listeners may indeed find verisimilar. In other words the

91 In a study of vocal attractiveness, Yi Xu et al. have collected data which suggest that breathiness, moderately high pitch, and the projection of a small body size (as simulated in “Hollywood” using formant shifting) increases attractiveness to male listeners (2013).
coherent logic of the narrative world affords the perception of verisimilitude and negotiating the protagonist’s reported experiences as if they were, or could be, actual.

### 3.6 Conclusion

In this chapter, we have discussed several ways in which the persona and phonographic environment of recorded popular songs may or may not resemble truth, particularly by either corresponding to or departing from the expectations of a reference worldview. As reception analysis has indicated in literary and film media, the extent of perceived verisimilitude may affect listeners’ engagement with songs in addition to the way in which narrative worlds are constructed.

Beginning with two forms of mimesis derived from Halliwell’s writing — that is, mimesis as resembling another world, and as evoking a sense of internal consistency or possibility within a fictional world — I have suggested that the use of mimetic strategies affords the building of verisimilitude into a narrative world. In this sense, we can follow Thurman and Mattoon’s scale between forms of virtual reality which have the resemblance of truth, and those which suggest greater departure. As Culler’s argument suggests, the degree of mimesis in a text goes some way to informing the way in which one might ‘naturalise’ it, that is, negotiate a narrative as though possible. Critical in this regard is the transaction between the three worlds raised in Chapter 2 for, as several narratologists have proposed, the realisation of a narrative world requires the incorporation of external schemata to supplement the worldmaking cues of a text. Listeners will, in other words, construct a narrative world by drawing on both the phonographic environment and reference worldview, and will naturalise areas of departure.
Although mimetic strategies may prompt listeners to build verisimilitude into a narrative world, we must recognise the schematic nature of recorded popular song. This is particularly relevant here as the brevity of song forms typically introduce considerable spots of indeterminacy in relation to worldmaking information and event sequencing. Ryan’s rule of minimal departure holds that what I have described as the reference worldview forms the basis of narrative worldmaking, and that correspondence with this reference worldview will generally be assumed unless a text explicitly departs from it in some way. This is also enhanced by a common position in which songs are written in ways which appear realistic or relatable.

Whereas a common aesthetic position in popular songs is to employ realistic subject matter, Zagorski-Thomas’ concept of sonic cartoons holds that recordings generally depart from acoustical realism through the mediation of production. In tracks that are not only recorded but conceived through studio technology, we may encounter increasing departure in terms of acoustical verisimilitude in addition to the combinations of sound-sources and so on. This departure, which is particularly common in electronic styles of popular music, indicates an oneiric aesthetic at odds with the use of generally realist lyrical themes. This oneiric aesthetic is further supported by issues such as the mediation of the persona, fluidity of place and time, and aspects of plot, as we have seen in Aqua’s “Halloween”.

My approach here, which is to describe verisimilitude through the employment of mimetic strategies or points of departure, moves away from various existing work on the topic of verisimilitude (or ‘realism’) in recorded music. Whereas the approaches discussed in this chapter have attempted to describe recordings in terms of identifiable categories based on various criteria, we can identify problems in the nomenclature used, the loose definition of different categories, and the possibility that the recording may afford different degrees of mimesis to different listeners. In
my own system, I have described four parameters which can be investigated through analysis of the persona and phonographic environment: the verisimilitude of the persona and narrative logic in terms of the correspondence to a reference worldview and internal consistency; the agency of human performers identifiable in the track; the coherence of the landscape and clarity of bonding with an actual or possible source; and, the mediation of the soundbox. Given the various possible relationships here, I have not proposed a typology of verisimilitude, but rather I hold that these four aspects all offer mimetic strategies which afford either the building of verisimilitude into a narrative world — and, consequently, accommodation — or points of departure.

In the following chapter, we will explore temporal sequencing and diegetic framing, concepts which support several points of this discussion so far. In particular, temporal sequencing may reveal significant spots of indeterminacy in relation to the narrative which prompt narrativisation in order to construct a possible sequence of events. Both temporal sequencing and use of sound at different diegetic levels also introduce additional considerations in relation to the oneiric qualities of recorded music. As we will see in Section 4.4, the mimetic strategies and points of departure observed in this current chapter provide useful clues which help to localise sound-events in relation to the narrative world.
4. Diegetic Framing and Temporal Sequencing

In the previous chapter, I introduced several ways in which recorded popular songs may depart from the expectations of a reference worldview due to oneiric aspects relating to the persona, acoustical space, landscape, and performative agency. This current chapter engages with two other aspects that may contribute to a fluid or ‘dream-like’ narrativity which I discussed in relation to “Halloween”. Firstly, we will consider how different forms of time correspond with Genette’s narrative/story/narrating trio and how this affords ways of mediating time in narratives. Secondly, we will consider how mimetic strategies and points of departure may suggest different ontological relationships between sound-events and the narrative world.

As an opening illustration, consider the blended space during the middle-8 of our earlier example, “We Are Never Ever Getting Back Together” (see Section 3.4.5). The persona is staged in a way which resembles the acoustic of a small tiled room whilst the protagonist appears to phone a confidante. Meanwhile, the musical discourse appears to occupy an alternative acoustical space which is conspicuously filtered using a rising low-pass filter. Whereas the chorus before it is situational and does not imply a temporal progression, the phone call in the middle-8 locates an implied listener in the same room as the protagonist and offers a window into the narrative world as the story unfolds. We can think of this blended space in terms of two ontologically separate ‘frames’: one containing this phone call, and one which encapsulates the musical discourse. The seemingly realist frame which Taylor occupies portrays what we might imagine hearing if we were in the room with her as we move from a perpetual present to a progression of time in the narrative world as
the call takes place. Conversely, the close staging of the musical discourse and the use of a rising filter suggest that it could not inhabit the same space as Taylor, whilst the repeating, undulating harmony and use of audio reversing in the guitar suggests less of a temporal trajectory.

In this chapter, we will begin in Sections 4.1-4.2 by exploring temporality in narrative, using the concept of scenes to describe passages of temporal and spatial coherence which correspond to discrete story elements. Thereafter, from Section 4.3, we will consider in greater detail how we might address the apparent bifurcation of the phonographic environment in cases such as “We Are Never Ever Getting Back Together” by exploring the contributions of sound and music to the construction of scenes, and their ontological relationship with the narrative world.

4.1 Temporality

Temporality is fundamental to both music and narrative discourse: time is a predicate for events, and so, too, for the event-sequencing which Herman and others hold to be a core component of narrativity. In Section 1.1.1, I raised Abbate’s criticism of musical narrativity from her claim that music “traps the listener in present experience” (1991: 53) and responses which contest Abbate’s view that an unfolding present tense prohibits narrativity. I, too, would like to argue that whilst music, as a temporally mediated art form is necessarily heard in the unfolding present — and that live instrumental music lacks the semantics required for narrative tense in the same way as literature — neither precludes the possibility of narrative tenses in recorded song. Abbate’s comments might hold for the experience of music — the media — but not narrative.
As John Blacking writes, “[w]e may say that ordinary daily experience takes place in a world of actual time. The essential quality of music is its power to create another world of virtual time” (1974: 27, my emphasis). Frith later supports this argument and reports:

to grasp the ‘rhythm’ of a piece of music [...] is at once a physical and mental process [...] By entering this world of ‘virtual’ or ‘inner’ time we effectively (willingly, trustfully) leave the world of ‘real’ time: hence the common experience of music as timeless (the common use of music to achieve the state of timelessness) (2002: 153, italics in original).

Both Blacking and Frith address the experience of rhythm and form in instrumental music, but noticeably reflect Gerrig’s remarks raised earlier concerning transportation to narrative worlds and partial inaccessibility to the world of origin (see Section 2.5). With this in mind, I would like to follow Negus (2012a) in exploring the possibilities of another form of what Frith calls ‘virtual time’: time as it relates to the narrative world.

As Negus (2012a) highlights, the issues of time in relation to musical narrative remains partially theorised, and scholarship on this topic relating to recorded popular song is particularly limited. As he argues, music mediates temporality in a way which departs from the ‘objective’ experience of lived time and, indeed, affords accommodation within narrative worlds with different temporal structures.92 In Negus’ article, he draws from the writing of Paul Ricoeur, particularly concerning the way in which time mediates perceptions of the past, present, and future. He continues by arguing that music affords ways of relieving tensions between these three forms of

92 This link with accommodation is also implied by Richard Elliott, who similarly puts it that “[m]usic can be a form of escape from the realities of time” (2015: 14), which offers some comparison with Gerrig’s (1993) sentiment that accommodation restricts access to external phenomena.
time and experiencing a sense of ‘timelessness’. To do so requires the recognition of
a second form of time, here corresponding to the narrative world. In this section, I
will describe three temporal relations by addressing the temporalities of the three
forms of narrative that Genette identified (namely: discourse, story, and narrating).

4.1.1 Time and Discourse

The first form of time that I would like to address is its representation in terms of
discourse — the word ‘discourse’ here referring to the text which forms the basis of
narrativisation. Consider Zager and Evans’ “In the Year 2525 (Exordium &
Terminus)”, a particularly extreme example in which a sequence of narrative events
spanning 10,000 years is compressed into the track’s duration of 3’15”. Throughout the
track, the narrator predicts a sequence of dystopian futures, advancing by 1,010 years
between each verse. Through this representation of 10,000 narrative years, we can
separate the temporality of the discourse from that of the narrative world.

In her entry on narrative temporality, Fludernik (2005) highlights the
pioneering work of Günter Müller, a German narratologist writing in the mid-
twentieth century. She regards Müller as the first to distinguish in narratological
terms between the two forms of time exemplified by “In the Year 2525 (Exordium &
Terminus)”, by using the terms Erzählzeit (“telling time”) and erzählte Zeit (“told time”).
Erzählzeit refers to the communication of the story. Meanwhile, erzählte Zeit
describes the form of time experienced by agents within the story, i.e. time unfolding
within the narrative world. For straightforwardness, I will refer to these by the (more
common) anglophone terms discourse time and story time, respectively.

Whilst helpful, Müller’s early conceptualisation of discourse time principally
addresses the duration of a narrative text but does not address the narrator. In
literature, discourse time can be expressed using the spatial metaphor of pages used
to describe a temporal succession. In recorded popular song, this can similarly be thought of as the lines of lyrics used by the persona to describe an event or state of mind, or the minutes and seconds of the recording itself. However, returning to the example of “In the Year 2525 (Exordium & Terminus)”, we can see a temporal dislocation between the narrator and story time that Müller’s categories do not address. With the exception of the years 5,555 and 10,000, each verse of the song is cast in the future tense, suggesting that the narrator is predicting a hypothetical future. When we reach the year 10,000, the narrator reports “now it’s been 10,000 years”, which implies an arrival at the narrator’s present. As the track closes, it subverts temporal linearity entirely as the narrator reflects on man’s poor judgement and eventual new beginnings, returning to the first verse (the year 2525) as the track begins to fade.

Using this example, we can distinguish between three forms of time: firstly, the 3'15" of discourse time; secondly the millennia of story time within the narrative world; and, thirdly, the temporal location of the narrator (narrator time). Using these three categories in place of Müller’s two also avoids a confusion arising from Erzählzeit (discourse time). Chatman, for instance, describes discourse time as “the time it takes to peruse the discourse” (1978: 62), but shortly after appears to voice a different view when writing:

[i]f the narrative is overt, there are perforce two NOWs, that of the discourse, the moment occupied by the narrator in the present tense (‘I’m going to tell you the following story’), and that of the story, the moment that the action began to transpire, usually in the preterite (ibid.: 63).

Here, Chatman appears to reinterpret discourse time (or, specifically, the ‘now’ of the discourse) as it relates to narration, rather than in terms of the reader which his previous definition suggests. Chatman then continues “[i]f the narrator is totally
absent or covert, only the story-NOW emerges clearly” (*ibid.*), which again does not account for the interpreter’s time spent negotiating a discourse. We should clarify a distinction here between the two forms of time which can be read from Chatman’s account. On the one hand there is what I hereafter refer to as discourse time, meaning the duration of discourse used to narrate some moment or progression of story time. On the other hand there is narrator time, which refers to the temporal locus of the narrator, which I shall address in section 4.1.3.

### 4.1.2 Time and Story

The second form of time which I would like to discuss is *story time*, that is, time as it unfolds within the narrative world. Given its relationship to the narrative world, *story time* is also related to accommodation, i.e. the sense of relocation to a narrative world through immersion within a given narrative. If we are not accommodated within a narrative world, the perception of time must necessarily relate to discourse time and our experience of reading, listening, and so on. However, if we are accommodated within a narrative world, Gerrig’s account of transportation (1993) proposes that we adopt a different spatio-temporal perspective.

As it relates to the story level of a narrative, we often encounter oneiric behaviours connected to this form of temporality. Unlike discourse time, for instance, story time does not necessarily require progression, but might instead pause to offer a sense of a suspended present. In addition to the suspension of story time, we often also encounter temporal jumps in which passages of a song move between different temporal settings. Indeed, it is the crucial decoupling of story time from discourse time that affords oneiric event-sequencing, as with “In the Year 2525 (Exordium & Terminus)”, and non-chronological formulations of the fabula.
Whether discourse time refers in a particular case to the pages of a book, verses of a song, or seconds of a recording, the concretisation of time in the form of a medium affords various relationships between story time and discourse time. In “In the Year 2525 (Exordium & Terminus)”, for instance, the track describes a very compressed timeline. The relationship between the progressions of discourse time and story time is discussed by Genette (1980) under the term ‘duration.’ He goes on to categorise four types of disparity between these two forms of time: summary, in which narrative events are expressed in the space of a more rapid discourse time; pause, in which discourse time advances but story time does not; ellipsis, in which a significant portion of time in the narrative is omitted in the discourse; and, scene, in which there is a parity between the progression of discourse time and story time.

I would like to group Genette’s categories of duration into two sorts of story time which we may encounter in recorded popular song. The first, synchronic temporality, describes a stasis or ‘perpetual present’ typical of situational, rather than event-driven, writing; story time does not progress.91 The second form of story time, described as diachronic, involves some form of temporal progression and hence, event-sequencing. To exemplify this distinction using examples so far, “Hollywood” is synchronic, as there is no temporal development of the story, whereas “Stephen” and “Waterloo Sunset” include diachronic elements.

---

91 Both Moore (2012b: 183) and Tagg (2013) describe a normative position in popular song wherein the persona relays experience ‘in the moment’, exemplifying synchronic story time. Nicholls (2007) also notes that most popular songs are often situational rather than consisting of elaborate event-sequencing.
4.1.3 Narrator Time

The third form of time which I would like to discuss — narrator time — has been raised briefly above in my discussion of Chatman’s writing: the temporal relationship between the narrator and the events or situation which he/she narrates. To introduce some possibilities, I would like to employ a further example here: “In Your Arms” by Nico & Vinz. In the song, we join a protagonist at two stages in his life, firstly as a young man looking to his life ahead and, from verse three, as an older man looking back on the intervening years. A temporal setting is established during the first verse, in which the protagonist expresses his determined outlook for an imagined future. Verse 2, on the other hand is cast explicitly in the present as the protagonist reflects on the world at that point in (story-)time. This is also extended by a bridge to the refrain which follows. However, during this bridge and refrain, the tense becomes ambiguous as he sings “sometimes I feel like I can’t run…”, but that he nevertheless finds solace in the arms of his addressee.

Between verses 2 and 3, there is a conspicuous temporal schism due to the first appearance of narration in the past tense in verse 3, which also coincides with the change to a different singer. This new vocalist sings with a noticeably breathier voice and less powerful delivery, both of which are features that we might expect of an older individual. It is as though the hypothetical future which the protagonist described in the first verse has by now been and gone whilst this second persona voices the protagonist once older. Indeed, verse 4 begins by looking back in resignation on what has passed by this stage of his life, before returning to narration ‘in the moment’ as the protagonist reflects on what his life has become.

“In Your Arms” highlights two important areas of discussion. Firstly, as we have seen, the subject matter of each verse moves between different modes of narration as the protagonist addresses the present, past, and future. The second
important conclusion concerns the shift in singer, which appears to mark an ellipsis in story time corresponding to a different mode of narrator time. This apparent shift emphasises the distinction between discourse time, story time, and narrator time. With closer attention to the phonographic environment of “In Your Arms”, one could further argue that the ambiguous narrator time in each refrain is supported by a blended space. In each hearing of the refrain, both the opening singer, who takes the role of the younger protagonist, and the second singer, who seemingly takes the role of the protagonist some years later, sing together. This leads, then, to a temporal blend of the narrator at two points in (story) time, enhancing the impression of timelessness which arises from the refrain.

To describe these relationships between a narrator and events more systematically, I will adopt categories proposed by Genette (1980) that describe the temporal location of the narrator relative to the story time. He straightforwardly identifies subsequent, prior, and simultaneous, which are typically associated with the use of the past, future, and present tenses, respectively. Finally, in cases of temporal blends, we can refer to interpolated narration, which combines any two or more of the previous categories. “In Your Arms”, illustrates the use of each of these four relationships between the narrator and story time, summarised in Table 2.

Table 2: Temporal movements in ‘In Your Arms’

<table>
<thead>
<tr>
<th>Time</th>
<th>Lyrical Content</th>
<th>Narration</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intro 0”</td>
<td>- n/a -</td>
<td>- n/a -</td>
<td>- n/a -</td>
</tr>
<tr>
<td>Verse 1 9”</td>
<td>Description of the present self and aspirations for the future</td>
<td>Prior</td>
<td></td>
</tr>
<tr>
<td>Interlude 26”</td>
<td>- n/a -</td>
<td>- n/a -</td>
<td></td>
</tr>
<tr>
<td>Verse 2 35”</td>
<td>Description of social surroundings and oppressive views of others</td>
<td>Simultaneous</td>
<td></td>
</tr>
<tr>
<td>Bridge 52”</td>
<td>Description of the self</td>
<td>Interpolated</td>
<td>In the moment (ambiguous tense)</td>
</tr>
<tr>
<td></td>
<td>Time</td>
<td>Lyrical Content</td>
<td>Narration</td>
</tr>
<tr>
<td>----------------</td>
<td>-------</td>
<td>-------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Refrain</td>
<td>1’11”</td>
<td>Celebration of “when I lie down in your arms”</td>
<td></td>
</tr>
<tr>
<td>Verse 3</td>
<td>1’28”</td>
<td>Comparison of past with present</td>
<td>Subsequent</td>
</tr>
<tr>
<td>Interlude</td>
<td>1’45”</td>
<td>- n/a -</td>
<td></td>
</tr>
<tr>
<td>Verse 4</td>
<td>1’54”</td>
<td>Description of new and final living condition, contrasted to the past</td>
<td>Simultaneous</td>
</tr>
<tr>
<td>Refrain</td>
<td>2’11”</td>
<td>- as before -</td>
<td></td>
</tr>
<tr>
<td>Bridge</td>
<td>2’28”</td>
<td>- as before -</td>
<td>Interpolated</td>
</tr>
<tr>
<td>Refrain</td>
<td>2’47”</td>
<td>- as before -</td>
<td></td>
</tr>
<tr>
<td>Refrain</td>
<td>3’06”</td>
<td>- as before -</td>
<td></td>
</tr>
</tbody>
</table>

### 4.2 Song Scenes

As “In Your Arms” demonstrates, diachronic event-sequencing affords non-chronological representations of the fabula. The notion of linear chronological ordering in popular song is critiqued by Tagg (2013: 392-4), who argues for the consideration of diachronic forms in terms of a centripetal/recursive process. Tagg argues that, whereas narratives tend to be visualised in terms of a unidirectional progression, songs tend instead to involve cyclical recurrence due to repeats or reprisals.

In order to describe sequences of events in recorded popular song, we must divide songs into discrete blocks of story time. To this end, Tagg (ibid.) employs the term *episodes*, which helpfully connects music with a narrative paradigm, although

---

94 This is especially pertinent to verse-chorus structures, in which the interpolation of choruses often disrupts a linear flow of story time.
this word choice presupposes the presence of a temporal development which is inappropriate for synchronic writing. I would like to instead adopt the alternative term “scenes” to describe a passage of continuous story time in a recorded song, whether this contains a temporal development, in the case of diachronic scenes, or a single circumstance, in synchronic scenes. Tagg also proposes the episodic marker, which refers to a musical gesture that is used to demarcate successive episodes. In this sense, episodic markers denote some form of temporal or geographical schism. In “In Your Arms”, for instance, the sound of a mark tree acts as an episodic marker which coincides with the end of each section listed above and corresponding shifts in content and mode of narration.

Lukas Graham’s “7 Years” also offers an effective example regarding episodic markers and the sequencing of scenes. In this case, the track is presented as though a protagonist is watching a series of home movies, in which a film projector is used to delineate between scenes, capitalising on the connotations of temporal development and age associated with dated playback technology. The temporal trajectory of the story is linear until it briefly revisits the opening material at the end of the track. The song begins as the protagonist recounts his life, advancing through story time to overtake narrator time as the protagonist muses on a possible future (in which the sound of projector reels is absent). The song eschews a conventional verse-chorus structure in favour of lyrics which adopt a ‘stream of consciousness’, whilst the personic environment alternates between two similar forms of thematic material. Whilst the lyrics feature little repeating material, they can be divided into hypermetric groups of four bars that generally alternate between citing an age and a reflection on this stage of life. These are also differentiated in terms of the melodic contour which the persona performs.
The majority of these sections in “7 Years” can be thought of as distinct scenes. The song begins, for instance, as the narrator recalls his experiences as a seven-year-old. The persona is joined by chimes which provide some melodic embellishment in the form of a repeating gesture that descends by step from $\text{3}^\text{rd}$ to $\text{7}^\text{th}$ before alternating between $\text{5}^\text{th}$ and $\text{7}^\text{th}$ in the key of G Aeolian, in addition to the piano that provides some bass and harmonic support.\textsuperscript{95} This section is underpinned by the ticking of the film projector, which appears to reach the end of its reel as the following section begins, corresponding, it seems with a shift in spatio-temporal setting. This new section is characterised by a move from the protagonist speaking from his own perspective to speaking about how he and his friends encountered the world at the time. The musical environment here consists of a piano and, as the section segues into the protagonist’s 11-year-old self at 0′50″, strings are also added. In place of the projector, the environment is joined by the distant, reverberant sounds of children’s excited voices, perhaps from a playground. A distinction is thus made between two phonographic scenes: our indeterminate starting position in which the protagonist may be watching a home movie, and the world it depicts.

The following eight bars describe the protagonist’s experiences as an 11-year-old and can again be divided into 2 four-bar scenes. The first offers little in terms of detail, whilst the second describes how the protagonist began as a songwriter. Once again, they are separated by the subtle sound of a projector. At 1′06″, though, a string crescendo and drum entry coincide with the sound of the projector running out of film, as the protagonist recalls how he began songwriting. By 1′24″, he has moved on

\textsuperscript{95} The repeating idea heard on the chimes reoccurs later in several further passages as the protagonist introduces his 11-, 20-, and 30-year-old self, as well as the repeats of his 60-year-old passage (from 3′15″) and 7-year-old passage, suggesting a connection between this musical idea and temporal movement.
to his 20-year-old self, when he first found fame which is embellished at 1’35” with the sound of an emcee announcing the singer as a crowd begins to cheer.

Throughout the 20-year-old passage, the texture continues to thicken until at 2’11” the protagonist soberly looks towards turning thirty years old, which marks that story time has overtaken narrator time. Both the strings and drums are removed, leaving the piano, chimes, and vocals, returning to the combination heard during the first scene (the 7-year-old) and, once again, the sound of a projector fades in at the end of the hypermetric group. We soon learn that by this point in story time, the protagonist is a father and, as strings fade back in, the protagonist begins prior narration by looking towards turning 60 years old, as he shifts back to the first person singular. As the protagonist moves to prior narration, the sound of the projector — associated so far with looking back over old home movies — is unsurprisingly absent. Once again, the musical texture thickens and becomes more rhythmically active, leading to a dramatic climax in which the voice of the persona becomes audibly strained as if competing with the volume of the environment until it mostly dissipates and the protagonist repeats his first remarks about life as a 7-year-old from 3’31”. The voice here, though, is staged with more reverb than the first passage, perhaps giving a sense of the empty, isolated surroundings which he expressed concern over and the temporal distance here between narrator time and story time as though being stuck in the past. Once again, the projector fades in and the track itself closes as the projector can be heard spinning down.

“7 Years” can be divided into a collection of sequential scenes which chart a one-directional temporal trajectory from the protagonist’s recollected childhood, present day, and possible future, before returning, albeit briefly, to his childhood. Furthermore, it demonstrates the uses of phonographic and musical materials to distinguish between different settings.
4.3 Diegetic Functions of Music

Having discussed the temporal dimension of narrative worlds in relation to event sequencing, I would like to consider the related issue of the narrative world’s ontological structure. In “7 Years”, we saw that, whilst narrator time remained consistent, the persona described several different moments of story time, several of which were also illustrated by the sounds of the playground and concert emcee. With this illustrative role of the phonographic environment in “7 Years” in mind, consider our earlier example of “Welcome to the Black Parade” (see Section 3.4.5) and the blended space at 25" - 1'16" between the entry of the marching band and the disappearance of the original piano. In this case, the protagonist is once again temporally dislocated from the story-now because he is relaying an event that took place earlier in his lifetime, whilst the phonographic environment sonifies this recollected memory.

In both examples above, the dislocation of the protagonist from the story results in an ontological separation reflected in different relationships between sound-sources and the story. In other cases, such as the middle-8 of “We Are Never Ever Getting Back Together”, the mimetic staging of the voice suggests that its ontological status here has altered and that it — temporarily — emanates directly from the narrative world. In order to account for such relationships that are possible between the phonographic environment and story, the following sections will discuss the ontological architecture of narrative worlds.

Genette describes the base level of the story as the diegesis, a term which I will use hereafter to refer to the level of a narrative world which agents and objects inhabit
in the story-now. In many cases, songs employ simultaneous narration in which the protagonist/narrator may appear to occupy this diegetic level. However, as we have seen in “Welcome to the Black Parade” and “7 Years”, the use of subsequent or prior narration indicates a dislocation between the diegesis and the (non-diegetic) level of narration. This has since been developed in relation to film, in which Claudia Gorbman (1976; 1987) and others understand the world seen onscreen as a filmic diegesis and explore the ontological relationships of sound-events with this diegetic frame.

Genette (ibid.) proposes a scheme of three possible relationships between elements of a story and the diegesis. To refer to events that occur within the primary narrative, Genette adopts the terms diegetic or intradiegetic. Genette’s extradiegetic category refers to a position external to the diegesis, through subsequent or prior narration, or otherwise reporting from outside the world described. Finally, metadiegetic describes the construction of an additional narrative frame embedded within the diegetic level through a narrative agent becoming the narrator of an additional story that is subordinate to the first.

In order to adapt Genette’s approach to analyse narrative worlds of recorded music, I would like to begin from Gorbman’s typology of diegetic relationships of film sound. Gorbman (1976; 1987), too, describes the diegesis as a primary narrative: the world which can be seen on-screen. Accordingly, she poses three diegetic relationships of sounds by adapting Genette’s tripartite typology thus:

96 Diegesis here is used in the sense of “story” rather than a form of narrativity (cf. mimesis). Unfortunately, Genette’s original French terminology, diégésis (the narrative mode) and diégèse (referring to the levels of a story), are not differentiated in English translations.
• **Diegetic**: sounds which emanate from sources within the narrative and that can be heard by characters and the audience;

• **Meta-diegetic**: sounds which can be heard by one character and the audience, but do not emanate from a source within the diegesis and so are inaudible to other characters;

• **Non-diegetic**: sounds which do not emanate from diegetic sources and are only audible to the audience.

Following Gorbman, Bordwell and Thompson (1985; 2010) proposed a similar distinction to Gorbman's diegetic/meta-diegetic categories between two forms of what they consider to be diegetic sound: internal and external. As they write, “[e]xternal diegetic sound is that which we as spectators take to have a physical source in the scene. Internal diegetic sound is that which comes only from the mind of a character; it is subjective” (1985: 193). Of course, by offering a combination of seemingly ‘objective’ diegetic sounds alongside a sonification of a character’s subjective, internal psychology, soundtracks afford potentially complex ways of worldmaking.

### 4.3.1 Sonic Diegeses

To apply or adapt Gorbman’s categories to recorded popular song, we must conceptualise the diegesis in a way which is compatible with acousmatic media. To begin, I would like to take an example which illustrates some key issues at play. The Scissor Sisters’ “Let’s Have a Kiki” is a particularly interesting case in which we might speak of a sonic diegesis alongside other sound-events with different apparent relationships with the diegesis. The track begins with a scene in which we join an answerphone recording from a character named Pickles asking the caller to leave a
message, followed by an indicatory beep. In response, Ms. Matronic begins a lengthy monologue in which she bemoans being refused entry to a nightclub by the police whilst it rains. Beginning at 5”, ahead of a full entry at 7”, a musical layer enters, consisting of kick drum, tambourine, bongos, hihat, and synthesiser. The scene is also furnished with subtle police sirens. Matronic’s phone message ends abruptly on the open statement “I know exactly what we need”, before coinciding with the refrain “let’s have a kiki” in response.

In this first scene, the landscape and arrangement of sound-sources act as mimetic strategies which position us as though listening from the perspective of a spectator within Pickles’ apartment. Indeed, the phonographic staging of the scene situates it within a small to moderate interior acoustical space, whilst the voices of the characters are both filtered to sound somewhat tinny, distorted, and compressed, to create the impression of sounding from an answerphone speaker. To enforce this, both are also panned within the centre of the stereo field. Ms. Matronic is accompanied by the subtle noise of rain, again staged as though the noise is captured as she leaves the message. Together, these elements establish a sonic diegesis, as the vocals and other sounds appear to be bonded to a source (the answerphone) located within the story-now.

Whereas mimetic strategies suggest the establishment of a sonic diegesis centred around Pickles’ apartment, the percussion and synthesiser are clearly staged within a different space and show a greater degree of departure: the reverb characteristic is inconsistent; a delay effect is selectively applied to the bongos at the end of the passage; and, these elements are panned wider than could be achieved in the case of an actual reflective acoustical space. At first glance, the police sirens may appear to be diegetic, for they complement the setting of a late-night New York City apartment, and are staged quietly, towards the right of the soundbox, as though
passing by on the streets below. However, the sound is abruptly cut at several moments to metrically coincide with the musical discourse, such as at 14″ in which the siren is gated to accentuate the final kick of the hypermetric group. Hence, these sirens cannot be diegetic, but do not cohere with the musical instrumentation either, as they are paramusical.

So far, we can observe three areas of interest: sounds which appear to be mimetically bonded to the narrative world; musical elements which occupy a different acoustical space; and, paramusical elements that do not appear to occupy the diegesis. However, we encounter two important oneiric moments in which this architecture is reconfigured. In response to Ms. Matronic’s statement at the end of the call, the refrain begins at 1′08″, marked by the removal of paramusical sounds and the bypassing of the telephone effect applied to Ms. Matronic’s voice. Joining her, several multitracked male voices are introduced, panned either side of the soundbox in the acoustical space of the musical discourse. At this moment, the vocals appear to diffuse from diegetic to non-diegetic, firstly because they become closely integrated into the logic of the musical discourse and, secondly, consist of a collage of different short samples. A further moment of transition between diegetic relationships occurs after a passage of instrumental musical discourse that is joined at 2′24″ by the protagonist as she exclaims her enjoyment. Here, the sounds of the phonographic environment pause, as if in anticipation for the protagonist to continue. When she does so, several additional vocal layers are overlaid that can be narrativised as the party guests. The guests’ chatter reintroduces a diegetic level, again from the perspective of an implied listener within the apartment, whilst the non-diegetic music continues alongside it. Until this moment, the combination of sound-sources and their phonographic staging have suggested an ontological distinction between the non-diegetic musical elements and paramusical elements which, for the most part, appear diegetic.
However, this binary is disrupted at 2′45″ as the partygoers begin to sing along to the musical discourse whilst remaining staged within the diegesis.

In this example, we can observe aspects of the phonographic environment contributing to a base level of the narrative world which we can think of as a sonic diegesis. This offers useful worldmaking affordances because of mimetic representation, particularly in relation to specifying a particular narrative space. In this way, sonic diegeses also provide a means of differentiating between scenes within a diachronic structure and hence allow the interpreter to infer a sequence of events or temporally/spatially-distinct situations from a text. Elsewhere, Mads Walther-Hansen has also argued that recorded music may create a sonic diegesis through causal listening and the metaphor of the recording as a virtual performance. Yet, as he broaches issues relating to the oneiric character of recorded popular song which I noted in Chapter 3 (see Section 3.2), he observes that “the diegesis [in recorded music] is not always clearly defined in listening”, and that “sound events do not necessarily have the same ontological status throughout a track” (2015: 38). Before we continue, then, I would like to clarify my position and address how diegetic relationships may change throughout a track.

Indeed, Lacasse’s (2006) reading of “Stan” highlights the use of sounds which afford imagining particular settings that imply temporal relocation. In his reading, diegetic sound is used to differentiate between the settings of an interior location in which Stan is writing to Eminem and vice-versa in the final verses, and the setting of Stan’s car as he drives along the motorway.

Whilst he contextualises his discussion with the notion of ‘sonic narrative’, Walther-Hansen’s account chiefly discusses the ontology of recorded sound in relation to a musical performance embodied in a track and not explicitly address how this contributes to narrativity.
In the discussion which follows, I regard the diegesis to be the base level of a narrative world inhabited by narrative agents in the story-now. Accordingly, sonic diegeses should encompass elements in the phonographic environment which are staged coherently and function mimetically to afford the furnishing of a narrative world. Sonic diegeses may be envisioned in response to source-bonding, as we have encountered in “The Revenge”, or to particular forms of phonographic staging, such as the voice in the middle-8 of “We Are Never Ever Getting Back Together”. This understanding moves beyond a metaphor of the recording as a virtual music performance space, to instead interpret the sonic diegesis as a virtual narrative space. Accordingly, my discussion also presupposes a more associative listening mode in which sounds afford narrativisation and worldmaking. This is particularly pertinent to cases such as Queens of the Stone Age’s “You Think I Ain’t Worth a Dollar, but I Feel Like a Millionaire”, which Walther-Hansen uses as his opening example, that is situated within a narrative space of a car as a driver gets in, turns on the ignition, and starts the radio.

Moving to the second issue raised by “Let’s Have a Kiki”, namely the changing relationships between sonic elements and the diegesis, we can look to Rick Altman’s idea of audio dissolve. In his study on The American Film Musical, Altman explores the issue of diegetic-localisation from a position that is informed by Gorbman’s and Genette’s discussions of diegetic levels. Altman regards the diegesis straightforwardly as the centre of a narrative world but describes just two diegetic levels constructed by the soundtrack: what he calls the diegetic track, or the “bearer of realistic sounds”; and,

99 In shifting from Walter-Hansen’s metaphor of the virtual performance space, I also challenge the ontological position which he assigns to music. Here, I disagree with his understanding of the majority of a virtual musical performance to be diegetic if it instantiates a virtual acoustical space.
the *music track*, or “bearer of an instrumental accompaniment” (1989: 62). As he continues, he writes that the *diegetic track* “reflects reality (or at least supports cinema’s referential nature)”, while on the other, the *music track* performs poetically by “lift[ing] the image into a romantic realm far above this world of flesh and blood” (ibid.: 63). Hence, he associates the *diegetic track* with affording ways of worldmaking through mimetic strategies and the *music track* with a quasi-diegetic mode of narration reminiscent of Herman’s *what it's like*. Altman proposes the concept of *audio dissolve* to describe cases in which the *musical* and *diegetic tracks* diffuse into one another, “blu[rring] the borders between the real and the ideal” (ibid.: 63). This occurs, for instance, when a diegetic character sings within the diegesis but is joined by an imaginary orchestra.

### 4.4 Three Frames in Recorded Popular Song

With “Let's Have a Kiki” in mind, I would like to discuss how we might adapt the ideas raised by Genette and Gorbman and apply them to recorded popular song. Whilst it might suffice for Altman to distinguish between *diegetic* and *musical tracks* in film musicals, the distinction between what appears diegetic or not in recorded popular music is open to greater ambiguity, particularly in cases where the relationship between sound-events and the diegesis appear to change. The film projector acting as an episodic marker in “7 Years”, as one example, also problematises a strict application of Altman's binary, as it operates within a musical syntax but is paramusical and does not occur in the story-now. Below, therefore, I propose a tripartite approach after revisiting Walther-Hansen’s discussion.

Walther-Hansen distinguishes between diegetic frames in the recording by considering the sonic quality of sound-events and their perceived temporal relationships. For Walther-Hansen, “[m]eta-diegetic sound events are indexes of
events that took place before the diegetic events, while extra-diegetic sound events are indexes of events that take place after the diegesis” (2015: 35). This emphasis on temporal relationships indicates a closer connection with Genette’s model of levels of narration rather than Gorbman’s concern for sonic ontology: whilst he regards meta-diegetic sounds as a form of recall, Gorbman’s meta-diegetic category concerns — more broadly — imagined or hallucinated sounds that are not necessarily recalled. Similarly, sound-events which are not staged as part of the diegesis could also be considered non-diegetic without necessarily occurring temporally after the story-now, although this conflicts with Walter-Hansen’s position.¹⁰¹

The position which I adopt draws most from Gorbman’s approach, which does not require non-diegetic sounds to be temporally dislocated from the diegesis. Take here Walther-Hansen’s example of Pink Floyd’s “Wish You Were Here”, which

¹⁰⁰ He also makes a distinction between diegetic on-stage “[s]ound events in the perceived performance space” and diegetic off-stage “[s]ound events absent from the perceived performance space, but temporally present” (ibid.: 37). To clarify this distinction, consider the opening moments of Aqua’s ‘Halloween’ once again, which incorporate a telephone, dialogue, howling wolves, and thunder. Although the howls and thunder are not muffled as we might expect from an indoor perspective, they are staged some distance in the soundbox to suggest that they occur ‘off-stage,’ in contrast to the ringing phone and protagonist’s side of the call, which would be best described as diegetic on-stage sound-events in Walther-Hansen’s system. Whilst Walther-Hansen’s examples all incorporate on-stage and diegetic voices, I contend that we may encounter occasional exceptions to this norm. Regarding Kate Bush’s “Waking the Witch” (see Section 4.5), for instance, I propose a reading in which the staging of the voice suggests that the protagonist is not diegetic.

¹⁰¹ Bordwell and Thompson (1985) provide examples of internal sound which correspond with the past, present, or future of the world seen on-screen.
suggests that music filtered to adopt the characteristics of an AM radio is meta-diegetic. The presence of the radio, though, establishes a new narrative scene. In this scene, the radio and the sounds of its listener — who can be heard later breathing and playing along — both indicate a new story-now, and so, a new diegesis. As this interpretation places the radio within the story-now, it is not meta-diegetic, as Walther-Hansen hears it, but diegetic.\textsuperscript{102}

In the subsections below, I outline three possible forms of narrative frames within recorded popular song, in order of greater distance from the diegesis: 1) a \textit{diegetic paramusical frame} consisting of all elements bonded with sources within the story-now; 2) a \textit{musical frame} consisting of elements which operate within a musical syntax and do not appear to be bonded with sources within the story-now; and, 3) a \textit{non-diegetic frame} consisting of elements which are not bonded with the story-now, nor classified within the musical frame.

\textbf{4.4.1 Diegetic Paramusical Frame}

I described earlier (see Section 3.4.2) the use of paramusical sounds as a mimetic strategy, as they generally allow us to readily furnish a narrative world. In these cases where paramusical sound-events appear to be bonded with a source in the story-now, we can describe them as diegetic. The initial moments of the track are a common point in which we encounter what I refer to as the diegetic paramusical frame. Indeed, Herman makes a similar observation when he emphasises the importance of story openings on an interpreter’s initial construction of a storyworld and, he suggests,

\begin{footnotesize}
\textsuperscript{102} Part of Walther-Hansen’s argument here notes that the radio occupies only the right channel of audio and so is not staged to reflect placement within an actual space. Rather than disqualify the radio from the diegesis on this basis, I propose that this exemplifies Zagorski-Thomas’ sonic cartoon aesthetic.
\end{footnotesize}
accommodation. He writes that “story openings prompt interpreters to take up residence (more or less comfortably) in the world being evoked by a given narrative” (2009b: 79). In the following chapter, we will see that the inclusion of diegetic paramusical sounds to establish an ambience or space can be a powerful cue for worldmaking by localising the implied listener.

Thomas Dolby’s “Spice Train” offers a helpful example of a diegetic paramusical frame from the track's beginning, which features a boarding call accompanied by a train whistle. Both these paramusical elements serve to consolidate the setting of the narrative world aboard the spice train. Similarly, Nizlopi’s “JCB”, which wistfully recounts the protagonist’s youth riding a JCB with his father, begins straightforwardly with a JCB that appears to drive off and recede in the soundbox as the first verse begins. During the final seconds of the track following the final chorus, the sound of the JCB fades back in as we hear the engine running down as it turns off.

Whilst source-bonding may particularly effectively afford localising sound-events within the diegesis, other mimetic strategies may also play a role in localising sound-sources. In Britney Spears’ “Oops!...I Did It Again”, a passage at 2′11″-2′29″ between the second and third hearings of the chorus offers a particularly effective example. At this point, an episodic marker of a harp glissando — often used in film to accompany flashbacks or hallucinations — enters as the musical environment almost vanishes. In place of the former musical material, the track foregrounds paramusical sounds that are bonded with a train conductor’s boarding call, the whistle and puff of a steam train, and a spoken dialogue between what appear to be two lovers. Here, once again, the mimetic source-bonding allows us to localise these elements within the story-now of this scene, but Zagorski-Thomas (2009) argues further that this moment of audio dissolve is enhanced by the staging of this passage, which is equalised in a way that is reminiscent of dated cinema playback, to
complement an intertextual reference to the film *Titanic*. Here then, staging is used as a mimetic strategy which localises the story-now within a movie theatre.

In an alternative example, the use of the telephone effect in Ke$ha’s “Stephen” that we observed in Chapter 1 (see Section 1.3) suggests a dislocation of the protagonist at a crucial point in the narrative reading I proposed. As stated, the music of the middle-8 can be narrativised as a wedding scene whilst the protagonist asks whether Stephen will not call because he feels she is deluded, unattractive, or because he does not love her. A telephone effect applied to the voice here acts as a mimetic strategy that bonds it with a diegetic source: Stephen’s answerphone. Whilst the voice, then, appears diegetic, the dislocation from the wedding which it implies reveals that the musical frame here is sonifying a scenario *imagined* by the protagonist.

### 4.4.2 Musical Frame

Whilst “Stephen” provides an example in which the personic environment appears to sonify the internal psychology of a character, we have seen other cases already in which the environment appears to take on a diegetic role, or depict a non-diegetic setting. With this undecidability in mind, I would like to follow Altman for now in describing a *musical frame*, in which the personic environment is situated.

Noting the variety of stylistic or idiomatic approaches to instrumentation, Allan Moore notes that in discussions of musical texture, “[i]t is more fruitful not to begin from the instruments themselves, but from the functions they perform (the ‘layers’ they constitute) within the musical fabric” (2012b: 20). Accordingly, he advocates textural analysis in terms of four functional musical layers: the functional bass layer, melodic layer, ‘harmonic filler’ layer, and the explicit beat layer. By identifying these four musical functions, Moore provides us with the basis for a discussion of the musical frame in recorded popular song. When these elements
appear to occur within a different acoustical space to diegetic elements — as they often do — we might choose to follow Altman’s suggestion that musical elements are not, strictly speaking, diegetic.

In many cases, the diegetic locus of the musical frame is indeterminate. In the uncommon scenario that the persona directly addresses performers or in some way draws attention to the mediation of the situation through the act of a musical performance, we can argue that the protagonist/narrator and musicians occupy the same diegetic frame. It is not always clear, however, whether the agent voiced by the singer is aware that he/she is singing, nor that the musical environment could be audible to them. As Jeanette Bicknell writes, “the variety of functions [of singing in music drama], and the different scenarios that can arise from them, means that the ontological status of singing can vary from drama to drama” (2015: 96). Indeed, I have noted previously that the use of singing suggests an oneiric form of narrativity because it acts as a point of departure (see Section 3.2). The ontology of singing has important ramifications in terms of diegetic locus and Bicknell continues by highlighting two common categories of singing in relation to music drama, after Cone (1989).

The first of Cone’s categories is what he calls realistic song, in which characters sing as a matter of plot (that is, the story incorporates circumstances in which an agent sings). Singing in realistic song, therefore, is diegetic and can heard by other characters. One analogue of this position in recorded popular song is P!nk’s “Raise your Glass” because the persona misses its entry into the final chorus, indicating that the musical frame is audible to it. The opening section of “Welcome to the Black Parade” which I described earlier could also be seen in this light, as the song opens, it seems, with the protagonist playing the piano, in which case the musical frame is audible to him, and may also explain his singing.
The second category that Cone proposes, operatic song, describes cases in which characters do not appear to be aware of singing. Instead, singing in operatic song is used in place of more typical spoken conversation. This therefore describes the normative position of recorded popular song. Operatic song charges the listener to either envisage an alternative narrative world logic in which singing is a normal form of communication, or accommodate this as a momentary lapse in the logic of the narrative world resulting from a momentary reconfiguring of the diegesis. In Cone’s (1974) earlier discussion of singing personae, he prefigures his commentary on operatic song by suggesting that characters in opera are not aware of singing and that they do not hear their musical accompaniment. Importantly, he suggests instead that musical accompaniment is imagined by the character or exists on a subconscious level and shared with the audience.\(^{103}\)

The diegetic locus of the musical frame, then, is open to considerable discussion. After Altman, we might argue that the musical frame is generally non-diegetic. Conversely, after Cone’s operatic song and earlier remarks, we could think of the musical frame as a sonification of the mental interior of the persona. Finally, after the concept of realistic song, we could argue that, in some circumstances, the persona

\(^{103}\) Kivy (2013) also explores the applicability of Cone’s realistic song to film soundtracks and proposes three modes of realistic song in film: ornamental, in which songs are not mapped closely onto the drama; embedded, in which songs are more closely mapped onto the film’s drama and are a means of advancing the plot; and, integrated, in which an embedded song also recurs through instrumental performance elsewhere in a film’s soundtrack as a form to support or develop an understanding of the on-screen drama. Kivy further proposes the category of music-track song, which in this case is an example of what Cone calls operatic song and describes songs or music more generally which functions as a Greek chorus, or an unseen character which is heard by the audience (but not by the characters of the drama) and actively contributes to the unfolding drama.
and musical frame are both diegetic. The situation, as Bicknell has written, varies between texts. To resolve this indeterminacy, we will revisit the diegetic locus of the musical frame in the following chapter in relation to perspectivity, or the relationship afforded between the musical frame and the protagonist/narrator (Section 5.4).

4.4.3 Non-Diegetic Frame

*Non-diegetic frames* consist of elements which are not taken to be diegetic and also do not appear to function within the four functional musical layers Moore describes. We have already encountered some examples of this in cases of episodic markers, such as the projector in “7 Years” earlier in this chapter (see Section 4.2). In this example, the episodic markers are formally integrated within the track, but like the narrator, are dislocated from the story-now. Accordingly, the projector is non-diegetic. We have also seen related cases in which paramusical material in The Scissor Sisters’ “Let’s Have a Kiki” (the ambient sirens) and Madonna’s “Hollywood” (the birdsong) is cut in a manner which suggests a point of departure indicating that these materials are non-diegetic.

Another prominent instance of non-diegetic sound-events is the use of sonic logos to indicate the involvement of certain producers, such as the sample of a cell gate closing used by Akon within both his own work and his features on other artists’ releases.\(^{104}\) We can regard this sample as paramusical — in that it bears little relation to its musical contexts beyond straightforward pitch shifting and staging — acting on an intertextual level. Rather than connecting the narrative world with a reference

---

\(^{104}\) The sample in question can be heard in the opening seconds of Akon’s “Right Now (na na na)”, “Don't Matter”, and “I Wanna Love You”, in addition to his collaboration with Gwen Stefani “The Sweet Escape” at 9″, Leona Lewis’ “Forgive Me” at 4″, and Michael Jackson’s 2008 version of “Wanna be Startin' Something” at 3″.
worldview, it is unrelated to the narrative world and hence is an example of paramusical sound which does not necessarily contribute to the construction of the narrative world.

Gnarls Barkley’s “Crazy” offers our last example of non-diegetic sound-events. Initially, the non-diegetic frame is constituted of vinyl noise. As the vinyl noise loops each bar, it may result from the use of sampled material from earlier recordings, yet nevertheless constitutes a non-diegetic texture and — if narrativised — suggests the presence of an implied listener. As we approach each refrain, additional white noise with a resonant moving filter is added, a sound effect we might associate with wind in an outdoor space, until it is abruptly cut at the end of each refrain. Again, there are no mimetic strategies which bond the noise to a diegetic source, and it does not function in terms of Moore’s four layers.

4.5 Case Study: Kate Bush - “Waking the Witch”

As we have encountered several times throughout this chapter and the last, both songwriting strategies and the utility of contemporary music technology afford ways of creating tracks which suggest a dream-like approach to aspects such as plot, temporal sequencing, and phonographic constructions of space. Kate Bush’s “Waking the Witch” offers an example par excellence of such an oneiric aesthetic and, accordingly, invites considerable discussion concerning temporality and diegetic framing.

---

105 The community-maintained website WhoSampled (2016) indicates that the producers sampled two tracks from 1968 and 1972.
The song itself is presented as a trial in which a protagonist is accused of being a witch and appears to be put to death by drowning. In an interview, Bush offers some background to this scenario:

[what fascinated me in doing [“Waking the Witch”] was the idea of a witch-hunter hiding behind the priesthood, as a guise, and coming to get this woman who isn’t a witch, but he wants to make her so. The girl closes her eyes to get away from it and goes to a church where it's safe and secure [...] but the priest turns out to be the witch-hunter. [...] It’s based very much on other people’s imagery of Roman Catholicism which I've found fascinating — you know, the kind of oppression, even madness, it can create, I suppose, in some people (Bush 1985).

The track can be coarsely divided into three sections: an introduction that consists of various vocal samples of different people telling their addressees to wake up lasting t'19"; approximately 2'40" of musical discourse which forms the main body of the track, depicting a witch’s trial; and, a closing outro of sampled material laid over the last elements of the musical discourse as it fades out.

4.5.1 Introduction

The introductory segment situates us in a fantastic phonographic space which is clearly oneiric in character, particularly in terms of landscape and acoustical space. A series of vocal samples emerge from different locations within the phonographic space, and each time reconfigure the apparent dimensions and acoustical properties of the soundbox. This is reinforced by the incongruity of the vocal samples, which come from different speakers that mostly tell their addressee to wake up. Moreover, 106

106 Although the variety of speakers and discrepancies in phonographic staging give the impression of being sampled from elsewhere, the vocal samples were recorded by Bush’s family, mixing engineers, and actor Robbie Coltrane (Moy 2007: 50).
in the opening seconds, the piano is reversed to create an opening swell and is staged with a stereo delay effect that would not be acoustically possible outside of the recording studio.

In addition to the piano and vocal samples, whale song can be heard in the background, further supporting the construction of a dream-like atmosphere. At 1′03″, a crowd can be heard reciting “We are of the going water and the gone. We are of water in the holy land of water”, a phrase which reappears later in the album on “Jig of Life”. This crowd becomes particularly significant later in the track as it is revealed that they are the jury at the protagonist’s trial. With both the combined landscape and staging of the samples and piano in mind here, we can consider how they appear to relate to the narrative world in terms of diegetic framing. As the vocal samples are largely incongruent in terms of speakers and addressees, and, like the piano, are staged within a phonographic acoustical space which could not be reproduced without production technology, I take them to be external to the diegesis. Hence, the vocal and oceanic samples appear to constitute a non-diegetic paramusical frame, and the piano forms a musical frame. However, I shall return to discuss the crowd, which, I believe, prefaces the construction of a diegetic frame.

4.5.2 The Witch’s Trial

From 1′19″, clear metre and harmonic underpinning are established as we enter the main body of the track which depicts the witch’s trail. In place of the samples and piano of the introduction, drums, guitars, vocals, and synthesisers enter the soundbox. During this passage, a verse-chorus structure is eschewed by the presence of limited episodic musical or lyrical material. We can, however, sketch a sequence of distinct scenes by examining the interplay between different vocal sources and their phonographic staging. There are three vocal sources at play. The first is Kate's
persona, which voices the protagonist and can be heard both singing and speaking at different points in this passage. Her first introduction comes at the very end of the introduction at 1’19” as she pleads for help whilst staged using a rapid stereo tremolo effect. For the most part, however, she sings whilst staged in the centre of the soundbox. The characterisation of the protagonist as a suspected witch is established by the persona’s identification as a “blackbird” and calls for help because, she says, there is a stone around her leg. The second vocal source is a more masculine-sounding vocal part which takes on the role of the antagonist, a judge presiding over the trial. According to Moy (2007: 50), this second vocal part was also recorded by Bush and transposed down into a male speaking range, an effect which gives the voice a demonic character. The third vocal source is the crowd which takes on the role of a jury at several points, particularly when asked by the judge to pass a verdict.

Considering this combination of narrative agencies, reconfigurations of the soundbox, and lyrical cues reveals the following sequence of scenes:

Table 3: Formal Sketch of ‘Waking The Witch’

<table>
<thead>
<tr>
<th>Scene</th>
<th>Time</th>
<th>Agents present</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Protagonist</td>
<td>Judge</td>
<td>Jury</td>
</tr>
<tr>
<td>Introduction</td>
<td>0”</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pleading</td>
<td>1’19”</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Courtroom¹</td>
<td>1’26”</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Chant¹</td>
<td>1’37”</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Courtroom²</td>
<td>1’53”</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Chant²</td>
<td>2’01”</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confessional</td>
<td>2’23”</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Courtroom³</td>
<td>2’36”</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Drowning</td>
<td>3’46”</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Courtroom

As the tabular sketch above shows, the main recurring element during the body of the track is a courtroom scene featuring the protagonist and judge. During these passages, the judge makes a series of somewhat threatening and largely rhetorical statements whilst Kate sings. At first, her lines appear to bear no obvious relation to the judge: as he remarks almost playfully upon her impending death sentence, Kate appears to sing about flowers. Appropriately, her lines seem to draw from an earlier shanty associated with seamanship, which A. L. Lloyd and Ewan McColl have recorded under the name “Blood Red Roses”. With little alteration, Kate’s lines closely follow Lloyd and McColl’s “oh, you pinks and posies” and “go down you blood red roses, go down”. Although the lyrics of “Blood Red Roses” are somewhat unclear in themselves, their adaptation in “Waking the Witch” supports the reported overarching concept to which the track contributes.

As with each of the scenes in the track, the musical content that constitutes the environment consists primarily of repeating material which is harmonically static. During the courtroom scene, a rhythmical noise part provides a constant semiquaver rhythm, which is joined by occasional moments of bass and gated drums.\(^{107}\) The harmony is static, outlined by an electric guitar panned left which pivots around C\(^\#\), and E with an occasional B, spelling a C\(^\#m7\)\(^{(-5)}\). As this instrumentation bears no apparent relation to, and would be highly unorthodox in, the context of a trial, these can be straightforwardly thought of as removed from the diegesis. In contrast, the judge’s use of spoken monologue and, particularly, the use of reverberation to stage

---

\(^{107}\) Moy (2007: 50) attributes the noisy semiquavers to the processed sound of a helicopter blade, which prefigures the outro.
him as though in a possible space both act as mimetic strategies that afford imaging him within a diegetic frame.

The trial is split between three scenes in the track. In its first iteration, the judge attempts to intimidate the protagonist into a confession, telling her she will not burn, nor bleed. In the second, he patronises her and alludes to her drowning. In the third, he accuses her and calls to the jury for judgement, to which they emphatically shout “Guilty! Guilty! Guilty!” In this third scene of the trial, both the judge and jury are staged in the same acoustical space. Given that they also both speak and the way in which the crowd appear to respond directly to the judge, we can argue that these elements constitute a diegetic frame. Indeed, in this scene, the jury declares the protagonist guilty and seals her fate. Meanwhile, the instruments which constitute the environment appear to share a different acoustical space which suggests that they operate ‘alongside’ the world-specific elements.

Chanting

The first elision between scenes takes place at 1’37” and abruptly cuts off the courtroom after seven bars, a bar short of the end of the hypermetric group. Here, the protagonist chants a Latin phrase which translates as “in the name of the holy spirit”. The second, meanwhile, enters at 2’08” after an eight-bar passage, as the protagonist appears to have changed tack, now chanting (again in Latin) “god and master of the underworld”.

Both times this chanting scene arises, there is a remarkable change in texture. The gated drums continue, though the bass, rhythm guitar, and helicopter lines are removed. Instead, the musical frame here is reconstituted through the addition of (unquantised) church bells as the protagonist chants in latin, and a chorussed electric guitar enters, panned right, playing an Em figuration around E, F♯, and G to create a
cross rhythm against the 4/4 metre and harmonically-ambiguous church bells and vocals.108

Confessional

Following the second passage of chanting, we move to an additional scene in which the environment returns to the combination originally heard during the courtroom scenes, with the rhythm guitar now principally moving between C# and B. The lyrics place the scene in the context of a confessional, for the antagonist's voice which had previously led the trial scenes invites the protagonist to speak, to which Kate responds by asking for a blessing and begging for forgiveness. This suggests, like Bush’s earlier comments, that the judge is a member of the priesthood and that the protagonist — inadvisably — entrusts him.

4.5.3 Outro

The passage of explicit musical discourse begins to close around 3′46″ following the judge’s referencing of the title. At this point, the music begins to fade whilst recordings of helicopters and the voice of coastguards enter, who caution their addressee: “get out of the water”. The inclusion of this passage suggests that the protagonist has indeed been put to her death by drowning and while the final outcome is left open-ended, we might see the gradual fade of the music (in which Kate is now no longer to be heard) as emblematic of her gradual decent into the depths of the water.

108 It may also be worthwhile to note that the pattern played by the guitar closely resembles the Dies Irae (“day of judgement”), which is used as a Roman Catholic hymn and part of the Requiem Mass.
4.5.4 Discussion

Drawing from these observations, I would like to offer some points of discussion in terms of how the apparent temporal sequencing and diegetic structure contribute to the formation of a narrative world and may inform narrativisation, before considering this reading in the wider context of the *Hounds of Love* album.

In keeping with the oneiric, dream-like flavour of the track, its diegetic framing appears somewhat fluid. During the introduction segment, for instance, the extent of departure in terms of landscape (the combination of speakers and whale song) and acoustical space (the stereo piano delay and staging of the vocals) eschews the identification of a clear diegesis. Several scenes of the musical discourse which follows, however, do afford the construction of a diegetic basis to a narrative world. In the trial scenes, for instance, the antagonist’s use of spoken language and the addition of reverb to the voice both act as mimetic strategies that suggest that he is diegetic. The church bells and chanting demarcate a separate scene in the setting of a church, which is later extended by the confessional scene with its spoken dialogue. The guitars, drums, and so on, however, are not spatialised within the same apparent space of the diegetic frame and so cannot readily be classified, though we might think of the helicopter semiquavers as non-diegetic due to the way in which they prefigure the closing segment and are born from a paramusical sound-source.

Two important moments of audio dissolve can also be observed in the track. The first concerns the crowd during the introduction. Here, they appear alongside a variety of samples which do not appear to be diegetic. Yet, as the crowd return later in the track, it becomes clear that they are the jurors for the protagonist’s trial, making it apparent that they are, at this second point, diegetic. Similarly, the coastguard helicopters that fade in from 3’46” begin as non-diegetic but, as the music and previous diegetic layer fade, rapidly seem to move us to a different diegetic space.
In the reading developed above, the use of diegetic sound-events are crucial to both worldmaking and narrativisation. Indeed, they are integral to the unfolding trial. However, the track also provides an interesting case in which elements which do not appear to cohere with the narrative world — the helicopter and lifeguard samples — also make important contributions to plot, for this final scene allows us to infer that the protagonist was put to death by drowning following her guilty verdict.

The diegetic locus of the protagonist invites additional consideration. Although she responds to the judge, the phonographic staging of the voice within the soundbox suggests that she occupies the same space as the musical elements. When the voice is staged using stereo tremolo, it similarly does not appear to occupy the diegetic frame in terms of acoustical representation. This offers a subtle suggestion that the protagonist may in some manner be removed from the diegesis. If we favour this interpretation, this becomes a crucial point that she is not engaged (and presumably not represented) in the trial scene at all. For this to be the case, her guilty sentence is surely inevitable.

The dislocation of the protagonist from the trial scene also supports Bush’s own accounts on the construction of the album. Reportedly, the track forms part of a suite entitled ‘The Ninth Wave’ which spanned the second side of the *Hounds of Love* record. In her own words, Bush explains the concept:

> for me, from the beginning, The Ninth Wave was a film, that's how I thought of it. It's the idea of this person being in the water, how they've got there, we don't know. But the idea is that they've been on a ship and they've been washed over the side so they're alone in this water. And I find that horrific imagery, the thought of being completely alone in all this water (1992).
“Waking the Witch” comes part-way through this concept, at a time in which, Bush conceives, the protagonist has fallen asleep and has begun to have chilling nightmares (ibid).

Narrativising the song — following Bush’s comments — in the context of a dreaming character lost at sea allows us to think of the reading so far as metaphoric and encapsulate it within another level of narration. Beginning with the introductory section, which perhaps represents the most explicit moment of phonographic departure, the passage could be narrativised as the character’s liminal consciousness as they fall into heavier sleep. In which case, the section represents its own audio dissolve of sorts as we begin by being accommodated in an intermediate dreamworld before the diegesis is established during the musical discourse. Thereafter, the protagonist’s voice — which I have taken to be meta-diegetic — can be taken to represent the mental interior of the dreamer, who has no agency within the dream world. As observed by Kruse, though, the jumbled vocal lines produced using the tremolo effect at the opening of the second passage — which I have identified as non-diegetic, i.e. not belonging to the dream world, nor the mental interior of a character — “sound like Kate Bush’s cries for help as she bobs in and out of the water, gasping” (1990: 462). Similarly, the dissolve between the musical discourse and outro could be reinterpreted as moving between the diegesis of the witch trial to the higher level of the character at sea, now awake, as aid is sent to him/her. Using Bush’s commentary on the work, then, in conjunction with the analytical tools developed so far, we can address how the track affords narrativisation in relation to a dream world in the larger context of the dreamer’s own ongoing predicament.
4.6 Conclusion

Following our previous discussion of oneiric aesthetics in recorded popular song, this chapter has developed a finer-grained approach to the ways in which tracks mediate narrative time and ontological boundaries. As I have shown above, Genette’s understandings of narrative/story/narrating which I raised in Chapter 1 (see Section 1.1) are governed by different forms of time. The first, discourse time, relates to the duration taken for narrative discourse to describe an event or series of events. Secondly, story time refers to the temporal setting of a particular event or situation. Finally, narrator time can be used to address the temporal relationship between the protagonist/narrator and the event or situation reported. We can describe four possible types of narrator time after Genette: prior, simultaneous, subsequent, and interpolated.

Of the three categories of time raised, we have focused particularly on story time, which is most pertinent to the narrative world itself. Here, I have proposed that we can segment story time into scenes, i.e. passages in which story-related aspects are temporally or geographically demarcated. An important distinction can be drawn here between two types of scene. A common position, as several scholars have noted, involves synchronic writing that depicts a single moment in which story time is suspended. Alternatively, scenes may be diachronic, involving a passing of story time. It is common for songs to ‘toggle’ between these two forms of scene. In “Let’s Have a Kiki”, for instance, we open with a diachronic scene of an answerphone message being left, though this is followed by a synchronic scene in which the refrain is repeated and story time is suspended.

Aside from the sequencing of scenes, we have explored the ontology of sounds as they relate to a given scene with references to several mimetic strategies and points
of departure identified in Chapter 3. In recorded popular song, we can observe three categories of sound which appear to emanate from within or at some remove from the diegesis. Through particular paramusical sound-sources or phonographic staging, it is possible to afford the construction of what I have called a diegetic frame, which consists of components that appear bonded to sources within the story-now. The second form consists of musical discourse, which typically occupies an alternative acoustical space to the diegetic frame and suggests that it does not emanate from a diegetic source. Indeed, in the next chapter, I shall argue that a dominant use of the music is in support of a persona’s reported view and hence the music in this case can be read as internally perceived or as a depiction of a character’s internal thoughts or feelings. The final case is the non-diegetic frame, which consists of elements which may be connected to the narrative world but are in some way removed from the diegesis. This category includes sound-sources which may be integrated within a musical syntax, such as the projector sample of “7 Years”, as well as those which appear at a further remove from the diegesis, such as the addition of sounds relating to playback which position the listener as a voyeur upon a fictional listener’s experiences.

Having discussed worldmaking on the level of each scene, in the following chapter, we will move to address the related concern for the construction of the implied listener and protagonist/narrator. In particular, we will see how lyrical construction and the organisation of the phonographic environment lead to particular perspectives upon/within the narrative world.
5. Perspectivity

In an investigation of the possible readings of the word “here” in song lyrics, Allan Moore makes an intriguing observation of Frank Sinatra and Dinah Shore’s recording of ‘The night is young’:

[b]oth voices are in an intimate space, and are very dry, compensated by the orchestra that has a measure of room reverberation — the instrumentalists are in a clearly different space, not interacting at all with the singers [...] Sinatra’s and Shore’s ‘here’ is marked out as private, but inhabits a privacy into which we are invited to eavesdrop (2010b: 178).

By noting the distinction between the staging of the voices and orchestra, Moore’s writing suggests that the construction of the phonographic environment evokes a spatial blend. With the three frames put forward previously in mind (see Section 4.4), we could view this as the sound of the lovers occupying a diegetic frame whilst the orchestra provides a musical frame which is not diegetic; a situation that we have already seen in the middle-8 of “We Are Never Ever Getting Back Together” (see Section 4.0). However, Moore’s expansion on this issue by noting that the phonographic environment implies the position of an eavesdropper — an agent which I refer to as the implied listener — introduces the final parameter of narrative worlds which we will discuss: perspectivity.

Whilst the phonographic environment may evoke a particular point of view through the manifestation of an implied listener, the issue of perspectivity extends also to the persona. So far, we have seen examples of protagonists who demonstrate different levels of access to the internal psychology of characters and perceptual access. Whilst “Hollywood” (see Section 3.5), for instance, is sung from the perspective
of the protagonist, the narrator of “Waterloo Sunset” (see Section 2.6) is restricted in terms of line-of-sight but demonstrates knowledge of the lovers’ internal feelings.

Both issues introduced here — i.e. the perspectivity of the implied listener and that of the persona — share a common basis in the hypothesis that narrative worldmaking is mediated by the perspectives available to the world’s inhabitants. This suggestion is reflected by Ansgar Nünning, who argues: “in addition to the making of events, stories, and emplotment, narrative worldmaking also involves another important aspect which has, however, faded into the background, and undeservedly so: the extensive importance of perspectivity or point of view” (2010: 205). To develop this line of thinking, I would like to consider the perspective adopted by the persona, and the relationships between the diegetic loci of the persona and phonographic environment. Hence, we will ask here whether the persona reports from a diegetic perspective (i.e. simultaneous narration of the base level of a story), or is removed from the circumstances put to us (such as through prior or subsequent narration, or the telling of a subsidiary story); and, whether the phonographic environment appears situated in the same diegetic locus, or instead appears to sonify another situation or aspect of a character’s mental interior.

5.1 Focalization Theory

In proposing that the listener’s imagination of a narrative world is informed by the perspectivity of its inhabitants, I have broached an analogue to what Bal calls the *focalizor.*

For Bal, the focalizor is an agent who perceives the narrative world of a text and colours it through his/her point of view, which is put to the reader. It is

---

109 Focalization is a large topic of discussion within narratology with a number of approaches which I cannot satisfactorily summarise in depth here. For an introduction to dominant approaches, see Jahn (2005).
important to note the purposeful distinction here between the *perceiving* consciousness and the narrator who *communicates* these perceptions, as this underpins the distinction between *fabula* — a chronological sequence of events — and *story* — the formulation of the fabula as perceived by a focalizor and reported by the narrator. Bal summarises focalization in relation to narration with the phrase “A says that B sees what C is doing” (2009: 149).\(^{110}\) In this case, A is our narrator, and B is the focalizor who perceives the actions of a narrative agent, C. In fiction, this distinction is often made clear using subsequent narration. Harper Lee’s *To Kill a Mockingbird*, for instance, is narrated by the central character, Scout, some time after the story. In this case, the adult Scout is our narrator (A), the younger Scout acts as focalizor (B), who perceived the actions of the local community throughout the story (C). Conversely, Virginia Woolf’s *Flush* presents the biographical tale of a dog, Flush. In this case, we are told by an indeterminate narrator (A) the perceptions of Flush (B) as he interacts with his family and the other characters of the story (C).

Bal’s discussion of focalization develops Genette’s (1980) three-fold typology that considers the narrator’s access to characters’ internal psychology and its embodiment or apparent scope of sensory perception.\(^{111}\) His first category, *zero*  

\(^{110}\) Although Bal employs a visual metaphor for the purposes of this example, her work illustrates broader concerns for taste, hearing, and other forms of perception which Nelles (1990) argues is similarly relevant to the concept.  

\(^{111}\) Whilst Genette described focalization as synonymous with the terms “point of view” and “perspective”, which both suggest an attitudinal meaning, Niederhoff (2011) stresses an essential distinction between focalization and narrative voice. Focalization determines the narrator’s access to information; the lens through which the fabula is shaped into a story. Elsewhere, Rimmon-Kenan (1983: 77-82) has suggested expansion to account for: perceptual focalization, which relates to aspects of space and time; psychological focalization which she
focalization, refers to omniscient narrators that are capable of accessing information known by other characters within the story. Genette’s second category, internal focalization, restricts narrated information to the spatial limits and mental interior of a single party. Finally, external focalization, describes a situation in which narrative information is constrained both in terms of space and psychological access, as though perceived by a bystander who is not party to the knowledge of other characters. As Deleyto (1991) also argues, though, the distinction of voice from perception, and hence narration from focalization, requires the presence of an agent who focalizes: Bal’s focalizor. Bal further reconsiders focalization in terms of diegetic localisation, i.e. whether a narrative is focalized from within the diegesis or from a position outside of a narrative event. As she writes, “[w]hen focalization lies with one character which participates in the fabula as an actor, we could refer to internal focalization. We can then indicate by means of the term external focalization that an anonymous agent, situated outside the fabula, is functioning as focalizor” (2009: 152, my emphasis).

Following Bal’s reformulation of focalization to account for diegetic locus, Edmiston (1989) has argued for a revision to focalization theory concerning first person narrators. Edmiston addresses a common logical inconsistency: he suggests that scholars tend to approach first-person narrators as internally focalized and so model the narrator as diegetic. However, as he rightly points out, first person narrators are not necessarily diegetic. This problematises internal focalization, as internally focalized first-person narrators can by definition only access their own internal psychology in their present, not the past or future. In response, Edmiston proposes four categories of focalization of first-person narrators which expand on

relates to cognition and emotiveness; and, an ideological component corresponding to an attitudinal perspective.
Genette’s understanding of focalization in terms of perceptual and psychological access:

- **paralipsis** - revealing less than is/was knowable to the narrator in the story-now;
- **internal** - revealing only information that is/was knowable to the narrator in the story-now;
- **external** - revealing information that was unknowable to the narrator during the story-now, perhaps through the use of corrective statements or the inclusion of knowledge gained subsequently;
- **paralepsis (zero)** - revealing more than could be known to any character; the narrator is omniscient.

In the following section, we will apply these four categories which Edmiston highlights to the role of the persona in a song.

### 5.2 Focalization of the Persona

For the persona to adopt the position of a *protagonist/narrator*, limits are generally placed upon the knowledge or psychological access and perceptual information which it relays to the listener. The persona, therefore, is focalized. The role of focalization is especially clear when a persona voices multiple characters within the same track or otherwise calls attention to its mediated nature. Eminem’s performances as Slim Shady, for instance, demonstrate that he has adopted perceptual and cognitive traits of a fictional character. We can compare this to “When I’m Gone”, which is written as a monologue in which Eminem takes on a different persona to address a young daughter named Hailie (after the artist’s real-life
daughter) and acknowledges a distinction from his other persona by referring to his Slim Shady alter-ego.

As the persona generally takes on the role of a protagonist, they function in a manner which is analogous for the most part to a first-person narrator and so Edmiston’s system of focalization offers us a helpful basis, particularly in the various cases of non-diegetic protagonists.

### 5.2.1 Paralipsis

Edmiston describes paralipsis as a situation in which the narrator is focalized through the internal psychology and perceptions of a single character but omits known information. We could also see this in popular song personae, albeit infrequently, as it requires some indication that information has been omitted. “I Didn't Know the Gun was Loaded”, as popularised by The Andrews Sisters, offers one such example in which lyric analysis reveals paralitical focalization. During the song, the singers tell the tale of the murderous Miss Effie. The persona begins by voicing a narrator in verse one and contextualises the story by introducing Miss Effie, who would reportedly drive men insane and shoot men somewhat indiscriminately. Verse 2 explains that Miss Effie shot the sheriff and was prosecuted in court, before being declared innocent by the jury in verse 3 and killed by the sheriff’s indignant wife. In each chorus, however, the Andrews Sisters adopt the role of the shooters as each pleads her innocence — firstly Miss Effie and, in the final chorus, the sheriff’s wife — by reciting the titular refrain. This presents us with two forms of focalization: the verses demonstrate omniscience and so we can regard this as an instance of what Edmiston calls paralepsis or zero-focalization. However, taking this narration to be reliable, the accounts during the choruses exemplify paralipsis, as the persona is character-bound and withholds information from the listener.
Whilst “I Didn’t Know the Gun was Loaded” demonstrates paralithical focalization, this is only possible because of the additional presence of an informed narrator. Otherwise, it is challenging to argue for this form of focalization because the listener has no means of assessing the extent of knowledge known to the protagonist which is undivulged, unless this is made explicit within the lyrics or through the arrangement. To make this information explicit requires the persona to act as a narrating third party.

Considering the persona within its broader musical context may offer support to discussions of focalization. Gotye’s “Somebody that I Used to Know” offers an example in which the music qualifies the focalization of the persona. The song’s lyrics relay a story of a former romantic couple following an acrimonious separation from two personae voiced by Gotye and Kimbra. During the first two verses, Gotye’s persona adopts the role of a protagonist, who describes the loneliness and sadness he reportedly felt during the relationship. In the chorus which follows, he exclaims to his addressee that she didn’t need to break all ties with him and attributes the blame squarely upon her. Thus far, Gotye’s persona appears internally focalized, as it voices the internal psychology of the protagonist. However, in the verse which follows, we hear a second persona voice Gotye’s addressee. She highlights in particular the protagonist’s failings as a lover, alluding to his pedantry and manipulativeness. The protagonist may not be as blameless as he suggests.

Kimbra’s persona sheds new light on what the male protagonist had previously reported. As I have noted, we could argue that both personae are focalized internally through the characters’ subjective positions but closer attention to the construction of the song is instructive here. The song, which consists of three verses and two instances of the chorus, is proportioned in clear favour of the protagonist: following Kimbra’s verse, Gotye immediately returns with the chorus. It is
particularly ironic that Gotye should interrupt here, as this opening line concerns being ‘cut off’. In total, he sings two verses and both choruses, whereas Kimbra’s role is far smaller, a disparity in discourse time which may suggest an indication of the protagonist’s guilt, as though Kimbra is being silenced prematurely or unfairly. Hence, the musical context sheds doubt upon Gotye’s accuracy and suggests he is paraliptically focalized.

5.2.2 Internal

Most often, the persona in popular song takes on a first-person, confessional mode which reports the experience of the protagonist from its subjective, character-bound perspective. This echoes Edmiston’s understanding of internal focalization, in which the narrator speaks from the perspective of a narrative agent that is spatially constrained, located in the story-now, and accesses the corresponding character’s knowledge. As this is a dominant position, there are a great many examples of internal personic focalization, and so I would like to highlight two instances which illustrate an association between the personic environment and focalization, and shifts in focalizor between scenes.

“Say Say Say”, as performed by Michael Jackson and Paul McCartney, provides one example of internal focalization and demonstrates the role of the musical environment to support this discussion. Two personae alternate, both addressing a disinterested romantic target. Both clearly speak from first person points of view but at several points, Michael doubles Paul as a backing vocalist, which could suggest that the two personae are focalized as the same character. However, the personic environment once again offers some clarification, as a distinction is made between the musical environments of Michael’s and Paul’s personae: Paul’s environment is characterised by a bright synth bassline which plays repeated
semiquavers, and additional percussion; Michael’s features a syncopated chorussed guitar and, later, brass. Moreover, the two personae are separated by their stereo placement towards either side of the soundbox.

For a second example of internal focalization, I would like to return to our previous discussion of “Waterloo Sunset” (see Section 2.6). As noted before, the last verse involves a change in lyric in which the narrator begins to sing about the feelings of Terry and Julie rather than his own. As the narrator here appears to begin through his own perceptions, looking at the world outside, we would characterise the persona as internally focalized. As we arrive at the final verse, this continues, judging by his likening of pedestrians with a swarm of flies. The persona continues to be internally focalized, although now Terry and Julie become the focalizors.

5.2.3 External

Edminston’s formulation of external focalization involves the narrator being diegetically dislocated from the circumstances they describe, typically narrating from a perspective after the events they recount and therefore capable of knowing more than his/her corresponding self of the story-now. Despite implicit access to more information through temporal privilege, this category of focalizor remains restricted to a particular space and a single character’s internal psychology. Like paralipsis, this requires non-standard lyric construction in order to acknowledge a distinction between the narrator’s present and the story-now, often through subsequent narration.

Maggie Rose’s “Looking Back Now” provides a helpful example of an externally-focalized first-person persona, as is clear from the very first line, derived from the title, which explicitly uses subsequent narration. Here, the persona demonstrates access to the internal psychology of a non-diegetic narrative agent as
she recalls a series of events leading to her present. In the first verse, she continues by describing a scene which concludes with her cheating husband's provocation to shoot him. By implication from the chorus, we understand that she shot him and indeed, the second verse shifts closer to the narrator’s temporal location, now describing the protagonist’s experience in jail. In this verse, she explains how a guard “got sweet” on her before he became aggressive, which prompted her to grab his pistol and shoot him. As these both involve subsequent narration by a non-diegetic narrator, the verses are externally focalized. Only in the final verse do we move to internal focalization via a shift into the present tense and made explicit with the first mentions of her experienced perceptions: as she is put to death with the lethal injection of sodium theopentol, she tells us that she is afraid, that her hands are growing cold, and her vision is fading.

We can also observe this form of focalization in Bon Jovi’s “Blood on Blood”. The track is narrated from the perspective of one of a group of friends who promise lifelong friendship to one another. In the second verse, the protagonist recalls the hero of the group taking the blame for stealing cigarettes and being taken to a cheap motel by a girl who “turned [them] into men”. The third verse shifts to simultaneous narration and explains that now the two other friends have grown up to take mundane jobs whilst the protagonist sings in a band but pledges his continuing support for his friends nevertheless. As the protagonist primarily employs subsequent narration, with an inflection into prior narration in the refrain, it is clear that he draws on knowledge from after the story-now. Hence, the narrator is non-diegetic and externally focalized.

This form of focalization also occurs in Just Jack’s “The Day I Died”, but is concealed by the construction of the verses. In each verse until the ninth (in which he leaves work), the persona, appearing to voice an experiencing protagonist, seems to
employ simultaneous narration, giving the impression of a protagonist being ‘in the moment’. In the chorus, however, the temporal setting appears to shift to an alternative present as the persona reports that the protagonist’s last day alive was the best of his life. This is concretised in the ninth verse, cast in the imperfect present tense as the protagonist hurriedly describes leaving work and picking up cigarettes from a local store before crossing the road and remarking on not noticing an oncoming taxi. With the protagonist’s death now confirmed, it is clear that whilst the verses appear to be internally focalized, this is in fact a non-diegetic narrator and hence is focalized externally.

5.2.4 Paralepsis (zero)

Paralepsis or zero focalization, involves an omniscient focalizor whose line of sight extends beyond what could ever have been known, whether or not the first person narrator is temporally dislocated from the story-now.

The Beatles’ ‘Eleanor Rigby’ offers an example of paralepsis in its brief tale of two key characters, Eleanor Rigby and Father McKenzie, who are introduced in verses 1 and 2, respectively. In the first verse, Eleanor picks rice from a church ground after a wedding and, we might infer from the mention of dreaming, imagines the role of the bride. Yet, we are told immediately after that she spends her time indoors by the window, presumably waiting for acknowledgement from a passer-by. The second verse joins Father McKenzie as he darns his socks alone one night. For Father McKenzie to be alone here, the narrator must be omniscient. Indeed, the narrator is assured that the father’s sermon will go unattended and unheard. In the final verse, the two characters come together, but only after Eleanor has died and Father McKenzie leads her burial. As the narrator explains, no one attended the funeral and
nobody was rescued (appearing to comment on the ceremony’s lack of impact), again demonstrating his omniscience.

Similar narrative tactics occur in Kate Bush’s “Babooshka”, a song in which the persona takes on the role of narrator to tell the tale of a wife in an uninspiring marriage seducing her husband under an alias, firstly by sending scented letters and later meeting him in person. “Babooshka” is paraleptically focalized, as made apparent in this case through the narrator’s demonstration not only of omniscience but also her perceptual and psychological access, capable of relaying the husband’s peculiar thrill receiving her letters and, more tellingly, the uncanny resemblance he felt upon meeting his wife in disguise.

Kavinsky’s “Pacific Coast Highway”, provides us with a final example. The persona is focalized as an omniscient narrator, whose simultaneous narration presents a police chase with a phantom driver, as though adopting the role of a newscaster. However, despite presenting information ‘in the moment’, the narrator speaks from an informed position and at several points employs prescient prior narration. His first line, for instance, alludes to prior odd happenings at night but states that there are more peculiar things to come. Indeed they do, for this driver is reportedly capable of not only outrunning the police but at one point appears to disappear entirely. Throughout the track, the persona also employs prior narration to highlight the last sighting of the driver and the route that he takes, revealing this persona to be focalized through an omniscient narrator.

5.3 Point of Listening and the Implied Listener

I would like now to turn to a second form of focalization which concerns the music that we hear. In the field of art music, Meelberg (2006: 66-71) argues that a musical
performance represents a focalization of the events represented in a musical score and, following this argument, I would like to suggest that we could also make a case for the role of record production as a form of focalization, as this involves the construction of a particular temporal and spatial arrangement of sound-events. In his discussion of radiophonic plays, Beck lays out supportive ground for this thinking by explaining: “in BBC production, directors and studio managers (panel operators) talk of the 'sound picture' and what is 'in' and what is 'out'. The radio drama play scene operates by a series of sound pictures and each of these has a perspective and an acoustic” (2009: 7.2). Beck’s identification of an aural perspective is pertinent, too, to recorded popular music, which necessarily constructs an aural perspective through the selection and organisation of sound-sources. In Beck’s discussion, these contribute to what he refers to as a point of listening.112

Beck develops his understanding of point of listening from Chion’s (1994) earlier discussion of film soundtracks, in which Chion coins the term point of audition to describe the location from which we appear to hear the diegetic level of film sound, an opportunity for comparison with the point of viewing, i.e. the perspective asserted by the positioning of the camera. For Chion, the point of audition is manifested from two parameters: a location within space in relation to the narrative world onscreen; and, secondly, he asks whose aural perspective is put to the audience, i.e. which or what form of narrative agent appears to be ‘hearing’ what we listen to.

With Chion’s two factors of point of audition as my basis, I would like to reflect on two components of the focalization of the phonographic environment: point of

112 Beck’s other concern for the “acoustic”, or the acoustical properties of a phonographic space that qualify a location in which the narrative events are taken to unfold, is less suitable for my discussion than perspectivity because, as Doyle (2005) emphasises, popular music production often favours close-miking techniques.
listening, to refer to the organisation of sound-sources and the corresponding sense of location within a space; and *implied listener* to describe the agent or experiencing consciousness which one could imagine hears sound-events (see Section 2.1). The *implied listener*, therefore, is an anthropomorphism of the *point of listening*, narrativised through phonographic aspects of the track, which acts as an aural focalizor.

The idea of an implied listener is also supported by Link’s discussion of noise in recorded music, in which he suggests noise in a track creates a mimetic projection of an intermediary’s listening experience. Effectively, this positions the listener as a voyeur upon the listening situation of another and so makes an *implied listener* explicit. In such cases, he puts it that “the recording becomes a fictional space that allows us the indirect experience, perhaps, of listening to someone else listen” (2001: 38). Link’s “fictional space” here can be thought of as a non-diegetic frame and, importantly, he highlights the manifestation of an *implied listener* by a recording. With the three short examples which follow, I would like to show some other ways in which we could imagine an *implied listener*.

A helpful opening example can be found in Bright Eyes’ “At the Bottom of Everything”. As the track begins, a first person narrator appears to sip a drink and start telling a story using subsequent narration. The reverb character is typical of a modestly-sized room, rather than a close-miked studio session - the speaker even momentarily appears to move away from the microphone, presumably in order to put down his drink and pick up his guitar. The persona speaks, rather than sings; is not situated within a musical environment; and, is temporally dislocated from the story-now due to subsequent narration. The narrator, therefore, is non-diegetic and the point of listening within the same room correspondingly suggests a non-diegetic
implied listener. This is supported by the acoustic profile of the phonographic environment and the apparent proximity to the speaker.

In some cases, the implied listener could alternatively appear to occupy the diegesis. Take Queens of the Stone Age’s “You Think I Ain’t Worth a Dollar, but I Feel Like a Millionaire”, which begins with a landscape that evokes a driver as he/she enters a car and tunes the radio. This radio content is staged through a bandpass filter to simulate the tinny quality of small speakers, which continues until 1′00”, part way through the introduction of the musical track, at which point the filtering of the drums and electric guitar is rather unexpectedly bypassed, and they become staged with a fuller bandwidth. In “Song for the Deaf”, which comes later in the same album, the radio presenter’s voice returns, again in the story-now, which indicates a diegetic implied listener. Imogen Heap’s “Propeller Seeds” includes a similar moment in which the implied listener (here non-diegetic) appears to mute crowd noise at the push of a button to coincide with a line in which she refers to deep conversation as though the other guests were muted. Heap diverges, though, from Queens of the Stone Age because prior to this moment of world deformation, what appears to be a diegetic frame had already been established, moving through several scenes such as an opening interior space, woodland, and restaurant.

The phonographic environments of both the Imogen Heap and Queens of the Stone Age tracks suggest that we are party to the listening experience of another: that is, an implied listener is manifested by sound-sources within the phonographic environment. The third case which I would like to raise moves us away from the idea of a possible listener and can be narrativised in cases where the phonographic environment can be heard as sonifying the mental interior of a character. From 1′ in

---

113 This is as opposed to the corresponding diegetic position on board the plane ride described later by the persona.
“You Think I Ain’t Worth a Dollar, but I Feel Like a Millionaire”, for instance, electric guitars are panned either side of the stereo field, aggressively distorted as the persona delivers several lines screaming. This supports the protagonist’s opening lines in which he compares himself to a conquistador. Similarly, from t’39” in “At the Bottom of Everything”, a musical frame enters following the narrator’s count-in. The staging of the voice changes here, and is relocated far closer to the point of listening and so the verisimilitude of the phonographic environment implied by the construction of acoustical space is lost. Over the verses which follow, the protagonist touches on various sombre themes and so we might think of the musical environment here as oppositional given its major tonality and brisk strumming but the persona, too, does not sound miserable. Indeed, when reporting that the plane plunges into the sea, there is noticeable excitement in the voice. In the final line, the protagonist makes his happiness explicit and so we may instead think of the phonographic environment as supporting his general attitude and, as with “You Think I Ain’t Worth a Dollar, but I Feel Like a Millionaire”, as sonifying the mental interior of the protagonist.

Here, then, we can see manifestations of an implied listener that exhibit three different relationships to the diegesis: a diegetic relationship in which the phonographic environment depicts the story-now; a non-diegetic relationship in which the implied listener appears to inhabit a frame temporally dislocated from the story-now; and, a meta-diegetic relationship in which the phonographic environment appears to sonify the internal psychology of a character. In each case, the combination and presentation of sound-sources establish a point of listening which can be anthropomorphised as being perceived or sonified by an implied listener.
5.4 Focalization of the Implied Listener

The examples raised above support a hypothesis that the phonographic environment presents the listener with a perspective, which we can think of as being perceived — and hence, focalized — by what I have termed the *implied listener*. I would like to follow Bal’s concern for diegetic locus and revisit the possible diegetic loci of musical frames. In this section, I shall discuss how the *implied listener* may be focalized in different scenes with attention to the musical environment, landscape, and staging within the soundbox.

My approach compares the diegetic locus of the implied listener with the protagonist/narrator, leading to three possible relationships. Firstly, drawing from Chion’s theory of subjective internal sound, we can argue that the phonographic environment sonifies the mental interior of a narrative agent in cases of musical environments which are consistent with the persona. Secondly, the phonographic environment may incorporate mimetic strategies that bond sound-sources to diegetic elements audible to an implied listener which is coherent with the diegetic locus of the narrator. Finally, a third possibility which we have encountered involves the phonographic environment appearing to depict an alternative space which is diegetically dislocated from the protagonist/narrator.

5.4.1 Meta-diegetic Focalization

In many recorded popular songs, we can observe a normative position in which the point of listening is not closely localised in relation to the diegesis. In such cases, Moore’s (2005) typology between the musical environment and persona suggests a relationship with the internal thoughts or feelings of the protagonist or other
narrative agent. Similarly, Bordwell and Thompson’s (1985) concept of internal sound, and Chion’s (1994) later understanding of ‘subjective internal sound’ refer to some sonification of a character’s internal psychology, reminiscent of Cone’s (1974) contention that music can be likened to the inner thoughts or feelings of a composer’s persona (see Section 4.4.2). This situation, in which the phonographic environment is meta-diegetic, models music as the sonification of the mental interior of a character (whether or not it is audible to the character). I will refer to this as *meta-diegetic focalization.*

Young connects the sonification of internal thoughts or feelings which I refer to as meta-diegetic focalization with phonographic departure: “[b]ecause they create a sense of detachment from known physical Reality [surreal or abstracted sonic environments] may be taken as a metaphorical representation of the inner world of the imagination, where free and fantastic associations between objects and experiences can take place” (1996: 73). With this in mind, the creative use of phonographic staging in Eminem and Rihanna’s collaboration “The Monster” provides a helpful starting point. In this track, Eminem raps verses before switching with Rhianna, who sings the chorusses. As both employ a first-person perspective, cover similar subject matter, and share the line “I’m friends with the monster”, it appears that both Eminem’s and Rhianna’s personae are focalized through one

---

114 This is the “received view”, as Moore puts it, “of the role of an accompaniment, which is to conform to the ostensible meaning of the lyrics and, in doing so, to support, or perhaps illustrate, the meaning of the song” (2005: par. 9).

115 Chion proposes this to distinguish from ‘objective internal sound’, which is generally physiological and so could be audible to other characters.

116 This understanding departs from Gorbman’s original understanding of meta-diegetic sound, which describes sound which is heard by a given character through imagination or hallucination, for instance.
The lyrics, though, highlight a frenetic state of mind as Eminem’s verses express anxiety surrounding the protagonist’s own mental stability and emphasise the multifaceted nature of the protagonist’s psyche. This, then, is neatly reflected by the incorporation of two singers, but enhanced by the staging of the musical frame. In the second verse, for instance, the vocal line is briefly interjected at **1′37”** by a moment of formant-shifted rap which transforms Eminem’s voice beyond recognisability. Later in this verse, gating is applied to the musical materials giving the impression of audio drop-outs. The voices are further differentiated by the treatment of Rhianna’s voice with a stereo delay effect which suggests an alternative acoustical space to that inhabited by Eminem. The phonographic presentation of these aspects evokes an irrealist space representative of the fraught psychology which the protagonist describes. The protagonist, then, also functions as the implied listener here, as the track offers an opening into the protagonist’s mental interior.

“The Monster” demonstrates a typical case in which the protagonist is the focalizor of both the personae and the implied listener. Alternatively, Coldplay’s “Paradise” offers an example where this is not the case. Once again, the phonographic environment suggests meta-diegetic focalization, but focalized through the addressee. Although the persona voices a non-diegetic narrator which uses subsequent narration, it is focalized internally through his addressee, capable of presenting to us the thoughts and experiences of a woman in her childhood and as she ages. In the first verse, the narrator sets the scene of the woman in her youth, her unfulfilled expectations, and her dreams of paradise. In the musical frame, the persona is joined by a piano, which provides a rudimentary harmonic backing before adding more decoration to coincide with the mention of encountering paradise when closing her eyes, and a drum kit. The voice is positioned slightly right of centre, the drums towards the left (especially the hihat, which is most prominent), and the piano
towards the right with some stereo spread. The piano, though, is conspicuously equalised to give it a bright, but somewhat thin, character. As the persona reaches the mention of paradise, a discernible shift occurs in the staging of the voice, which appears to recede within the soundbox and pan slightly from left to right of the centre each time “para-” is sung. Two important aspects to note are that despite the slight movement of the voice, these sound-sources remain fairly tightly clustered within a fairly empty soundbox, and the staging of the piano which connotes age by schematising the limited frequency response of dated broadcast media.

At 1’29”, with the mention of closing eyes, a fuller-bandwidth Csus4 chord on strings is spread across the soundbox, as a synth or processed electric guitar flits between left and right, and a single jingle bell gesture is panned hard-left. In the verse which follows, the string ensemble joins the piano to offer harmonic support, though their timbre, like the piano, sounds somewhat tinny, especially at the start of the verse. Later, in preparation for the full refrain, the texture thins to lead vocal and piano, but the soundbox once again becomes significantly fuller as strings return, which are doubled by a synthesiser pad, along with additional percussion and a bass synth. The repeated coincidence between the ‘richer’ and ‘warmer’ fuller use of spectral space during the refrain suggests that the implied listener — in this case, the addressee — is meta-diegetically focalized, commensurate with the addressee’s experience which the narrator relays.

5.4.2 Coherent Focalization

In some cases, the phonographic environment in a given scene may incorporate diegetic elements or employ other mimetic strategies alongside the musical frame and locate the implied listener within the same diegetic locus as the protagonist/narrator. Take the opening moments of Jason Derulo’s “Cheyanne”.

192
this case, the track begins with a muffled distorted electric guitar and crowd noise. We then hear a knock and a voice calling the singer to come out, ready for the show to begin. From this information, we can construe that the phonographic environment has been focalized through a diegetic implied listener, as the point of listening has been constructed inside a room backstage. Thereafter, when the diegetic frame is removed and a new scene begins at 25″, we move to meta-diegetic focalization, which continues for the remaining duration of the track.

Successive scenes might involve different focalizors. As we have seen in Section 5.3, Imogen Heap’s “Propeller Seeds” makes liberal use of diegetic sound-sources. In this case, the binaural mixing in order to create a more convincing 360-degree acoustical image demonstrates a perceived desire to afford the listener immersion within the acoustical space, which consists of several identifiable locations (a restaurant, an outdoor forest, and the recording studio) through the use of a diegetic frame. This too, then, offers an example of coherent focalization of the implied listener. Yet, at 1′40″, the clicking sound of a button to mute the conversation of the restaurant scene implicates the presence also of a non-diegetic focalizor.

In some cases, the persona may also offer cues which support coherent focalization of the implied listener. Consider, for instance, P!nk’s “Raise your Glass” and Carly Simon’s “You’re so Vain”, which both involve some form of reference to the fact that the persona is singing, an analogue to Cone’s ‘realistic’ form of opera. In both cases, the suggestion that the musical environment is audible to the persona indicates that the focalizor is not meta-diegetic, but rather is outside of a diegesis, consistently with the persona.

One final example of a coherently-focalized implied listener can be found in Pink Floyd’s Wish You Were Here album at the end of “Have a Cigar” and the
beginning of the title track. During the last section of ‘Have a Cigar,’ the soundbox is comprised of drums, guitars, and synthesizers presented in their full frequency bandwidth. At 4′50″, though, a loud sweep of noise is introduced, after which the previous instrumentation is filtered to suggest the frequency characteristics of music being played from a small AM radio, now staged in mono with added room ambience. The filtered music then gradually moves to find a position on the right of the soundbox. At the start of “Wish You Were Here”, the music of the previous track is interrupted by snippets of noise and speech staged in the same way, which create the effect of tuning the radio to a new station. A station is found featuring a guitar which, at 57″, is joined by a full-bandwidth guitar staged very close to the point of listening and away from the radio, as though the listener is playing along to the music on the radio. In this sequence, then, the last seconds of “Have a Cigar” can be described as meta-diegetic as the diegesis is deformed, until the new diegesis involving the radio listener is concretised and we enter into coherent focalization.

5.4.3 Dislocated Focalization

The third form of phonographic focalization I would like to propose involves implied listeners who appear to be set in a space which is diegetically dislocated from the position of the protagonist/narrator. One example of this position occurs in the case of Daft Punk’s “Giorgio by Moroder”. As briefly noted previously in Section 2.4, the track begins with the sounds of conversation, cutlery, and Giorgio recounting the beginnings of his musical career to his addressee. This diegetic frame affords constructing a narrative world centred around the interior of a restaurant. Some distance away, a band can be heard playing softly. Due to this placement and the coherent acoustic, we might straightforwardly also locate this musical frame as diegetic, as though the entire phonographic environment here is focalized diegetically by Giorgio’s addressee. From 34″, though, the phonographic
environment is reconfigured and we move away from coherent focalization to dislocated focalization as we enter a new musical scene in which Giorgio starts to describe his early relationship with German discotheques around 1970. As he does so, the music which the band was playing moves closer in the soundbox and two guitars are added, placed on either side of the soundbox and processed using a wah effect, which can be found prominently in disco music. The musical frame at this point, then, appears to depict a past setting from which Giorgio is diegetically dislocated.

The opening of My Chemical Romance’s “Welcome to the Black Parade” also provides a compelling example of dislocated focalization. The track consists of a diachronic progression between introductory settings before a lengthy synchronic passage which lasts the remainder of the track. It begins as we join the protagonist at the piano as he plays a descending figure and recounts his childhood. This somewhat mimetic phonographic environment invites us to accommodate ourselves alongside the protagonist as he plays. Thus, the phonographic environment affords a location in the time of narration, which we can take as an instance of coherent focalization.

The protagonist continues narrating, detailing his childhood experience of being taken to see a marching band. As he does so, an ensemble of snare drums, a bass drum, marching wind instruments, and glockenspiel enter in time with the piano. By the time the protagonist repeats this at 1'15'', the marching band has entered in full and the persona is now audibly shouting, suppressed within the mix, as if depicting the protagonist lost amongst the parade and unable to make himself heard over the band. We begin, then, with the piano corresponding to the narrator’s present as he begins to establish a story time in his past. As he continues, the phonographic environment dissolves into a recollection of the past to match the story time rather than narrator time; when he mentions the marching band, the phonographic environment, now blended, superimposes the two temporal settings as if calling into
mind the sound of the marching band and offering a perspective from the
protagonist’s inner thoughts. Upon his repeat, the piano has either exited or faded
beyond clear audibility, as the phonographic environment, with the possible exception of the unexpected guitars, and persona now appear to depict the same time setting. In other words, we start with an implied listener located at the piano, before this dissolves into a new scene in which an implied listener is located in the protagonist’s recollected past; from coherent to dislocated focalization.

5.5 Case Study - Eminem and Dr. Dre: “Guilty Conscience”

To further illustrate the role of focalization in recorded popular song, and how it may inform worldmaking and accommodation affordances in conjunction with diegetic framing, I would like to discuss in greater detail Eminem’s “Guilty Conscience”, featuring Dr. Dre. The track was released in 1999 on Eminem’s Slim Shady LP and is of particular importance as the first of several tracks which Eminem performed with Dr. Dre, who was executive producer for the album and producer for various later Eminem releases. Slim Shady LP also led to Eminem’s commercial breakthrough and introduced Eminem’s Slim Shady persona -- originally conceived on the previous Slim Shady EP — to a wide audience.

In a rhetorical analysis of Eminem’s music, Sellnow and Brown (2004) adopt an approach which partially resembles Moore’s persona-environment paradigm by considering the relationship between music and lyrics, which they characterise as incongruent and tragic. This relationship occurs in “Guilty Conscience” due to a

117 In this analysis, I refer to the explicit album version. An alternative clean edit was produced for radio in which the second scene is rewritten and the junctions between scenes are different.
combination of what the pair identify as intensity patterns in the music and tragic patterns in the lyrics. These intensity patterns in music result from unresolved harmony, “irregular” tempos, syncopation, disjunctive melodies, short gestures, and ‘competing sounds’ (ibid. 23-24), as they highlight especially in ‘Guilty conscience,’ which “employs vigorous intensity patterns that never seem to resolve until, finally, they are silenced by a series of abrupt gunshots. Curiously, the gunshots come as a welcome relief [...] Musically, then, the madness is over” (ibid.: 24). Meanwhile, tragic lyrics involve “the protagonist coping with his or her fate” (ibid.).

In “Guilty Conscience”, as with the majority of the album, Eminem raps as Slim Shady, as characterised by the use of exaggerated violent and misogynistic themes. “Guilty Conscience” features two further personae: Dr. Dre, and an uncredited session vocalist. The latter is focalized as an omniscient narrator, who provides a contextualising commentary at the start of three successive scenes of which the track comprises. Together, Slim Shady and Dr. Dre perform a morality play of sorts by enacting immoral and moral sides of the protagonist’s internal psyche, respectively. To do so, we could describe these two personae as internally focalized, as they share the given protagonist’s perceptions in each scene and occupy his internal psychology.

I favour a reading of Eminem’s and Dr. Dre’s personae as being paraliptically focalized for the most part, as the protagonist in both cases acts as focalizor but the personae illustrate different cognitive and emotive orientations which Rimmon-Kenan (1983: 77-82) has highlighted as aspects of focalization. Indeed, to this end, each presents one side of the focalizor’s personality and so it would perhaps be hasty to categorise them as internal. Given that the personae are paraliptically focalized and speak from the position of the protagonist’s internal psychology in each scene, they are also not diegetic, but rather represent a subordinate layer, as if acting as ‘inner
voices’ - a detail which is supported by the close proximity of the voices to the listener in the soundbox. At the end of the song, though, Eminem's persona ceases to be focalized as he appears to address Dr. Dre directly.

As informed by the narrator and the use of a diegetic frame within the phonographic environment, the track constructs three discrete situations. In the first scene, the protagonist, Eddie, considers robbing a liquor store. The second scene takes place in the context of a fraternity party as the protagonist, Stan, contemplates intercourse with an underage girl. Finally, Grady comes home to find his wife in bed with another man as Eminem and Dr. Dre debate killing her, before turning on one another, and ultimately agreeing as Grady fires shots and presumably kills his wife and her partner.

5.5.1 Scene 1 (Eddie)

The phonographic environment of the first scene employs a combination of a diegetic frame, musical frame, and non-diegetic frame. The scene begins with the voiceover of a narrator who reports that the frustrated 23-year-old protagonist, Eddie, is considering robbing a liquor store but begins to battle with his conscience as he enters the building. From the phrasing of the narrator, it is clear that the corresponding persona is focalized paraleptically. For one, the narrator begins “meet Eddie”, seeming to directly address the listener. Furthermore, the narrator is able to report Eddie’s feelings and the sudden interference from his conscience. Although the narrator employs simultaneous narration, the way that he appears to directly address the listener and phonographic staging in close proximity in the centre of the stereo field without acoustical processing to suggest correspondence with an actual space both illustrate dislocated focalization of the narrator’s voice.
Alongside the dislocated, non-diegetic voiceover, the protagonist, Eddie, inhabits a frame of paramusical Foley and short vocal samples which sonifies a *coherent*, diegetic frame. Here, the sound of footsteps and inhalation/exhalation afford the narrativisation of a scene in which the protagonist nervously paces before making his choice to rob the store. A bed of quiet noise and brief indiscernible talking to the left of the soundbox some distance from the point of listening also suggests an outside space nearby to a busy road with bystanders passing by. Short vocal samples staged closer in the soundbox also depict Eddie as he wrestles with his decision, saying “can’t take this no longer” and “gotta do this, gotta do this”.

Finally, alongside the narrator’s non-diegetic frame and Eddie’s diegetic frame there is a musical frame which simply consists of high pitched oscillating string chords,\(^{118}\) which lend an ominous mood to the scene. Despite the simplicity of the musical material, it offers a moment of complementation, as it prefigures Eddie’s reckless and damaging possible behaviour and seems to offer a window into Eddie’s internal psychology. Considering this musical frame, the manner in which it supports the mood conveyed by the narrator but is acoustically staged differently to the diegetic frame suggests that it is focalized meta-diegetically.

At 16”, the phonographic environment is reconfigured; the previous diegetic material is mostly removed, the narrator is substituted for Dr. Dre’s persona in the first instance (followed later by Eminem), and the musical frame becomes led by a bass, piano, and sampled drum kit. We subsequently learn that Dr. Dre is voicing the protagonist’s conscience, which urges caution. However, this is disputed by a second persona which enters later, voiced by Eminem, who instead encourages Eddie to rob

\(^{118}\) This string movement appears to be sampled from “The Back Room” from Pino Donaggio’s soundtrack to a 1979 horror film entitled *Tourist Trap*. 
the store. At this point, then, both personae appear to have been focalized paraliptically to each accommodate one facet of Eddie’s moral compass.

Despite the majority of previous diegetic material being omitted from this passage, Eddie can be heard briefly seeming to ask who the character voiced by Dre is. This small detail demonstrates that whilst Dre’s vocals are staged in a different acoustical space than Eddie’s voice (which is staged with some room ambience), Dre is audible to Eddie, reinforcing the meta-diegetic phonographic focalization of Dre and Eminem. Accordingly, we may also think of the remainder of the musical frame as being meta-diegetically focalized, to offer some reflection of the protagonist’s psyche as the musical environment follows a straightforward oscillating bass pattern which pivots around A and B which lends a forbidding mood and sense of perpetuity as Eddie makes his choice. Curiously, the scene concludes as Eddie agrees with his good conscience to not rob the store. He can even be heard shouting in agreement. A brief exchange between Eminem and Dre also follows, in which Eminem explicitly addresses the good conscience as Dr. Dre in the line, disagreeing with his outlook.

5.5.2 Scene 2 (Stan)

Following the scene involving Eddie, the music is interrupted by the sound of noise, acting as an episodic marker which suggests a schism in the diegesis to a different place/time, as if switching between preset channels on a radio. This association with the burst of noise implicates the presence of a dislocated implied listener, though a new scene is quickly introduced involving a different protagonist: a 21-year-old man named Stan. As with Eddie’s, the phonographic environment of Stan’s scene begins with various paramusical material, including the sound of jubilant party guests and, closer to the point of listening, a male voice encouraging his uncertain female addressee to join him as he opens the door of a room, reassures her, and the pair kiss.
Once again, musical material can also be heard, although here it is staged using a bandpass filter and reverb, mixed into the distance of the soundbox to suggest that it is coherently focalized.

The narrator from the previous scene returns mid-way through the passage of Foley, again staged in close proximity and without the ambience applied to the diegetic elements, here constituting a non-diegetic frame. The narrator introduces Stan and reports that he is attempting to seduce his addressee, but that she is underage. As the narrator speaks with access to information about both agents and speaks assuredly of the events as they occur, the narrator is once again shown to be focalized paraleptically.

As before, the phonographic environment is reconfigured following the narrated passage: the piano, guitar, bass, and drums return, alongside the two personae. Dre and Eminem are also focalized paraliptically again, as Eminem voices Stan’s antagonistic conscience and, using particularly crude language, tells Stan to spike the woman’s drink and have intercourse until she loses consciousness. Dr. Dre reprises his role as the protagonist’s righteous moral compass and, as before, alternates with Eminem. This time, however, in the diegetic frame, Stan’s addressee can be heard sharply inhaling with some discomfort, illustrating that the protagonist followed Eminem’s advice.

5.5.3 Scene 3 (Grady)

Between the second and final scenes, a burst of noise at 2′02” once again indicates a schism between place and/or times. In this third scene, the point of listening moves abruptly from the Stan’s party into what appears to be an outside space: a diegetic frame consists of the sound of a motor running against the backdrop of crickets. As the motor stops, a car door opens and some brief musical material can also be heard
that stops as the protagonist gets out. As he gets out, a voice also mixed in this diegetic frame remarking on the pleasantness of arriving home, supports the narrativisation of this scene as his arrival in the suburbs after working all day. At this point, the paraleptically-focalized narrator returns one final time to set the scene.

The narrator explains that the subject of this scene is a construction worker named Grady who, upon entering his mobile home discovers his adulterous wife together with a different man. At this point, the diegetic elements exit the soundbox as the musical frame and two personae return, beginning as Dre attempts to calm Grady, interpolated with passages voiced by Eminem which exaggerate the situation and mention several ways in which Grady could kill his wife and lover, despite Dre’s caution to think of his (presumably unborn) baby. In this scene, the antagonism between the two paraleptically-focalized personae escalates but leads to Dre’s eventual relinquishment to Eminem’s case. In the final line, Dre asks where Grady’s gun is located and a diegetic frame is briefly reintroduced as Grady can be heard picking up the gun and firing two shots.

5.5.4 Accommodation

Throughout the track, we can observe several focalization strategies which afford accommodation within the narrative world. Firstly, the very presence of a diegetic frame and indeed, a narrator, provides a wealth of worldmaking cues. In the case of Grady, for instance, coherently focalized paramusical sound is used to afford an evening scene as Grady gets out of his car and walks into his home. Meanwhile, the narrator provides us with additional character cues which the listener may use to form a more specific image of Grady in terms of age and employment, and to support narrativisation through the mention, for instance, of his hard day’s work.
In each scene, the paraliptical focalization of Eminem and Dr. Dre affords deeper accommodation within the psyche of the corresponding characters within the narrative world. Hence, in rapid succession, the track affords accommodation as a dislocated listener, then as a coherent witness within the diegesis, and then within the consciousness of a single character. Interestingly, in the third scene, as the two personae once again antagonise one another, Slim Shady cites events external to the story of the track and addresses his partner as “Mr. Dre”. Here, then, we can observe the use of minimal departure which further affords accommodation within the world for listeners familiar with the events which Shady references.

5.6 Conclusion

In this chapter, I have argued that two processes of focalization can be observed in the case of recorded popular song. The first, which is analogous to the role of focalization in literature as described by Genette and developed by Edmiston, is the filtering of knowledge and perception which a persona accesses relative to some narrative agent. This understanding of focalization underpins the distinction between the persona (which is a component of the phonographic environment) and the protagonist/narrator (the agent which inhabits a narrative world). A dominant position within popular song is for the persona to be focalized in terms of the narrative’s protagonist through internal focalization, whereby the persona accesses the perceptions and experiences of one character. In other circumstances, though, the persona may be focalized externally as a non-diegetic party without access to diegetic characters (this is particularly common in instances of subsequent or prior narration); paraleptically with access to unknowable information; or, paraliptically by constraining the information accessed by the persona to a partial account of what is known.
The second form of focalization is made possible by the recorded nature of popular music tracks and again returns to the agents identified in Chapter 2 (see Section 2.1.1) and draws on the issues of ontology explored in Chapter 4 (see Section 4.3). Whereas literature is principally diegetic in the manner in which information is told to the reader, recorded media involves an important mimetic dimension which makes possible the implied listener. In recorded music, the focalization of the implied listener is manifested in the configuration of the soundbox, which affords several possible relationships with the diegesis. Through a development of Link's observation surrounding the implication that we listen to another listener when noise is added to a recording, my approach has considered the relationship between the focalization of the implied listener in relation to the diegetic locus of the protagonist/narrator.

The first relationship which I have proposed, *meta-diegetic focalization*, describes a common circumstance in which the implied listener is not focalized in terms of a diegetic or non-diegetic agent, but instead appears to depict the mental interior of the protagonist/narrator. The second relationship introduced in this chapter involves an implied listener which shares the diegetic locus of the protagonist/narrator, i.e. *coherent focalization*. Finally, *dislocated focalization* involves an implied listener which is diegetically dislocated from the protagonist/narrator.

Both of the processes of focalization which I have highlighted offer important worldmaking affordances. Regarding the persona, focalization allows us to consider the subjectivity of the protagonist or narrator and its involvement with the world and the situation being put to us. The focalization of the implied listener, meanwhile, allows us to consider the implied position between us and the world. Similarly, the pair both
afford accommodation within the narrative world. As Moore (2012b: 184-185) points out, songs afford empathy with a particular agent or agents within the narrative, such as the protagonist or addressee, or afford us the position of a neutral onlooker. Part of this process of accommodation, I believe, can be attributed to the perspective afforded to the listener through focalization.
6. Interpretation of the who’s
Quadrophenia

Having outlined theoretical perspectives on narrativisation and aspects of worldmaking over the past five chapters, I would like to take the opportunity in this chapter to illustrate, with the use of a longer case study, an integration of the approaches discussed so far. On the one hand, this more expansive analysis will support my premise that the narratological concepts introduced in this project can be successfully applied to recorded popular song. On the other hand, we will see how the changing perspectivity, verisimilitude, and diegetic/temporal architecture across a concept album reward study by affording sophisticated forms of narrativisation and worldmaking. In this chapter, I will focus on The Who’s Quadrophenia in response to Nicholls’ (2004) earlier discussion of the album in relation to ‘virtual opera’. I would particularly like to challenge Nicholls’ apparent view that a narrative understanding of Quadrophenia requires familiarity with a prose narrative which accompanies the album and that “the actual songs of Quadrophenia say little in narrative (or even conventional dramatic) terms” (ibid.: 144).

As I have previously noted in reference to Nicholls’ (2007) later work, I favour a cognitive understanding in which narrative is understood as a mental construction by the interpreter (Ryan 2004; Fludernik 2009; Herman 2009a) (see Section 1.1.2). Nicholls, by contrast, appears to take narrative as an encoded property of the text, which implies a particular intended ‘correct’ meaning and, with it, a predetermined degree of narrativity. In the account below, I describe a particular narrativisation of the recorded songs of Quadrophenia which, contrary to Nicholls’ assertion above, shows that narrativity can be evoked by the album if one is prepared to interpret it in a narrative manner.
As Nicholls (2004) explains, the album is a monodrama that follows a protagonist named, in the accompanying prose, Jimmy. My interpretation below is grouped into five passages which, importantly, are not wholly chronologically sequenced. In brief, I argue that the opening tracks take place chronologically last in the story, whereas the second passage depicts the earliest part of the story with the protagonist leaving home. In the third passage, the protagonist attempts to build a new life without marked success, until he visits the seafront in the fourth passage in a bid to rediscover his mod roots. In the fifth passage, the protagonist takes a boat whilst under the influence of drugs and alcohol to find some isolation on a rock at sea, where the first grouping of tracks is set as the protagonist reflects on the events which brought him to the rock.

6.1 A Narrative Interpretation of *Quadrophenia*

Before I offer my reading of *Quadrophenia*, it is particularly important to address the approach I will take, particularly with respect to Herman’s (2009a) *situatedness*. As described in Chapter 1 (see Section 1.1), this refers to formal properties of a medium and the communicative context of narrativity. Throughout this section, I will approach the recording as a text in itself, separately from the prose and photographs which accompany the album, rather than pursuing a transmedial narrativisation. I also depart from Nicholls’ reckoning that “so apparently removed are the songs [of *Quadrophenia*] from the dramatic narrative that many include acousmatic sounds (of sea, rain, jukebox, train) in order to provide both narrative contextualization and external contrast to the otherwise internal drama of the lyrics” (2004: 155), which appears to suggest that the contributions of paramusical sound-sources are less valid than the lyrics. Rather, I shall follow the general approach set out in Chapter 1, particularly informed by Lacasse (2006) and Liu-Rosenbaum (2012) (see Section 1.2.1),
which considers the contributions of both the persona and phonographic environment.

Herman’s concept of situatedness extends to the communicative context of the medium, which here is particularly important. The discussion which follows draws on multiple listenings of the album, which itself relies upon the fixity of phonographic recordings. This is necessary in order to identify a range of intra-diegetic connections which Dannenberg (2008) argues informs how easily an interpreter may negotiate a narrative world, and which contribute to the building of internal consistency into a narrative world (see Sections 3.1 and 3.4.1).

Although my interpretation of the album involves a narrativisation which does not match the chronology of the tracks, my discussion below will address each track in the order in which they appear on the album. The first grouping opens the album with two instrumental tracks and the first song, “The Real Me”, which provides our first introduction to the protagonist and several helpful contextualising cues that support preliminary worldmaking. In each of the following sections, a brief overview of each track is provided, with greater focus on diachronic examples, before outlining a narrativisation of the tracks based on these observations.

6.1.1 Opening

“I Am the Sea”

The album opens with prominent paramusical elements, including crashing waves and rain, alongside sparse vocal textures and some tuned musical material: three piano lines, an electric guitar modulated by a talkbox, and a vocal line which whispers the title of the track. Whilst the paramusical texture continues throughout, the piano
and electric guitar soon thin, to be followed by sparse quotations of motifs which are repeated elsewhere but heard in full in four particular tracks:

1. a french horn motif derived from “Helpless Dancer” (58”);

2. the “Is it me” idea from “Doctor Jimmy” (1’12”);

3. the titular call from “Bell Boy” (1’29”);

4. and, the refrain from “Love Reign o’er Me” (1’40”).

At 2’03”, a segue to the next track begins by introducing the vocal refrain from “The Real Me”, which sweeps across the soundbox.

Although there is clear performative agency (see Section 3.4.3) in the realisation of the musical frame and one could argue for a performative role of recordist gesture (after Bennett 2015) in the panning of the piano, there is a greater degree of departure in relation to source-bonding and phonographic staging. For one, the piano lines begin distributed across the soundbox but later pan gradually (as does the electric guitar), effects which of course could not be reproduced outside the studio without some effort. Secondly, the balancing between, say, the whisper and waves indicates considerable departure. In several cases, a delay effect is also applied to create an artificial repetition of vocal material, whilst the final two excerpts (from “Love Reign o’er Me” and “The Real Me”) both pan across the stereo field.

“The Real Me”

“The Real Me” introduces our protagonist for the first time. The song is structured with four verses, each separated by a refrain directed towards different addressees upon each hearing: a doctor, his mother, a girl, and a preacher. The beginning is marked by the entry of a drum kit, which provides a steady 4/4 beat alongside a bass
guitar, electric guitar and, later, vocals. When the persona enters, we learn during the opening two verses that the protagonist had previously sought professional psychological support on several occasions, with limited success. His mother, reportedly, had suggested that the protagonist’s mental distress is typical of his family. Verse 3 draws an unusual connection between gaps between paving slabs and veins, before illustrating a degree of paranoia when the protagonist suspects being watched by others from within their homes. The protagonist here also introduces a woman he previously loved but who now reportedly is ambivalent, if not deliberately distant. In the final verse, the persona returns to subsequent narration as the protagonist reports seeking support from an evangelical preacher.

During the verses, cast in C Aeolian, the electric guitar harmony rests primarily on the tonic, with occasional alteration between VII-i, whereas the refrains introduce brass that suggest an upward movement by moving from i-i7 to iv and then iv with octave doubling of the root. During verse 3, though, the harmonic filler of the electric guitar is omitted until the following chorus, which leads to a marked increase in the pace of harmonic change and a sense of added urgency in verse 4. As a particularly fitting segue into the following track, the final “me” is delayed and replicated multiple times, increasingly overlapping and distributed throughout the stereo field.

“Quadrophenia”

In a nod to “I Am the Sea”, “Quadrophenia” incorporates various themes which preempt later tracks. The track begins with a confident strummed guitar pattern over a pedal F in the piano at a brisk 123bpm. Atop this pedal, the piano follows a non-diatonic pattern of major chords moving downwards by fifths from I-IV-VII-♭III, before moving upwards to IV-V-♭VI-I. The lead electric guitar enters in the right of
the soundbox, followed shortly after (at 13") by the drums and synthesised brass. At this point, the brass reintroduces and develops the horn motif leading to a half-time passage at 40” as the “is it me” motif is played on the guitar over a synth pad. Shortly after, this passage is repeated, transposed down, and begins to segue to a new section, marked by a restatement of the first theme on the french horn.

At 2’12”, a piano pedal returns (now playing the tonic in the key of E Mixolydian) alongside a synthesiser figuration which closely follows the melody of “Helpless Dancer”, from which the horn motif is derived and which also features a piano pedal. At 2’36”, the bass and guitar begin imitating the contour of the synthesiser with a varied rhythm, arriving at an oscillation between \(^5\) and \(^1\) which is interrupted by an increase in momentum at 2’59” offered by a semiquaver-based inverted synth pedal and the reintroduction of the drums.

Following a prolonged segue, a final section begins at approximately 4’02” as the synthesiser fades out and descending broken chords enter on the piano. As the texture grows with the support of the synthesiser and electric guitar, we reach a passage which references the “Love Reign o’er Me” theme. As “Quadrophenia” concludes, the sounds of waves return, followed by a motor which recedes within the soundbox away from the point of listening, which crossfades into the sound of rain, now staged as if heard from an interior space.

Narrativisation

At this early stage, the first three tracks allow us to draw some early impressions of the protagonist, his setting, and his corresponding state of affairs. The opening track, “I Am the Sea”, illustrates Brøvig-Hanssen and Danielsen’s (2013) concept of spatial simultaneity through a blended space that combines pianos and other instrumental sound sources with exterior sounds of the outdoors (see Section 3.4.5). To naturalise
this blend, I will refer to my development of Walther-Hansen’s (2015) position that I have set out in Sections 4.3-4.4, by considering the phonographic environment here in terms of diegetic, musical, and non-diegetic frames.

As I have argued in Chapter 4, mimetic strategies (see Section 3.4) may bond sound-events to sources within the story-now to establish a diegetic frame. Alongside this paramusical material runs a musical frame of instrumental sound-sources throughout the first 48″, in which there are several notable points of phonographic departure, particularly in relation to acoustical space through the balancing and mobile panning of sound-sources. Finally, from 58″, the four quotations from later tracks are all staged in different acoustical spaces, indicating dislocation from both the diegetic and musical frames, to instead constitute a non-diegetic frame.

From this arrangement of sound-sources into three frames, I would like to consider the focalization of the phonographic environment. Following Beck (2009) and Chion (1994), I have suggested that mimetic strategies create a point of listening within the diegesis. With Wishart’s (1986, 1996) concept of landscape — the combination of imagined sound sources — in mind, the combination of waves and rain establishes the first setting within the narrative world and positions the implied listener at a windy seafront. The musical frame, meanwhile, is oneiric in character due to the multiple pianos which move across the stereo field and the balancing between instrumental sound-sources and the amplified whisper. This oneiric musical frame and the presence of a non-diegetic frame which quotes later material together suggest meta-diegetic focalization (see Section 5.4.1) which — to draw from Gorbman’s (1976) understanding of meta-diegetic sounds — sonifies the surreal, perhaps hallucinatory mental interior of the protagonist as he reflects on events
which occur earlier in story-time but are only revealed to the listener later in the tracklisting.\(^{119}\)

As we move to “The Real Me”, it is suggested by the persona that the protagonist exhibits or performs multiple personalities. This detail is complemented particularly by the points of departure in the segue to “Quadrophenia” as multiple samples of “me” are overlaid, creating a surreal landscape whilst also evoking an incoherent acoustical space. Meanwhile, the harmonic filler and general musical environment in “The Real Me”, commensurate with shifts in the mood of the lyrics, indicate meta-diegetic focalization of the musical frame which reinforces the apparent conflict between the protagonist’s self-image and his addressee’s perceptions of him. Throughout the track, the form of lyrical narration continues the previous suggestion of reflecting on past experiences. Using Genette’s (1980) terms, the verses are relayed through subsequent narration, whilst the choruses, in contrast, employ simultaneous narration (see Section 4.1). As Bal’s (2009) and Edmiston’s (1989) theories of focalization remind us, the use of subsequent narration in this case dislocates the protagonist from the diegesis, meaning that the temporal location of the narrator at this point remains unknown. Although the choruses employ simultaneous narration here, the rhetorical phrasing of the questions suggests that the protagonist remains non-diegetic.

Moving to “Quadrophenia”, the incorporation of motifs from later in the album (in the same respective order) in a strophic form gives an impression of summarising the events which come later in the album, as though the musical frame here is sonifying a focalizor reflecting on a sequence of events. The closing mimetic elements of “Quadrophenia” offer a helpful way of making sense of story time, as the

\(^{119}\) Indeed, Lacasse’s (2000) reception test data indicates that delay (as used on several elements in the non-diegetic paramusical frame) evokes associations with the past.
return of the ocean suggests a grouping of the first three tracks into a continuous progression of story time. Meanwhile, the motor driving off here acts as an episodic marker (Tagg 2013) to suggest a departure from this space and so a schism to a different temporal and/or spatial setting within the narrative world.

6.1.2 Leaving Home

“Cut my Hair”

The persona is reintroduced in “Cut my Hair”, which offers some additional insight into the protagonist’s background and life at home. It begins — per the title — with the protagonist contemplating changing his haircut in order to fit in with his peer group, relayed now using simultaneous narration. As he continues in the first verse, he explains that he lives with his parents but it emerges in verse 2 that his relationship with them is strained and that he finds them more overbearing than his classmates’ trendier parents. In this same verse, he reports that his mother had discovered a collection of hidden ‘blues’ and casts doubt over whether he will be allowed to stay in the family home.

In each verse, the musical frame is somewhat subdued. A harmonic layer is offered by the piano which follows a IVma7-V7-iii7-vi-ii7-V-I pattern in C Ionian with descending voice-leading. The piano and voice are joined by sustained monophonic guitar notes, and the bass, which also offers a degree of melodic decoration. Throughout the first verse, only the cymbals are used, but the drums provide a 4/4 explicit beat layer throughout the later verses. In the choruses, the protagonist turns to the subject of fashion by referencing zoot suits in preparation, he reports, for a fight at the beach. Unlike the verses, the musical frame in each chorus is upbeat and positive, cast in C Mixolydian at a brisk tempo, as the drums provide extended tom-
tom fills in place of a repeating beat and cymbal hits to emphasise the rhythm of the vocals. The voice here is sung with greater energy, and the passage is also punctuated by staccato synthesiser chords.

In a middle-8, which begins at 2'30", the protagonist continues to question his identity, beginning by asking why he should change his appearance in order to be respected, but continuing, conversely, by asking why he attempts to find belongingness within his peer group. As with the chorus, the musical frame here is upbeat and set in C Mixolydian. The middle-8 is then repeated and modulates upwards again, this time to D Mixolydian, until shifting to A Harmonic Minor in the final verse as the cadential ending of each verse phrase is changed from Dm7-G-C (iim7-V-I in C Ionian) to Dm7-E7-Am (iim7-V7-i in A Harmonic Minor). Commensurate with this harmonic shift, the melodic span of the voice becomes far narrower, the piano appears softer and recessed within the soundbox, the bass is far more static than previous verses, and the synthesiser returns but now plays broken chords. During the last minute of the track, a diegetic frame is added, featuring a male speaker appearing to read the news of disturbances in Brighton, London, and the South Coast as if from a radio broadcast. The programme highlights rising tensions between gangs of rockers and mods whilst a kettle can also be heard boiling towards the right of the soundbox.

“The Punk and the Godfather”

Although “The Punk and the Godfather” employs the same general instrumentation of previous tracks, the reverberation effect applied to the vocal track here is evocative of a much larger space. Per the title, we are presented with a dialogue between two characters in the track, both voiced by Roger Daltrey. In each of the verses, we are presented with the godfather, whereas the punk can be heard in each chorus.
Through the dialogue between the two characters, the godfather criticises the punk’s notions of grandeur by proclaiming his addressee’s failings, limited autonomy, and purported redundancy. The punk refutes these accusations in the chorus, which describes his loftier impression of self-worth despite being a stuttering “punk in the gutter.” Each chorus is followed by a brief bridge in which the drums take on a more decorative role rather than providing an explicit beat layer alongside picked guitars and brief melodic contributions from the bass. A vocal track is present here, but is processed using a talk box as the lyrics appear to read “my generation”.

Following the second chorus, the bridge leads to a middle-8 in which the vocal track is panned towards the right of the stereo field and later joined by a second, which is panned symmetrically left. The persona at this point resumes voicing the punk, who expresses uncertainty and vulnerability in comparing himself to a clown and denying his ability to teach or lead. The middle-8 leads to the first introduction of a diegetic frame at 3’22” involving the sound of a cheering crowd. This is swiftly followed by a shout of “ok”, staged some distance away in the soundbox, which leads into a repeat of the opening verse. The cheering crowd also return at the end of the track at approximately 4’50” as the ‘my generation’ refrain fades out.

“I’m One”

As we move to “I’m One”, the musical frame takes on a rather more contemplative feel, consisting of a picked acoustic guitar and electric guitar swells as the protagonist reflects upon his life. The track is synchronic and formed of three verses, principally using simultaneous narration, which are each followed by a refrain. In the first verse, the protagonist reflects on annual feelings of insignificance. Here, he also mentions ‘coming down’ and the encroaching feeling of loneliness, but nevertheless, the refrain mirrors the title to mark some form of resolution. Throughout, the lyrics play on three
possible interpretations of ‘one’: firstly, being of one mind; secondly, being one among others; and, thirdly, being “the one,” i.e. some form of leader.

Following the first verse and refrain, drums enter and the electric guitars begin to incorporate a greater number of rhythmic embellishments in their harmonic line, whilst the second verse introduces references to fashionable clothing. However, the drums are omitted from the last verse, which takes on a more solemn character as the protagonist describes his partial success in fulfilling his expectations laid out in verse 2. As we move to the final refrain, the drums are reintroduced as the lyrics are sung with greater support from the singer’s diaphragm, evoking a more confident-sounding delivery.

Narrativisation

The second grouping of tracks in the album illustrate several further instances in which diegetic architecture provides narrativisation cues, although we begin, also, to see ways in which the tracks support the importance of issues such as external schemata, minimal departure, and gaps of indeterminacy (see Sections 3.0-3.1.1). In terms of their contribution to my narrativisation, these tracks offer a window into the protagonist’s relationship with others in the world, including details relating to his life at home and peer group. In this grouping, we can narrativise the first temporal developments of the story, beginning with the protagonist’s involvement in a protest and departure from the family home. Having done so, he attends a concert, which prompts him to reflect on his place in relation to his peers and mod culture.

The phonographic environment and lyrics of “Cut my Hair” together afford the clearer construction of a diegesis than the seaside setting of the three previous tracks. Paramusical materials once again establish a diegetic frame which affords localising the track in terms of a particular geography. In this case, we see the staging
of the rain used as a mimetic strategy that indicates an interior location, whilst the source-bonding of a kettle and radio later confirm that the implied listener here is focalized within the protagonist’s home.

The use of simultaneous narration in “Cut my Hair” indicates an alignment now between the protagonist and the story-now. With this basis, we can begin to narrativise a sequence of events in relation to the narrative world. I would like to draw particular attention to the final verse, in which the protagonist mentions coming home on the first available train, as the radio broadcast reports clashes in Brighton. This is also alluded to in the chorus, in which the protagonist can be heard lusting for a fight on the beach.\textsuperscript{120} Previously, the events preceding the story-now constituted a spot of indeterminacy (Ingarden 1973). These new details, however, allow us to resolve this indeterminacy by narrativising the protagonist’s arrival after attending a fight between mods and rockers in Brighton, as he contemplates leaving the family home.

Both the musical frame and the lyrics offer some further insight concerning the protagonist’s internal psychology. Here, I take the musical frame to be meta-diegetically focalized to distinguish between the protagonist’s state of mind in three particular contexts. Firstly, the verses — which are implicitly set at home — feature no explicit beat layer, the harmonic voice-leading simply descends, and the voice is sung with less power than the choruses which follow. By contrast, the choruses — in which the protagonist describes being under the influence of amphetamines — is remarkably more positive, with a more upbeat rhythmic backing, assured vocal projection, and use of the Mixolydian mode. Thirdly, in the final verse, the musical frame depicts the protagonist’s deteriorating mood as the effects of the drugs wear off, and indicating a growing discontentment with his life at home. Here, these

\textsuperscript{120}In “Bell Boy”, the protagonist also suggests that he has taken part in mod protests under the leadership of a former hero.
musical details are also supported by the constitution of spectral space (Camilleri 2010) which Liu-Rosenbuam (2012) also considers in his narrative approach to interpretation: the verses’ concentration of spectral energy is in lower registers; spectral energy is distributed more evenly in the choruses; and, in the final verse, the limited instrumentation creates gaps in the spectrum, particularly in higher frequencies.

Given the protagonist’s involvement in violence at Brighton, his future at home can be cast in some doubt, as the first verse mentions. Indeed, we can draw a connection here with “The Real Me”, which suggests a particularly problematic relationship between the protagonist and his mother, given her lack of sympathy for his apparent anxieties. “Cut my Hair” confirms this strained relationship by revealing the protagonist’s deep dissatisfaction with his parents, especially when compared to those of his peers.121 In the context of this relationship, the mother’s discovery of the protagonist’s ‘blues’, a slang term used to refer to amphetamines, affords an opportunity for narrativisation.122 I would like to argue here that this discovery

121 “Cut my Hair” provides the basis of several connections to be drawn later which afford the building of internal consistency into the narrative world which can be used to direct later narrativisation. During the chorus, for instance, the protagonist makes references to fashion, prefiguring a similar thinking in “Is It in My Head?”, here in preparation for a beach fight which “Doctor Jimmy” similarly alludes to. During the middle-8, the protagonist also refers to a “dancehall friend”, ahead of the later track “Helpless Dancer”.

122 The chorus also refers to leaping, which references a further slang term for amphetamines: ‘leapers’.
prompts the protagonist to leave home amid escalating tensions between him and his parents.\textsuperscript{123}

The staging of the voice and use of paramusical materials in “The Punk and the Godfather” indicate a change in setting. Here, we can observe several elements which constitute a diegetic frame due to the use of mimetic strategies: at 3′23″, the vocal shout out implies the presence of an audience; following this, the sound of applause enters; and, thirdly, I would like to argue that the additional reverb which stages the voice as though within a large acoustical space suggests that it emanates from the diegesis. The landscape here remains musicians singing and playing instruments, but these mimetic strategies simulate source-bonding with performers and audience at a concert and position the implied listener within a large arena. With this in mind, we can narrativise this track as our protagonist attending a concert.

My position that the protagonist is attending a concert invites us to return to Cone’s (1989) distinction between ‘realistic song’, in which singing takes place within the story and so is audible to other characters, and ‘operatic song’, in which characters do not appear to be aware of singing and so accompaniment is, generally speaking, non-diegetic (see Section 4.4.2). In this case, as the musical frame acts as a function of the story, we can say that it is diegetic. Following the categories developed in Section 5.4, then, the musical frame here is coherently focalized, as it occurs within the story—now where the protagonist is located. This focalization, however, appears to shift in the middle-8 (2′33″) to coincide with a change in phonographic staging and lyrical content. At this point, then, we can observe Altman’s (1989) audio dissolve as the

\textsuperscript{123} This is supported by the events of later tracks and, in particular, “Sea and Sand”, wherein the protagonist claims he ought to have left sooner.
(previously diegetic) voice becomes implicitly meta-diegetic due to a change in phonographic staging.

During the middle-8, the punk makes a sober admission which indicates that the lofty image of himself relayed during the choruses is unreliable. Within the phonographic environment at this point, the change in musical texture alongside the use of a talkbox on the voice and different panning suggests a shift from the concert setting. The point of phonographic departure here, then, can be understood as a shift from coherent to meta-diegetic focalization. This also affords some reflection on the focalization of the persona. Returning to Edmiston’s (1989) revision of Genette’s (1980) focalization theory, we might say that the persona during the choruses is internally focalized, as it accesses the knowledge of a single character and is relayed simultaneously (and thereby does not benefit from subsequent knowledge) (see Section 5.2). However, given the unreliability implied by the punk’s admission in the middle-8, we could more correctly view the choruses as paralitical, as information is seemingly deliberately withheld.

At several moments, tracks illustrate the importance of external schemata — references to external articles with which a listener may be familiar from his/her reference worldview — in narrative worldmaking. These include, for instance, the references to the BBC broadcasters, the city of Brighton, popular (although now outdated) slang terminology, and remarks to particular fashion trends. These external schemata prompt the listener to build verisimilitude into the narrative world and illustrate Ryan’s rule of minimal departure. That is, these references promote the view that the fictional world inhabited by Jimmy is, in fact, a bygone era of our actual world.
The persona’s remarks towards particular fashion trends throughout both “I’m One” and “Cut my Hair”, in particular, suggest that the protagonist has some affinity with mod subculture. This association is strengthened by references to amphetamines and the radio broadcast in “Cut my Hair”. These details also afford narrativisation and so, with this in mind, I would like to suggest that having attended a concert, “I’m One” marks a moment of contemplation for the protagonist in which he responds to the false idols critiqued earlier by resolving to not only be a part of the mod community, but also a leader of his peers whilst the effects of drugs wear off.

6.1.3 A New Life

“The Dirty Jobs”

The third group of tracks opens with an adaptation of the horn motif, now played on a brass synthesiser alongside piano, guitar, and bass. As the vocals enter, the protagonist employs simultaneous narration to describe working as a pig farmer whilst revealing a degree of regret in needing to show more caution in what he talks about. Yet, in the chorus which follows, he suggests, conversely, that he is discriminated against and reveals some concern for growing older. Verse 2 again employs simultaneous narration, although the protagonist moves here to describe working as a bus driver for miners, who, like him, find themselves on hard times as the mines are reportedly closed.

In a middle-8, the protagonist suggests that it is his responsibility to stop letting himself be taken advantage of and, in the final verse which follows, he directs his narration to addressees who, he argues, have abandoned their previous rebellious ways. Here, paramusical reverberant shouts enter the soundbox, followed by the
booming sound one might associate with an explosion or felling of a structure before crossfading into the sound of a brass band.

“Helpless Dancer”

“Helpless Dancer” conspicuously moves to a different acoustical space, beginning with a musical frame in isolation. The track opens with restatements of the horn motif, which also returns later. At 33″, a dialogue begins between two vocal tracks, panned midway either side of the stereo field, whilst the french horn exits the mix. The piano here also revisits the I-iv-VII-I movement which is stated in “Quadrophenia”. Instead of a strophic form used in previous tracks, “Helpless Dancer” consists of ten stanzas which alternate between the two personae, who describe increasing unemployment (and racial discrimination), unexplained disappearances (which the persona associates also with homosexuals), and little independence for working people. The track draws to a close as the horn motif is transposed down in sequence with the piano, leading to the vocal exclamation “you stop dancing!”. At this point, the musical frame is replaced by a sample of The Who’s earlier song “The Kids are Alright”, staged with notable reverberation as though playing in the large space of the dancehall, before a restatement of the ‘Is it me...’ motif, which is delayed and recedes within the soundbox.

“Is It In My Head?”

In contrast to the mood of the previous track, “Is It in My Head?” opens with a more contemplative feel, cast in a moderate tempo and G Major harmonic setting. As with several previous examples, the track is synchronic as the protagonist reflects on his present situation without arriving at a point of conclusion. The persona begins with the observation of a man seemingly free of difficulties but continues by describing
the country as “starved” and how people would mock when passing the protagonist. Between each verse, the title is restated as a refrain as additional vocal recordings enter the soundbox and the protagonist questions whether he is perceiving the world around him accurately. The final verse is particularly telling, because the earlier theme of mental instability returns as the protagonist explains his difficulties in expressing himself and his sense that someone is following him.

“I've Had Enough”

A drum fill at approximately 140bpm denotes the opening of a more purposeful musical passage in “I've Had Enough”. As the musical frame is supplemented with a bass and other guitars, an opening instrumental passage is underpinned by a tonic pedal D. Against this static pedal, parallel electric guitar chords rise from IV-V-♭VI-♭VII. The persona enters with two verses, beginning at 18”, in which the protagonist seems to describe an unqualified addressee’s lack of progress. Following these first verses, a textural change marks a shift to the chorus, in which the lyrics — now using prior narration — look to the protagonist’s fashion aspirations. As he considers his jacket, motorbike, and haircut, we move to A Ionian as the harmony generally rests on the tonic, pivoting occasionally. However, at 1′22”, the metre abruptly shifts from 4/4 to 6/8 as we move into A Aeolian.

In this new passage, a synthesiser enters, playing the arpeggiated harmonic backing heard previously in “Quadrophenia” to accompany a sung refrain from “Love Reign o’er Me”. Following several repetitions, the quotation of “Love Reign o’er Me” leads to the first hearing of a refrain, which is structured as a list of frustrations with opposing pairs (“living” vs. “dying”, “smiling” vs. “crying”, and so on), cast in the original key of D. This leads to a repetition of the structure thus far, with verses 3 and 4 introducing a bolder, ambitious side to the protagonist. The lyrics of the refrain are
also altered when repeated, and diverge from the list of pairs half way through, as the protagonist declares he is tired of mod fashion and emotions, most conspicuously love, which the persona shouts as the musical frame exits. At this point, a diegetic frame enters with the sound of a train passing.

Narrativisation

The third grouping of tracks follow the protagonist as he builds a new life and takes on a string of mundane jobs amid growing cynicism which reaches a peak with a realisation in “I’ve Had Enough”. Having left home, we join the protagonist in “The Dirty Jobs” as he attempts to find a satisfactory job, but takes no pleasure as a pig farmer, bus driver, and so on. During “The Dirty Jobs”, the use of simultaneous narration throughout a succession of different scenes suggests the rapid passing of story-time throughout the track, progressing by weeks, if not months, in the narrative world. From the last verse, we can conclude that the protagonist still aspires to become a mod leader, although he appears to believe his peers are largely hypocritical and not dedicated in the way in which he is. As the track draws to a conclusion, paramusical materials enter whose staging and source-bonding indicate the establishment of a diegetic frame. The diegetic frame appears to depict some form of protest — the protest on the Brighton seafront from the radio report in “Cut my Hair”, perhaps — that the protagonist recalls as he cajoles his addressees. As the story at this point is localised some time after the radio report in “Cut my Hair”, the protest here exemplifies dislocated focalization of the implied listener, as it takes place prior to the narrated story.

Over these three tracks, we witness the protagonist’s growing dissatisfaction with the mod community and wider society. The lyrics of “Helpless Dancer”, for instance, reveal several confrontational ideological points which highlight tensions
in the society of the narrative world and continue the protagonist’s critique from “The Dirty Jobs”. Using multiple personae voiced by the same singer and which complete one-another’s sentences, “Helpless Dancer” once again alludes to the protagonist’s fractured personality (although he is in agreement with himself here). Despite the optimism of “I'm One” just a few tracks ago, his pessimism continues into “Is It in My Head?” as the protagonist appears to suffer a crisis of confidence following the dancehall. In this case particularly, the multiple vocal tracks during the refrain further support the protagonist’s reported schizophrenia.

In Section 3.4.1, I introduced Segal’s (1995) concept of storyworld logic to describe the rules and principals which govern a narrative world. In this case, as the protagonist persists across the album, we encounter a narrative logic from his subjective perspective. Indeed, Ryan (2005) suggests that possible worlds are relayed and organised from the position of an experiencing consciousness. “Helpless Dancer” offers some insight here into the state of affairs for the protagonist and his peers as he sees them by suggesting that employers and society conspire against the working class and various minorities, positioning the protagonist as an underdog with little opportunity to, for instance, find a fulfilling job or love.

Between “Is It in My Head?” and “I've Had Enough”, a gap in story-time is implied as the protagonist returns to his earlier determination. As we might expect (see Moore 2012b: 117), the addressee in “I've Had Enough” is unqualified, which warrants the interpreter to narrativise an addressee (Schinko, Kiefer, and Huck 2011). Here, we encounter a gap of indeterminacy which affords two possibilities of note. On the one hand, we might hear the protagonist addressing his peers and the lack of

---

124 In this thinking, I follow Maus’ (1991) related suggestion in his discussion on the diegetic hypothesis of musical narrativity that breaks between musical movements may suggest temporal schisms.
momentum in the mod movement. On the other, we might understand this passage as rhetorical, with the protagonist talking to himself. In the reading of the album presented here, the latter understanding is most appropriate, as the recordings afford the construction of a narrative world in which the protagonist has left home without familial support and finds himself exhausted by a string of unfulfilling jobs and lack of romance.

By this point in the album, we can observe various intra-diegetic connections, which Dannenberg (2008) suggests afford the construction of belief in a narrative world. The references to fashion in “I’ve Had Enough”, for instance, returns to themes previously raised in “Cut my Hair”. The most pervasive connections, however, occur through musical quotation, as “The Dirty Jobs” and “Helpless Dancer” both employ the horn theme, whilst the latter also draws on harmonic material from “Quadrophenia” and the ‘Is it me...’ motif.125 Similarly, “I’ve Had Enough” prefigures material which is elaborated in “Love Reign o’er Me”. Following my argument in Chapter 3 (see Section 3.4.1), this supports the building of internal consistency within a narrative world, as it emphasises the congruity of the persona across the album.

6.1.4 Visiting The Seafront

“5:15”

“5:15” opens with a male voice staged with the character of a public announcement in a reverberant space alongside footsteps and metallic impacts. The track then restates the lyrics “why should I care?” from “Cut my Hair” before the environment shifts dramatically to an upbeat feel supported by quaver piano chords, brass, the

125 The lyrics provided in the accompanying liner notes also attribute ‘Roger’s theme’ to “Helpless Dancer”.

227
establishment of an explicit beat layer, and an active bassline. In the first verse which follows, additional cues are provided as the protagonist mentions girls, seating, and ushers. The second verse is voiced without backing vocalists and Jimmy here mentions ‘uppers’ and ‘downers’, referring to different forms of drugs. This is emphasised in the chorus that follows, in which the protagonist’s stream of consciousness questions where he has been as he appears to feel he has no home, concluding by suggesting that the protagonist is under the influence of drugs.

Verses 3 and 4 revisit several of the ideas expressed previously in the album, including references to the miners’ quarry of “The Dirty Jobs” and lorry driving along similar lines, as well as mentioning the ballroom covered in “Helpless Dancer”. Following a return to the refrain, an extended playout begins from approximately 3’15”. This leads to a second quotation of the opening from “Cut my Hair” as the track comes to a close.

“Sea and Sand”

“Sea and Sand” begins with the sound of waves and chirping birds. Each of the verses consist of two hypermetric groupings, in which the first lasts eight bars, and the second is extended. In the first and fourth verse, each grouping begins by dwelling on I(sus2) on a picked guitar, joined by the voice and diegetic frame, before drums and electric guitar enter from the fifth bar and move from IV-V-I. The first verse opens in A Ionian, in which the protagonist laments his life at home, with a particular note on his drunken parents, who reportedly threw him out of the house.

As the protagonist continues in verse 2, he introduces a romantic interest and her attention to fashion. Here, the harmony is based in C Ionian and is generally more mobile than the previous verse, and also ends by resting on a D major chord, lending a Lydian flavour. Unlike the previous verse, the drums provide a soft backing
throughout without the additional guitars. Here, the protagonist resolves to adapt to complement her taste and, accordingly, a bridge which follows repeats the lines first heard in “I’ve Had Enough” (returning to A Ionian) in which Jimmy describes his ideal outfit, scooter, and neat haircut.

The third verse opens a new scene set in a dancehall as Jimmy expectantly waits for his addressee to notice him, employing the same texture and harmonic patterning of verse 2. This is followed by a return to A Ionian for a further bridge, in which he asks why others are able to affordably adopt popular fashion trends and expresses frustration with women appearing more sophisticated than reality. This leads to a final verse, mimicking the texture of verse 1, as we return to the setting of the beach and the protagonist asks an addressee to join him to sleep overnight there together, celebrating their togetherness and asserting that he should have left home sooner. From approximately 3’30”, an extended playout modifies previous material whilst the lyrics mostly repeat the phrase “I’m the face if you want it”.

“Drowned”

In “Drowned”, we learn that Jimmy is in search of his past heroes, without success. Nevertheless, the following verses each comment on the restorative effects which Jimmy finds in the water. In each chorus, the protagonist calls to be ‘set free’ by the cold water of the sea, which — with the title in mind and a later clear lyrical suggestion — perhaps prefigures an attempt to drown. Yet, he is surprisingly optimistic and the personic environment, too, is upbeat and positive, led by drums, piano, guitar, and bass, with occasional brass. Following two verses and chorusses, extended instrumental sections are introduced from 2’05”. Later, the brass textures from “5:15” return. This instrumental passage leads to the third verse in which the protagonist explains that his memories of the past are returning, along with the
names of his former peers. A final chorus and extended instrumental playout follow, which conclude at 5’02” as a flanger effect is applied to the ensemble and the instruments drift towards the centre as they fade out. The track then ends with a diegetic frame containing the sounds of children, water, and footsteps on wet gravel which pan from left to right as a man idly sings the opening lines from the final verse of “Sea and Sand”.

Narrativisation

Having grown tired of his life after moving out of the family home, the penultimate grouping of tracks in the album can be narrativised as the protagonist deciding to leave town and visit the seafront in a bid to rediscover his mod roots. This passage particularly offers an opportunity to consider the incorporation of new information, some of which is contradictory, into a narrative world through what Herman (2009a) characterises as the ‘deformation’ of a storyworld.

With the sound of a train crossing the soundbox in “I’ve Had Enough” in mind, the diegetic frame which opens “5:15” affords imagining a train station with the protagonist on board as the guards prepare for departure. As the music begins, the opening quotation from “Cut my Hair” here creates an intra-diegetic connection which implicates some association between the two tracks. Following my reading earlier, this connection can inform a narrativisation in which the protagonist here is returning to Brighton. Whilst on the train, the lyrics inform us that he once again takes amphetamines, perhaps to pass the time.

“Sea and Sand” supports my contention that the protagonist here is visiting Brighton, as it opens with a diegetic paramusical frame of sounds bonded with waves
and chirping birds, positioning an implied listener at a seafront. Having found a companion to join him on the beach overnight, the protagonist appears markedly more content and suggests he should have left home sooner. Here, the meta-diegetic musical frame of the track offers a glimpse into the internal psychology of the protagonist as he attempts to validate himself and suggesting — in contradiction to “Cut my Hair” — that it was his own (purportedly correct) decision to leave home.

“Sea and Sand” exemplifies a departure from the logic of the narrative world thus far established: the protagonist’s apparent claim to have chosen to leave home when it appears more likely that he was thrown out by his parents challenges Lewis’ (1979) position on presupposition as it contradicts previous information. However, this is resolved in a passage which comes later in “Drowned”. At the end of this track, a change to the staging of the vocals indicates a point of audio dissolve between the musical frame and a diegetic paramusical frame as an implied listener is positioned as a third party on the beach as Jimmy passes by. As he passes, Jimmy sings a passage from “Sea and Sand” which reframes this earlier track in a rhetorical light that reinforces the contention above that the musical frame is meta-diegetically focalized, as though the protagonist was imagining inviting his addressee to the beach.

Although the explanation above suggests that the apparent contradiction can be accommodated within the logic of the narrative world, “Drowned” prefigures a deterioration of the protagonist’s mental state over the tracks which follow. On face value, the lyrics suggest that the protagonist wishes to die at sea, although this breaks with the logic of the narrative world established so far and the musical frame here is oppositional, in Moore’s (2012b) persona-environment typology. Accordingly, we can

---

126 Indeed, the persona explains in “Drowned” that the protagonist has come in search of lost heroes.
instead read these references as a metaphor for the sea being — as the protagonist sees it (although perhaps misguidedly) — a site of baptism and renewal.

6.1.5 Seeking Refuge

“Bell Boy”

The protagonist in “Bell Boy” begins by celebrating a new-found freedom, claiming that he is now where he can feel like he’s the only real man in the world. As he continues, he appears to recognise a figure from the “crazy days” of the past and asks this addressee whether they recognise him from 1963, when Jimmy was a devoted follower.

The second verse is voiced by a different persona (here voiced by Keith Moon), which takes on the role of Jimmy's addressee. The addressee explains that he has found employment and begun a new life: he now works in a hotel which the mods once attacked. Yet, in the chorus, he reveals some dissatisfaction with this newfound lifestyle. In a bridge at 2'04", we learn that he occasionally still visits the beach at night, joined by a sullen musical environment, in stark contrast to the upbeat repetition of Jimmy’s first verse. After repeating the earlier material, the addressee returns with a final verse in which he appears to try to convince Jimmy to also find a stable job.

“Doctor Jimmy”

“Doctor Jimmy” begins with several overlaid tracks of resonant filtered noise against a recording of the sea. This diegetic frame continues in isolation for approximately 32". The running idea of the protagonist's fractured personality returns once again here. In this case, a distinction is made between “Doctor Jimmy and Mr. Jim”
throughout each refrain. As the protagonist reports, these personalities are suppressed under the use of pills but exacerbated by alcohol. This also helps to explain the lyrical material before each refrain, in which he confidently abandons his morals and promises to steal, rape, and gamble whilst under the influence of drugs.

Following the second refrain, an extended instrumental bridge (2'44"-3'16") abruptly shifts key, whilst the explicit beat layer is removed as cymbal and drum rolls are used instead in a more decorative fashion. When the vocals reenter at 3'17", joined by piano and brass, they do so in halftime as we hear the vocal “is it me?” idea in a full 6-bar-long version, which in fact asks whether the protagonist is having a (momentary) period of clarity. At the end of this vocal segment, Jimmy says “the past is calling”. In the final verse which soon follows at 4'13", he resolves to return home and confront the “baboon” who previously fought him.

The prechorus returns following the third verse, repeated twice, although this time, each pair of lines is interpolated by four bars of instrumental material featuring a pedal fiddle (heard previously in “The Dirty Jobs”) and staccato synthesiser chords. Following a further instrumental bridge from 5'23"- 5'42", we hear the final statement of the refrain, followed once again by an instrumental passage leading to the ‘Is it me...’ alternative verse.

“The Rock”

“The Rock”, is the last instrumental track of the album. In the opening seconds, a musical frame of synthetic brass and piano are joined by a diegetic paramusical sound

127 The two characters noted here reference Robert Louis Stevenson’s Strange Case of Dr. Jekyll and Mr Hyde, in which the two titular characters refer to the upstanding Henry Jekyll and his sociopathic split personality.
of splashing water crossing the soundbox. In this introductory passage, the tom-toms are staged with considerable reverb which gives the impression of large space. A point of arrival can be narrativised at 54", announced by the episodic marker of drum rolls and termination of the previous harmonic pattern.

As we enter a new scene at 1'03", a synthesiser enters towards the right of the soundbox with a tremolo and delay effect. A guitar enters shortly after, quoting the ‘is it me?’ motif from “Doctor Jimmy”, which is developed before a repetition in a lower register. In the following scene (from 2'44"), material is again quoted, this time from “Helpless Dancer”, which grows increasingly chaotic as the texture thickens each repetition and the spectral space becomes crowded until a descending electric guitar at 4'38" prefigures an elision with a new scene which draws from “Love Reign o'er Me”.

The final scene begins with a synthesiser line adapting the earlier reference to “Helpless Dancer” alongside the earlier piano descending broken chords and vocals processed through a talk box which appear to say “listen to me” (although their intelligibility is limited). From 5'57", the drums reenter as the guitars establish a repeating gesture that unexpectedly terminates at 6'23", the start of the fifteenth bar, with a flanged, reverberant stab. At this moment, the musical environment is punctured and elides with the sound of a clap of thunder and rain which continues into the following track “Love Reign o'er Me”.

“Love Reign o'er Me”

The final track of the album opens with an intimate piano solo against the sound of rain with decoration from tom-tom fills and cymbals. As the piano falls, the rain fades out and the piano begins to play soft triplet descending broken chords later joined by the familiar synthesiser arpeggios (both as with “The Rock” and previously also
featured in “Quadrophenia”, although now as triplets) to establish a key of $E_{♭}$Aeolian. As the vocals enter, the protagonist celebrates the restorative power of love, incorporating several references to the coast, leading to the refrain “Love, reign o’er me” which is referenced in several previous tracks. During this refrain, the piano and electric guitar play an extended line falling by step from $^{2}$ to $^{\frac{1}{2}}$ across two octaves. A second verse follows in a similar vein, accompanied once again by the refrain, which leads to a new verse, now in the relative major and 5-bar hypermetre.

At 3’57”, the track returns to the previous texture of the verse and harmonic setting of $E_{♭}$Aeolian, here with the melody played on electric guitar without the voice until the refrain. The album reaches a tumultuous conclusion when, at 4’58”, with a final wailing “love”, tremolos are played on all the instrumental textures until a final sustained tonic chord, with added support from brass, distortion on the electric guitar, and metal clashing in the left of the stereo field.

Narrativisation

In this final grouping, our protagonist encounters a past hero from his mod days, who now, surprisingly, works as a hotel’s bell boy. Faced with this betrayal of mod values, the protagonist takes to sea in search of refuge and solace on an island where he contemplates his life amidst worsening weather. Throughout this grouping of tracks, the persona and phonographic environment evidence the protagonist’s wilful isolation. “Bell Boy”, for instance, announces this lyrically, whereas “Doctor Jimmy” opens with a lengthy passage of simulated wind, which evokes connotations of solitude.

“Bell Boy” incorporates the only direct mention of a particular time in the album: the year 1963. This raises two notable consequences: firstly, as subsequent narration is used here, it implies that the story at this point is temporally located some
time after 1963; secondly, the news report in “Cut my Hair” refers to fighting between mods and rockers on the Brighton seafront which happened in real life the following year, again suggesting minimal departure. From what we know of Jimmy, we may infer from this latter point that the addressee in “Bell Boy” is a former mod hero with whom he was once associated and presumably fought alongside, yet who now appears to have taken an entry-level job in a hotel. The addressee here becomes a further example of the compromised mod community in this narrative world, and his attempt to persuade Jimmy to follow this new stable lifestyle reflects the loss of his former values. Indeed, we know from the line “the past is calling” in “Doctor Jimmy” that the protagonist is reflecting on his former days, and that he feels an urge to return to this lifestyle rather than find a mundane job away from the seaside.

“The Rock” evokes one of the most crucial concluding events in the protagonist’s storyline. The opening of track can be narrativised as a small motorboat — with the protagonist presumably driving — crossing the sea. With the title in mind, we can infer that he is aiming here for a small island, though as “Doctor Jimmy” tells us, he does so whilst under the influence of alcohol. This setting is enhanced by the use of the tremolo and delay that give the opening synthesiser a shimmering quality, a visual anaphone (see Tagg 2013) of the sun breaching clouds over a rippling sea. As with Liu-Rosenbaum’s (2012) approach to narrative musical interpretation, we may narrativise the guitar mimetically here as a representation of the protagonist, with the return to the ‘is it me...’ motif indicating self-reflection now that he resides in the isolated surroundings of the rock. However, the repetition and development suggest a degree of uncertainty: the motif here is contextualised harmonically in B♭ Mixolydian (previously heard in C Aeolian) and, unlike its previous undulating contour, the melody here falls to arrive awkwardly on the major V7, where previously
it rested on $^5$, making a III+ chord here where previously it was a more consonant V(sus4).

In the closing section of “The Rock”, we find in the talkbox line a moment of resolution and determination, yet the abrupt shift coinciding with the clap of thunder moves us from this subjective, meta-diegetic focalization of the protagonist’s mental interior, to the more ‘objective’ coherent focalization that suggests Jimmy is, in fact, on the rock in the rain, leaving the fate of the protagonist unknown. Indeed, the closing track, “Love, Reign O’er Me”, is entirely synchronic and offers some insight into the protagonist’s thoughts but no further event-sequencing. Following Grossberg’s (1992) important link between affect and the phonographic voice, the strained vocal projection makes the protagonist’s desperation clear in the refrains of “Love Reign o’er Me”, particularly as the album comes to a close with Jimmy presumably alone and far from both land and home.

6.2 The World of Quadrophenia

With the above discussion in mind, I would like to reflect on the architecture of the narrative world afforded by the album. Nicholls classifies Quadrophenia as an example of monodrama, meaning that the songs are relayed from the experiences and position of a single protagonist. As such, the tracks offer various information which affords imagining Jimmy. We learn, for instance, about his views on society; his tastes in fashion, music, and women; the jobs he attempts; and, his inter-personal relationships. If we adopt a logic which minimally departs from an exterior worldview — as the tracks encourage the listener to do — we might also infer aspects such as his age from knowing his situation at home and recreational interests.
In Chapter 4, I noted Blacking’s (1974) suggestion that music instigates a ‘virtual time’ (see also Frith 2002; Elliott 2015) (see Section 4.1). Following Negus’ (2012a) extension to this area to address narrative time, I would like to highlight the way in which the recording temporally localises the narrative world. The excerpt of the BBC light programme in “Cut my Hair”, which also references Brighton and the Grand Brighton Hotel, affords situating the world in relation to the United Kingdom during the 1960s, at which time mod subculture had first emerged. Indeed, the broadcast directly references the clashes which occurred at Brighton in 1964 between mods and rockers. In other words, the track affords the construction of a possible world which coheres, for the most part with an actual context, affording some degree of accommodation, particularly for those for whom Jimmy’s perspective aligns with his/her own exterior worldview. Whilst this affords temporally localising the world around the mid-1960s, at least initially, the protagonist’s multiple failed jobs in “The Dirty Jobs” implies a time span of some months or perhaps years. Indeed, the protagonist’s phrasing in “Bell Boy”, which refers to “’63”, suggests that the narrative world encapsulates several years.

In Chapter 3, I raised Goodman’s (1978) argument that worlds are constructed from other worlds, supporting Ryan’s (1991, 2005) related contention for minimal departure. With this in mind, we can argue that Quadrophenia encourages the construction of a narrative world which minimally departs from an exterior worldview through the inclusion of identifiable actual events, subcultures, and geography which all call upon the interpreter to incorporate external schemata within the narrative world. Together, these points of intersection reflect Fludernik’s (2009: 55) suggestion that realist fiction draws from aspects familiar to its reader,

---

128 The BBC Light Programme was broadcast from 1945-1967 (BBC: 2017.).
which Busselle and Bilandzic (2008) and Green (2004) indicate may afford greater degrees of transportation.

With reference to Dannenberg’s (2008) concern for intra-diegetic connections as a means of enhancing believability, we might also consider the various links between songs which support the construction of belief and, accordingly, accommodation. In the first place, interpreting each of the songs (other than “The Punk and the Godfather”) as sharing a protagonist involves the construction of a single narrative world. These intra-diegetic connections also include the various references to the themes of the beach and water, dance halls, fashion, and so on. “Doctor Jimmy”, for instance, refers to the beach fight mentioned in “Cut my Hair”, which references the dance of “Helpless Dancer” too. Similarly, the news report in “Cut my Hair” cites beach fighting between mods and rockers, as depicted in the diegetic frame from 3’27” in “The Dirty Jobs”, and increasing security at train stations before the protagonist catches one on his way to the beach in “5:15”. Such links suggest congruity and so afford the supplementation of a narrative world using information from multiple tracks.

I have also highlighted how correspondences between the information in a recording and actual places, agents, or events afford building verisimilitude into a narrative world. In the case of Quadrophenia, tracks reference actual places, such as the Brighton seafront and the Grand Hotel, and relatable spaces such as the train station and seaside fair. Particularly interesting here is that the band also include intertextual references to “My Generation” in “The Punk and the Godfather” and “The Kids are Alright” in “Helpless Dancer”, which encourage further connectivity between a reference worldview and the narrative world for those familiar with the band’s earlier work. However, this also emphasises a disconnection between the real singer and protagonist, and with it, the fictional nature of the story.
Other links can be found in such details as the incorporation of motivic material throughout the album leading to their fuller exposition later. This in particular raises the question of the narrator’s perspectivity. For the most part, the persona is focalized internally to voice the point of view of the protagonist. However, the references in the phonographic environment to events or material outside of linear story time afford further discussion. This could be interpreted in two particular ways which I would like to highlight: firstly, if taken to be meta-diegetically focalized, the phonographic environment here could offer a window into the fractured psyche of the protagonist; secondly, as I take further below, we could hear these moments as dislocated focalization, suggesting that the narrator is non-diegetic and that the events depicted throughout the album are not being experienced but, rather, recalled.

6.2.1 Narrative Summary

I would like to propose an interpretation of the album in which the tracks describe a chronological sequence of events, as laid out above, with one important caveat. Despite “I Am the Sea”, “The Real Me”, and “Quadrophenia” appearing as the first tracks on the album, they are perhaps best understood as chronologically last in the story, with Jimmy looking back over his path to the rock (and his choice ahead). Indeed, this can explain the apparent incoherent schism between “Quadrophenia” and “Cut my Hair”.

In this reading, beginning from “Cut my Hair”, Jimmy leaves the family home after his mother finds amphetamines. Later, in “The Punk and the Godfather”, he attends a concert, where he confronts the superficiality of the mod subculture he admires and in “I’m One” contemplates his position in relation to his peers. “The

129 This reading is given further weight, perhaps, as the titular track coincides with a point of narrative emphasis: the end of the story.
Dirty Jobs” describes his various attempts to find a stable job in vain. Upon visiting a dancehall and reflecting upon his lifestyle in “Helpless Dancer”, he finds a moment of clarity which develops through “Is It in My Head?” until committing to change in “I’ve Had Enough”. Jimmy takes the train to the seaside in “5:15” in search — according to “Drowned” — of the prominent figures of the past. Upon meeting a former celebrated leader in “Bellboy”, however, he finds his hero has taken the job of a hotel lobby boy. Jimmy turns to alcohol in “Don’t They Know It’s Christmas?” and takes a boat to sea amidst a storm, finding land and screaming his sorrows to the world in “Love Reign o’er Me”.

At this point, we find one final opportunity for narrativisation in order to infer what happens next. Jimmy is in somewhat dangerous surroundings by the time we come to “I Am the Sea”: the sea is audibly treacherous, he is drunk, and so his safe return to land is uncertain. In the context of his dissatisfaction with his life and peers, and his altered mental state, we could propose that the plunge at 7” of “I Am the Sea” depicts Jimmy diving — to his eventual death — into the ocean and “The Real Me” as the protagonist’s response to those who failed him, his question intended as an epitaph to be interpreted as ‘now can you see the real me?’.

6.2.2 The Prose Narrative

Having outlined a narrative reading of the musical album, I would like to end this chapter by considering briefly how this compares to the short story provided in the liner notes which accompany physical releases of the album. Also relayed from the perspective of Jimmy, the prose narrative provides some supplementary detail, confirming that “The Punk and the Godfather” is intended to depict a concert by The Who themselves, as well as additional plot points, such as Jimmy destroying his scooter whilst hungover from amphetamines before he catches the 5:15 train. Coming to the end of the storyline, the prose narrative also offers the potentially important
detail that Jimmy's boat — which is apparently stolen — is not securely anchored, and so drifts off from the rock, leaving him stranded.

In general, the prose narrative supports the sequence of events described in the reading of the album above. In short, it describes Jimmy leaving home amidst tension from attending fights in Brighton, only to sleep on the streets. After attending a Who concert, he finds a job but soon decides to return by train to his spiritual home in Brighton. In Brighton, he encounters Ace Face, a former leading mod, who now works in a hotel, and decides whilst intoxicated to steal a boat to drive to a rock out at sea. This leads to a conclusion of the prose narrative that employs subsequent narration and, as with the reading above, frames the narrating as a recollection of events leading to Jimmy finding himself on the rock.

In some areas of the story, the tracks provide details not mentioned in the prose narrative. We are told, for instance, that Jimmy’s hero ended up working at the hotel that he had previously vandalised; only the recordings make it explicit that Jimmy’s hero had become a bellboy. Furthermore, the tracks devote greater time to passages of the storyline, such as the multiple jobs which Jimmy pursues in “The Dirty Jobs”, and highlights broader aspects of the narrative world such as tensions surrounding ethnicity and sexuality referred to in “Helpless Dancer”.

6.3 Conclusion

In this chapter, I have synthesised the positions explored in previous chapters in service of a narrativisation of Quadrophenia to show that together, the persona and phonographic environment can be readily interpreted in narrative terms. In this sense, the integration of the ideas set out in this project offers ways of finding narrativity in the album that Nicholls appears to discount in favour of the
accompanying prose. To do so, my approach has integrated the concepts discussed in previous chapters to shift the focus from narrativity as a property of the text to a more emergent quality informed by the narrative world imagined in response to the text. In this respect, attention to issues such as focalization, narrative framing, and verisimilitude, alongside lyrics and tonal aspects allow a more holistic understanding of narrativity which accounts not only for the sequencing of events, but also broader aspects which affect narrative readings, including the agents, forces, and settings involved.

In the General Introduction (see Section 0.3), I outlined the three hypotheses which underline this project, namely: songs can have narrativity; songs afford narrative worldmaking; and, the construction of a narrative world in response to a song affords accommodation. Using the analytical approaches developed in this thesis, we can see that a narrative reading is indeed possible through the narrativisation of the tracks, situated within a narrative world constructed in response to the tracks. The use of intra-diegetic connections and external schemata, furthermore, prompt us to furnish our narrative world with familiar articles and build verisimilitude into the narrative world, making it more navigable and affording accommodation within it. Indeed, to revisit Herman’s four basic elements that characterise narratives, using the ideas and concepts raised in previous chapters, I have accounted for the situatedness of the album, in the sense that events are described lyrically and phonographically; narrativised event-sequencing; highlighted narrative worldmaking based on the affordances of the album; and argued that the construction of a narrative world logic and phonographic representation of the persona affords an impression of what it's like.

This extended case study also provides evidence which confirms several arguments made throughout this project. Firstly, the interpretation of the album
through various listenings with the incorporation of external schemata support my contention in Chapter 1 that narrative is a processual, emergent, cognitive phenomenon and that narrative is not simply ‘found’ within a text, but that we instead narrativise the text. Secondly, in Quadrophenia, we can observe the interrelation of three worlds outlined in Chapter 2, narrativised through the combined roles of the six agents that I identified. The transaction between worlds is also supported by the adherence of the album to Ryan’s rule of minimal departure in practice and the use of mimetic strategies to naturalise oneiric elements. Finally, my analyses of the tracks illustrate that applications of narratological concepts, such as focalization, diegetic framing, and different forms of narrative time, can be applied to recorded popular song in the pursuit of new, analytically informed narrative readings.

Whereas Nicholls problematises the use of paramusical sounds, they provide an effective means of evoking different settings within the narrative world in conjunction with the lyrics. Yet, further attention to the role of focalization reveals ways in which the album affords ways of narrativisation and accommodation far differently than the short story in the album’s liner notes. For the most part, paramusical sounds are coherently focalized to afford the imagining of the diegesis and, further, situate an implied listener within this space. Meanwhile, the musical frame is largely meta-diegetic and can be narrativised to offer some insight into the mind of the protagonist. Given that music is temporally mediated, the focalization of the different elements in each track afford to the listener an opportunity to experience a story from different vantage points, whereas the prose narrative is largely narrated from the perspective of a non-diegetic narrator.

The narrativisation of this album draws on various passages, both synchronic and diachronic (see Section 4.1.2). In this respect, I also challenge Nicholls’ reasoning
for his position on the grounds that the songs are largely situational — that is, synchronic — and rely on the use of paramusical materials to denote changes of setting. On a local level of individual songs, we might observe short diachronic sequences, such as “Sea and Sand”, in which the protagonist sets out to attract a girl by attending a dance and inviting her to join him overnight at the beach, or more clearly “The Real Me”, in which the protagonist explains his visits to different confidantes in search of support.\textsuperscript{130} Returning to “Cut my Hair”, for instance, we can see clear event sequencing in Jimmy’s experiences of taking drugs and attempting to fit in, leading to his expulsion from home. Furthermore, with the additional details of “Doctor Jimmy”, which mentions an individual who wounded the protagonist in a fight, we can reimagine the sequence of events to accommodate Jimmy’s presence at the beach fight, which could also contribute to his parents throwing him out.

This example shows that recorded music affords means of engagement specific to the acousmatic format and, with these, opportunities for worldmaking and narrativisation that operate differently to written or told stories. The analytical tools developed in the previous chapters allow us to unlock new ways of hearing the tracks in a narrative way and consider a more sophisticated narrativisation which Nicholls appears to prematurely discount. Given this series of interconnected tracks, the notion of narrative worlds is especially pertinent if we are to interpret the album as a series of connected events. Indeed, being largely situational, the tracks afford various aspects of worldmaking through the details we learn about Jimmy and his relationships with others, his experiences with the mod community, and the settings in which the tracks take place.

\textsuperscript{130} A similar argument is made by Negus (2012b) in relation to Nicholls’ (2007) subsequent article.
The comparison between the album and prose narrative reveals areas of general agreement and some opportunity for conflicting narrativisations, though we must acknowledge a distinction between the two. On the one hand, the prose narrative provides a generally clear and linear form of engagement. On the other, the tracks provide a temporally mediated telling of a story that requires attentive and/or repeated listening in order to draw connections between sequences of events. Although the use of language in the prose narrative also provides greater specificity, the phonographic medium allows for greater degrees of immersion. In this case study, for instance, we have seen such affordances in aspects such as: the apparent emotional labour in voicing the protagonist; the use of mimetic strategies that establish a sonic diegesis; and, offering a window into the mental interior of the protagonist through meta-diegetic focalization.
7. General Conclusion

7.1 Summary

The overarching aim of this project is to develop a basis for the narrative study of recorded popular songs by illustrating the possibility for tracks to have narrativity and by investigating the ways in which tracks afford ways of imagining a narrative world in which stories unfold. In this respect, I have set out to show that it is possible to adapt and apply narratological ideas from outside of musicology in service of new possible analyses and interpretations of recorded popular song. The project provides a starting point to a greater body of possible research which I shall briefly sketch in the following section. First, though, I would like to reflect on the key contributions of the thesis.

In Chapter 1, we explored the compatibility of recorded popular song and narrativity, drawing from discourse in Western art music and narratology. As we have seen, there is considerable debate over whether music in isolation is capable of narrativity or whether one might alternatively interpret it in terms of drama. Furthermore, in the discourse which surrounds Western art music, much of the literature does not address the narrative role of the voice, and tends towards readings that are heavily informed by perceived intentions of the composer without acknowledging the individuality of interpretation. As Ryan (2004) and Herman (2009a) both stress, though, the possibilities of narrativity are informed by the medium through which it is interpreted. Hence, beginning from Nicholls’ model of narrativity in popular music as the interaction of the music or album packaging with a narrative discourse instantiated by lyrics, I have argued that recorded popular song is a distinct medium from Western art music performance in the sense that it is plurimedial. That is, the persona (the lyrical content of a song and aspects pertaining
to their delivery) and phonographic environment (encompassing all other musical and paramusical textures) are both sonically observable in a track and together afford means of narrative interpretation that are distinct from other musical traditions. This marks a development from work which either does not satisfactorily account for the narrative possibilities of recorded music or which suggests that narrativity is principally afforded through language.

The second key contribution offered by Chapter 1 is the proposal of a cognitive understanding of musical narrativity. In place of a conservative approach which understands narrative as an encoded property of a text, this alternative understanding, applying Fludernik’s concept of narrativisation, instead regards narrativity as emerging from a particular mode of listening and interpretation within a narrative structure. In this respect, the idea of narrativisation offers a reception-oriented stance that is not challenged by the likelihood that interpretations will vary between listeners and over time, and duly acknowledges the importance of the listener.

Taking the view that recorded popular song can be narrativised, Chapter 2 develops a world-based model to address the agents involved or implicated in narrativity, and how we can understand the process of forming these worlds. With Chatman’s model of narrative communication as a basis, I proposed a parallel with three roles relating to the singer discussed in existing literature: the real singer, persona, and protagonist/narrator. To these, I discussed three remaining agents in a parallel to Chatman’s model: the addressee, implied listener, and real listener. These six agents lay the basis of the key contribution of Chapter 2, which is to systematise recorded popular song in terms of three worlds attendant on narrativisation. To begin, we can make the straightforward assertion that the real singer and real listener inhabit an actual world, our experiences of which can be used to supplement and
make sense of narratives. As this is encountered subjectively, I refer to this actual world as the listener’s reference worldview. Secondly, the phonographic environment describes the sonic material of a recording which the listener narrativises. Finally, through narrativisation, the listener constructs a narrative world, i.e. a mental representation of the world inhabited by the protagonist/narrator and addressee.

To describe the role of the text in narrative worldmaking, Chapter 2 introduced the concept of affordance. In brief, the theory in this context holds that aspects of the phonographic environment may offer listeners particular ways of narrativisation and constructing a narrative world. In conjunction with the concept of narrativisation, this theoretical basis offers a form of a ‘listener-response’ model which Reyland (n.d.) has argued for in relation to art music, and may be used to address several of Kennett’s (2003) and Negus’ (2012a) concerns relating to the role of the listener in musical interpretation.

Chapter 3 opens a second section of the thesis that sets out to extend the theoretical model outlined in Chapters 1 and 2 into analytical approaches to textually informed aspects of narrative readings, beginning with the topic of verisimilitude and departure. In this chapter, I observed that recorded popular songs tend to exemplify an oneiric or dream-like aesthetic, but I propose that the listener nevertheless builds verisimilitude into a narrative world in response to mimetic strategies. My position follows Ryan’s ‘rule of minimal departure’, which states that interpreters tend to imagine narrative worlds which minimally depart from what I refer to as the reference worldview unless the text explicitly states otherwise. It is important to note here that the brevity of recorded popular song often introduces considerable narrative gaps which afford narrativisation of details or events that are not explicitly stated. To do so requires his/her incorporation of external schemata derived from a reference worldview to furnish or supplement a narrative world.
Having discussed typologies relating to verisimilitude in music, my approach in Chapter 3 considered instead the mimetic strategies which afford building verisimilitude into a narrative world. Regarding the persona, I have proposed mimetic strategies which function through the perceived correspondence between the persona and real singer or a convincing portrayal of the protagonist/narrator. Similarly, we considered the internal coherence of the narrative world and its correspondence to the reference worldview, which related studies in literature and film have suggested inform the degree of belief and transportation. Moving to the phonographic environment, we addressed aspects of source-bonding and landscape, acoustical space, and performative agency which may resemble or depart from the possibilities of a reference worldview. In production that resembles the possibilities of a reference worldview, verisimilitude might suggest a ‘trueness to life’ that supports accommodation within the narrative world.

With the oneiric possibilities of recorded popular song in mind, Chapter 4 addresses the temporal and diegetic architecture of narrative worlds. As literary theorists have demonstrated, important distinctions can be drawn between representations of time on the level of the story, narration, and discourse, though this is not generally applied systematically to popular song. In this chapter, I have used the concept of scenes to describe discrete passages of story content, which may employ prior, simultaneous, or subsequent narration (or some combination), and be either diachronic or synchronic. By “diegetic architecture,” I refer to the perceived ontology of sound-events and narration in relation to the story-now. Whilst recorded popular song is acousmatic, the phonographic environment may afford different diegetic levels through mimetic strategies and particular forms of staging or source-bonding. With this in mind, I have argued for the consideration of recorded popular song in terms of three possible frames: a diegetic paramusical frame, comprised of
elements which are bonded with elements in the narrative world in the story-now; the *musical frame*, of elements which function in terms of Moore’s functional layers and are staged outside of a possible space; and, the *non-diegetic paramusical frame* of elements which do not correspond to Moore’s functional layers or the narrative world in the story-now. Borrowing from Altman’s theory of *audio dissolve*, we have also seen how in the case of popular music, it is possible for these frames to deform across scene boundaries.

In Chapter 5, we explored how tracks position the implied listener and persona in relation to the narrative world through a discussion of perspectivity and the concept of focalization. By modelling the protagonist after a first person narrator, Edmiston’s categories of focalization can be applied to the way a persona takes on a character through withholding information (*paralipsis*); adopting the perceptions and constraints of a diegetic agent (*internal*); accessing perceptions of an agent whilst non-diegetic (*external*); or, knowing information which would be inaccessible to an agent (*paralepsis*). In this chapter, I also developed original categories of focalization of the implied listener in recorded music, which result from the construction of the phonographic environment and the establishment of a point of listening. As we have seen, the phonographic environment might appear to sonify the focalizor’s internal psychology (*meta-diegetic focalization*); portray a perspective audible to an agent with the same diegetic locus as the protagonist/narrator (*coherent focalization*); or depict a scene which is dislocated from the diegetic placement of the protagonist/narrator (*dislocated focalization*).

**7.1.1 So What?**

Before I come to address how this project can be supported by further narratological work, I would like to briefly address what Moore refers to as the ‘so what? test’, “that
is, how well [theoretically informed work] moves from analysis to useful, and usable, interpretation” (2009: 412). In this spirit, we should ask how the theories developed in the past chapters can be employed to valuable effect, and how they provide a practical advantage over piecemeal approaches to the study of narrative aspects in recorded popular song. The first of these issues can be addressed with reference to Chapter 6, in which the example of Quadrophenia allows a synthesis of the positions explored in this project. The chapter shows how attention to the issues addressed in this thesis allows for new interpretations and greater explanatory power of how music may afford ways of narrativisation and worldmaking. Whilst Nicholls has argued, for example, that the songs of the album offer little narrativity, my approach, which considers the role of both the persona and phonographic environment, reveals considerable narrativity afforded through intra-diegetic connections and changing verisimilitude, diegetic/temporal architecture, and perspectivity.

The key contributions of Chapters 1 and 2 are to connect recorded popular song with a cognitive understanding of narrative and show that music can be heard in a narrative way, recognising the importance of listeners in meaning-making. In Chapter 3, my approach to verisimilitude departs from existing strategies by considering aspects which afford the construction of belief rather than categories of phonographic correspondence to an acoustical reality. This shift allows one to address how a recording informs the building of verisimilitude into a narrative world which is necessary if it is to be used as the basis of narrative readings. In Chapter 4, I discuss the role of temporality which has previously been partially applied to recorded popular song, but without exploring the distinction between discourse time, story time, and narrator time, as I have done here, which offers a number of possible relationships not addressed elsewhere. Whilst this has been applied more or less directly from existing narrative theory, I make the connection here with the related
issue of diegetic framing to show how narrators in popular song, and the phonographic environment, may be removed from the story present. Finally, Chapter 5 contributes ways of considering the perspectivity of the persona and the ways in which the phonographic environment positions the implied listener in relation to the narrative world. The integration of these approaches makes it possible to describe in detail the narratological workings of a recorded popular song, and introduces new areas of inquiry surrounding, for instance, the contributions of the phonographic environment.

Returning with *Quadrophenia* as an example, the differences between the narrativity of the prose and tracks illustrate the necessity for a more nuanced, ‘media-conscious narratology’ (Ryan and Thon 2014a) if we are to more comprehensively address the ways in which recorded popular song may have narrativity. In this respect, piecemeal applications of existing narrative theories may offer some support but cannot account for the important connections across topic boundaries and may not be readily applied across repertoire. For this reason, the aim of this project has been to develop approaches as the basis of an integrationist, general strategy of analysis which can be applied to different forms of repertoire and afford new analytically informed interpretations. With this in mind, I would like to consider in the following section how we might pursue such a methodology further.

### 7.2 Future Opportunities in Popular Musicological Narratology

The concept of narrative worlds affords academic study beyond the possibilities of narrativisation by considering further the relationship between an actual world and the possibility of narrative worlds. As Negus and others have emphasised, songs exist and interact with broader cultural and social contexts. As such, the study of narrative
worlds in recorded popular song is particularly relevant to considering how tracks reflect society or culture. This may take the form of commentary on actual events, by illuminating how songwriters envision the inaccessible, such as the future, past, other worlds, or other individuals’ experiences. Alternatively, we might explore the issues this raises in relation to identity, from a queer or feminist perspective, for instance.

In Chapter 3, I have only discussed verisimilitude and the incorporation of external schemata in relation to the reference worldview. Further investigations, however, might explore a related hypothesis in which listeners instantiate a transmedial narrative discourse by negotiating between a song and another narrative world to explore issues of intertextuality or retellings of stories from different media through recorded popular song. This could include the creation of concept albums inspired by other, non-musical works, such as Camel’s The Snow Goose, the Alan Parsons Project’s I Robot, Pink Floyd’s Animals. One could similarly investigate ways in which songs appropriate and retell tales from elsewhere, such as the incorporation of Shakesperian characters in Taylor Swift’s “Love Story” and Radiohead’s “Exit Music (For a Film)” (both incorporating the characters of Rome and Juliet), or references to Orwell’s Nineteen-Eighty-Four in David Bowie’s “1984” and “Big Brother”.

Some further opportunity for study might involve the consideration of different media. I have elected not to investigate external media such as album packaging due to a proliferation of different formats which are packaged in different ways. The work presented here could readily be expanded to accommodate this, perhaps taking a cue from Cook’s (1998a) study of multimedial interactions. There is considerable promise, too, in the narratological study of music video. My focus throughout has been on the construction of an aural narrative world. Whilst cinema studies benefits from a lengthy tradition of scholarly criticism which is of more direct application, perhaps, in this regard, there is scope for work investigating, for instance,
the relationship between the narrative world suggested aurally, and the one seen on screen. Doing so calls particularly for consideration of the ways in which visual cues inform worldmaking and mediate or alter our impression of the diegesis. These extensions, of course, also relate to the study of narrative across media and stand to benefit further studies of how the affordances of musical media inform narrativity and worldmaking.

In addition to these opportunities for further narratological investigation, narrative world theories in recorded popular song invite consideration from other methodological approaches beyond narratological musicology. One particular avenue which may reward study in the context of popular music is the interrogation of worldmaking through creative practice. Although my concern in this project has been musical analysis and not intended as the development of compositional tools, a number of the issues discussed relating to verisimilitude and departure, the (de)formation of a sonic diegesis, and the mediation of perspectivity invites further research through compositional practice.

Finally, my approach in this project has been largely hermeneutic, although several areas may benefit from empirical studies in the form of reception testing. Dibben (2001b) and Lacasse (2000), for instance, each provide instructive examples of investigating the associations made by listeners with sounds and audio effects, respectively, yet related work from a narratological perspective has yet to be undertaken. To this end, reception testing might be used to confirm the hypothesis that perceived verisimilitude affords greater degrees of accommodation (as related tests have shown in other media), or explore worldmaking affordances in greater depth.

***
As Camilleri and Smalley write, “an important goal of analytical exploration is [...] to attempt to reconcile and relate the internal world of the work with the outside world of sonic and non-sonic experience” (1998: 7). To this end, the acknowledgement of narrativity and narrative worldmaking in experiences of recorded popular song invites considerable further attention as a means of illuminating philosophical issues beyond the realm of musicological analysis. In recorded popular song, we may interpret narratives which take different trajectories — and by different means — to literature or other conventional narrative media; which espouse different ideologies and views; and which could provide revealing details about its makers and listeners. The task now is to explore these further.
References


———. 2017. ‘A Re-Encounter with the Scorpions’ “Wind Of Change” Why I Couldn’t Stand It Then—what I Learn from Analysing It Now’. In Perspectives on German


Discography


P!nk. 2006. Centerfold. Spotify. I’m Not Dead. LaFace Records LLC.


Appendix: Glossary of Key Terms

Accommodation

After Herman (2009a): the sensation of taking up residence within a narrative world. Alternatively described as transportation (Gerrig 1993), recentering (Wolf 2004), and deictic shift (Zubin and Hewitt 1995).

Affordance

After Clarke (2005): the mechanism through which the phonographic environment offers interpretative outcomes or ways of interacting with the track.

Authenticity

After Moore (2002)’s ‘first person authenticity’: authentication of a real singer through perceived correspondence with the persona.

Blended space

A term developed by Fauconnier and Turner (2002) to describe the resulting mental space which results from conceptual blends of two or more input spaces. In relation to recorded music, blended spaces are taken in this project to refer to incoherent combinations of landscapes or acoustical spaces which are compositied.

Diachronic

A passage of discourse which consists of temporal development (c.f. synchronic). This requires one or more events.
Departure

1) The construction of a narrative world which departs from the logic of a reference worldview.

2) The conception of a phonographic environment which would not be reproducible without the use of electronic music technology.

Diegesis

1) The telling of a story, as opposed to its enactment (c.f. mimesis).

2) The base level of the story and centre of the narrative world.

Diegetic paramusical frame

Used to describe paramusical aspects of the phonographic environment which appear to emanate from the narrative world in the story-now.

Fabula

After Bal (2009): a chronological ordering of events which, when reformulated through focalization and narration, form the basis of the story.

Focalization

1) The restriction of perceptual information and psychological access available to the narrating agent; a process which forms one aspect of the persona’s adoption of a narrator/protagonist character.

Paralipsis: revealing less than is/was knowable to the narrator in the story-now.
Internal: revealing only information that is/was knowable to the narrator in the story-now.

External: revealing information that was unknowable to the narrator during the story-now, perhaps through the use of corrective statements or the inclusion of knowledge gained subsequently.

Paralepsis (zero): revealing more than could be known to any character; the narrator is omniscient.

2) The organisation of the phonographic environment in a way which suggests correspondence with the perceptions of a narrative agent (whether character-bound or implicit). See also: implied listener.

Meta-diegetic: the phonographic environment principally sonifies the mental interior of a narrative agent.

Coherent: the phonographic environment or given frame depicts a scene which is coherent with the diegetic locus of the protagonist/narrator.

Dislocated: the phonographic environment or given frame depicts a scene which is dislocated from the diegetic locus of the protagonist/narrator.

Implied listener

A concept developed from Booth’s (2006) ‘implied reader’. The implied listener refers to the implicit experiencing consciousness through which we perceive or narrativise the phonographic environment.
Landscape

After Wishart (1986, 1996): the imagined acoustical space and the sound-sources recognised and distributed within it.

Mimesis

Representation or imitation of the actual or possible, conventionally associated with drama, as opposed to narratives (c.f. diegesis). Mimetic strategies form the basis of verisimilitude.

Musical frame

A frame occupied by diegetically-indeterminate aspects of the personic environment, consisting of Moore’s (2012b) four functional layers.

Narrative

The interpreter’s representation of events or circumstances in the medium of consciousness. These events may be directly drawn from the text, or inferred through narrativisation.

Narrativity

A property of a text in which it evokes the interpreter’s imaging of a narrative.

Narrator

A narrative agent adopted by the persona who chiefly relays the experiences of others.
Narration

The telling of a story by the *protagonist/narrator*. After Genette (1980), we can observe four forms of narration:

- *Subsequent*: the narrator reports from after the story-now.
- *Simultaneous*: the narrator reports from within the story-now.
- *Prior*: the narrator reports from before the story-now.
- *Interpolated*: a blend of two or more of the above.

Narrativisation

The process of interpreting a work in terms of a *narrative* structure (see Fludernik 2010).

Narrative world

The reader’s mental model of a world in which a narrative unfolds, encompassing aspects of setting, temporality, and the agents or forces involved (see Herman 2009a on ‘storyworlds’; Ryan 1991 on possible worlds).

Narrator time

The temporal locus from which the narrator narrates. In cases other than *simultaneous narration*, narrator time is decoupled from *story time*.

Non-diegetic paramusical frame

Used to describe paramusical aspects of the *phonographic environment* which do not appear to emanate from the narrative world and/or in the story-now.
Oneiric space

A configuration of the soundbox which is explicit in its studio construction through little correspondence with a real location or performance.

Paramusical

After Tagg (2013): aspects of the phonographic environment which are not tonally or rhythmically incorporated in the personic environment (chiefly Foley or sounds which may appear to be captured unintentionally).

Performative agency

The apparent influence of human performers in a recording (chiefly over rhythmical and tonal aspects of a performance), contrasted with the use of quantisation, sampling, and so on.

Persona

After Moore (2005, 2012b): the mimetic representation of a singer's personality in a recording, concerning aspects of lyrics, timbre, and melodic contour.

Personic environment

After Moore (2005, 2012b): musical discourse in relation to which the persona is typically situated. Phonographic aspects of the personic environment are discussed in relation to the phonographic environment, which also encompasses paramusical elements which are not discussed in Moore’s writing on the personic environment.

Phonographic environment

The sonic discourse of a recording, described in relation to imagined sound-sources and the way in which they are staged. This includes the phonographic representation
of all sound-sources, including the persona, but not tonal or semantic aspects. See also: landscape, soundbox.

Point of listening

The implicit position of the listener from which proximity to sound-sources in the soundbox is judged.

Protagonist

A central narrative character enacted by a first-person persona.

Reference worldview

The actual world, as perceived by the interpreter, and the corresponding learnt schemata used to envisage a narrative world from a recording.

Scene

A passage of discourse in which the spatial and temporal aspects of setting are coherent. Scenes may be synchronic, in the case of explaining a single circumstance, or diachronic, in the case of describing an event or short sequence of events.

Soundbox

After Moore (2012b): a heuristic model of acoustical space in a recording which accounts for stereo panorama, perceived height, perceived depth, and time.

Staging

After Lacasse (2000): the manner in which sound-sources are phonographically represented within the soundbox. This includes aspects of spatialisation, in addition to timbral and dynamics processing.
Story

After Bal (2009): a particular manifestation of the fabula, as perceived by the focalizor and relayed by the narrator.

Story time

The progression of time as it relates to the events or situation relayed by the narrator. Unless simultaneous narration is used, story time is decoupled from narrator time. The term story-now is used to refer to the state of story time at a given moment.

Synchronic

A passage of discourse in which there is no temporal development of the narrative (c.f. diachronic).

Verisimilitude

1) The correspondence of the phonographic environment to an actual possible performance without the assistance of recording technology, especially in terms of landscape, performative agency, and staging.

2) The correspondence of a narrative to possible experience in relation to the reference worldview.