



# University of HUDDERSFIELD

## University of Huddersfield Repository

Jennifer, Smith

Worldbuilding Voices in the Soundscapes of Role Playing Video Games

### Original Citation

Jennifer, Smith (2020) Worldbuilding Voices in the Soundscapes of Role Playing Video Games. Doctoral thesis, University of Huddersfield.

This version is available at <http://eprints.hud.ac.uk/id/eprint/35389/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: [E.mailbox@hud.ac.uk](mailto:E.mailbox@hud.ac.uk).

<http://eprints.hud.ac.uk/>

# Worldbuilding Voices in the Soundscapes of Role-Playing Video Games

Jennifer Caron Smith

A thesis submitted to the University of Huddersfield in partial fulfilment of the requirements  
for the degree of Doctor of Philosophy

The University of Huddersfield

October 2020

## Copyright Statement

- i. The author of this thesis (including any appendices and/ or schedules to this thesis) owns any copyright in it (the “Copyright”) and s/he has given The University of Huddersfield the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.
- ii. Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the University Library. Details of these regulations may be obtained from the Librarian. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.
- iii. The ownership of any patents, designs, trademarks and any and all other intellectual property rights except for the Copyright (the “Intellectual Property Rights”) and any reproductions of copyright works, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

Material from Chapters three, four, five, and six have been published in a special edition of the journal *The Soundtrack* as: Smith, Jennifer. (2020) ‘Vocal disruptions in the aural game world: The female entertainer in The Witcher 3: Wild Hunt, Transistor and Divinity: Original Sin II’, *The Soundtrack*, 11(1–2), pp. 75–97. [https://doi.org/10.1386/ts\\_00006\\_1](https://doi.org/10.1386/ts_00006_1).

Material from Chapter four has been published on the blog site medium.com, April 4, 2019 as: Smith, Jennifer. (2019) ‘A Beautiful Song: The Adaptive Music of NieR: Automata’, Medium, 4 April. Available at: <https://medium.com/the-sound-of-ai/a-beautiful-song-the-adaptive-music-of-nier-automata-4b29e5d7f2f3>.

## Acknowledgements

Writing this thesis has not been without its ups and downs, but it is thanks to the people that have followed me through this process that I have made it to this moment of completion. The encouragement and advice from my supervisory team, Catherine Haworth and Steven Jan, has been especially incredible. I would like to thank them for keeping me on-track throughout the years and always believing in me and my work. The staff and students at the University of Huddersfield, and Huddersfield Students' Union, have also provided me with endless support and given me the opportunities to improve my teaching and academic skills that I value greatly.

The support I have received from the UK ludomusicology community has been immense across my career so far. Thank you to Tim Summers, Michiel Kamp, Mark Sweeney, and Melanie Fritsch for listening to my ideas and always providing me opportunities to have my research heard. Thank you also to my wonderful friends in the academic research community, without whom I would not have been so frequently inspired.

None of this would have been remotely possible without the emotional and financial aid of my parents Caron and Peter Smith. Neither would I have survived this process without my good friend Josh Dibble, and my partner Chris Knight who have both provided me with endless support, love, and advice without complaint.

To my Grandma who would have loved to have seen me complete my PhD.

## Abstract

Role-playing video games are designed to immerse their players in virtual worlds. In order to create these worlds, certain worldbuilding processes need to be followed to create a coherent, consistent, and organic space that feels authentic for the player. This thesis explores the role of voices within soundscapes that engage with these worldbuilding processes. Voices are examined in various case studies to explore their implementation and effectiveness when building an authentic game world in the, often lore-heavy, worlds of role-playing video games. I propose that the worldbuilding processes which incorporate voices take place across two fundamental areas: environmental worldbuilding, and worldbuilding characterisations. Environmental worldbuilding is the building of features such as climate, terrain, and location which integrate people and culture. Worldbuilding characterisations are the formation of character-based identifiers that the player can engage with, possibly leading to a reflection of self between player and character.

To involve the player in these complex environments, the game world must effectively 'build' itself through a combination of gameplay, visuals, and soundscape. An adaptive soundscape can build this organic game world by engaging with the player's actions, and the agency of their playable character. I identify how voices are implemented within the soundscape to engage with the player's physical and narrative position in a game world. I demonstrate how voices highlight player actions in order to immerse and engage them within an organic-seeming role-playing game-world, through techniques including the identification of voice as language-based-meaning, and vocalisations that act as communicators of emotion.

Alongside several shorter case studies, the following role-playing games are analysed in detail:

*NieR: Automata* (Dev. PlatinumGames, 2017, Comp. Keiichi Okabe)

*Transistor* (Dev. Supergiant Games, 2014, Comp. Darren Korb)

*Divinity: Original Sin II* (Dev. Larian Studios, 2017, Comp. Borislav Slavov)

*Final Fantasy XV* (Dev. Square Enix Business Division 2, 2016, Comp. Yoko Shimomura)

*The Witcher 3: Wild Hunt* (Dev. CD Projekt Red, 2015, Comp. Marcin Przybyłowicz)

## Contents

<b>Copyright Statement</b> .....	<b>1</b>
<b>Acknowledgements</b> .....	<b>2</b>
<b>Abstract</b> .....	<b>3</b>
<b>Contents</b> .....	<b>4</b>
<b>List of Tables</b> .....	6
<b>List of Figures</b> .....	6
<b>List of Examples</b> .....	7
<b>Abbreviations</b> .....	<b>8</b>
<b>Chapter One: Introduction</b> .....	<b>9</b>
Methodology.....	11
Case Study Scope .....	12
Overview of the Structure.....	13
<b>Chapter Two: Game Music and Sound</b> .....	<b>16</b>
Games: Worldbuilding and Immersion .....	18
Immersion .....	22
Outsourced Worldbuilding.....	24
Diegesis Worldbuilding: Game Interfaces and In-Game Audio .....	29
Interacting with Video Game Worlds .....	35
Player Embodiment.....	40
Modelling Worldbuilding: Environments and Characterisations.....	44
<b>Chapter Three: Voice, Vocalisations, and Voices in Multimedia</b> .....	<b>48</b>
Defining Voices .....	48
Voices in Multimedia .....	55
The Demands of Voice Acting .....	55
Voices as Environment and Characterisations.....	62
Voices in Video Games.....	72
Do Synthesised Voices Sing? .....	73
Recorded Voices.....	76
Player Voices .....	80
<b>Chapter Four: Environmental Worldbuilding</b> .....	<b>82</b>
Spatial Immersion .....	83
Traversing Environments .....	86
Worldbuilding People and Culture .....	90
Gendered Religious Environments.....	95
Adaptive Environments: <i>NieR: Automata</i> .....	100

Voice, Language, and the Adoption of the Human .....	102
Amusement Park and the Mad Songstress .....	105
<b>Chapter Five: Characterisations and Identifications .....</b>	<b>111</b>
The Process of Identification .....	111
Reflections of Self .....	113
Player Attributes: Representations of Identity .....	120
Issues of Characterisation and Representation .....	122
Muteness and Madness: In Transistor and Divinity Original Sin II .....	126
Combat and Comradery: <i>Final Fantasy XV</i> .....	133
Voice Acting and Direction.....	136
Relationship Building: ‘Stand by Me’ .....	138
<b>Chapter Six: The Witcher 3: Wild Hunt .....</b>	<b>147</b>
Environments and Accents .....	150
A Rare Glimpse of The Sun: Velen’s Marshland.....	154
Celticisms and the Other: The Skellige Isles.....	162
Characterisations: Geralt and Himself .....	168
Lack of Geralt’s Voices: ‘Ladies of The Wood’ .....	168
Lack of Geralt’s Voices: The Human Amongst the Inhuman.....	173
The Female Entertainer: Geralt’s Redemption .....	177
<b>Conclusions and Further Research .....</b>	<b>184</b>
Future Directions for my Research .....	187
<b>Bibliography .....</b>	<b>190</b>
<b>List of Text Sources.....</b>	<b>190</b>
<b>List of Video Game Sources.....</b>	<b>197</b>
<b>List of Film and Television Sources .....</b>	<b>201</b>

## List of Tables

Table 1. Examples of John Williams' 'Hedwig's Theme' as core canon, fringe, and metatexts. ....	20
Table 2. <i>Transistor</i> : Narrative, directional, and combat dialogue from the Transistor. ....	128
Table 3. <i>Divinity: Original Sin II</i> : Lyrical meanings in the verse and chorus' of 'Sing for Me'. ....	131
Table 4. <i>Final Fantasy XV</i> : Positive, neutral, and negative comments towards <i>FF XV</i> 's combat system. These comments were made on a GiantBomb forum relating to <i>FF XV</i> . All typos are made by the users. ....	136
Table 5. <i>Final Fantasy XV</i> : Lyrics for 'Apocalypse Noctis'. Translation partly by James Cook and partly gathered through an online fan translation. ....	143
Table 6. <i>Final Fantasy XV</i> : YouTube user comments from the official Florence + The Machine - 'Stand by Me' video. ....	145
Table 7. <i>The Witcher 3</i> : Lyrical Translation of '...Steel for Humans' from Latinised Bulgarian to English. ....	161
Table 8. <i>The Witcher 3</i> : Lyrical Translation of 'Fear a' Bhàta'. The Gaelic lyrics are used in 'The Fields of Ard Skellig'. ....	166

## List of Figures

Figure 1: The model identifies the process in which worldbuilding occurs, eventually influencing or being influenced by the soundscape. ....	45
Figure 2: The 'third space' between character and audience, as opposed to speaker and listener. ...	65
Figure 3. <i>Metal Gear Solid 3</i> : The game's excessively long ladder sequence. Screenshot by Author..	88
Figure 4. <i>Red Dead Redemption 2</i> : Dutch's gramophone on the edge of camp. Screenshot by Author. ....	93
Figure 5. <i>The Talos Principle</i> : From left to right – Roman garden ruins temple, Egyptian temple, medieval Christian monastery, and the Tower. Screenshots by Author. ....	98
Figure 6. <i>NieR: Automata</i> : environments from left to right - City Ruins, Desert, Amusement Park, Forest. Screenshots by Author. ....	101
Figure 7. <i>NieR: Automata</i> : Desert Environment – Apartment Complex – Beginning of the medium cue and vocals. Screenshot by Author. ....	103
Figure 8. <i>NieR: Automata</i> : An unexpected peaceful machine in the final Desert environment. Screenshot by Author. ....	104
Figure 9. <i>NieR: Automata</i> : The costume-clad machines of the amusement park environment that appear to be maintaining the environment's original functionality. Screenshot by Author. ....	105
Figure 10. <i>NieR: Automata</i> : Simone's plea to the game world to become 'more beautiful'. Screenshot by Author. ....	107
Figure 11. <i>NieR: Automata</i> : The retro bullet-hell style minigame presented to the player as they are 'hacked' by Simone's siren song. Screenshot by Author. ....	108
Figure 12. <i>NieR: Automata</i> : Simone's reason for attempting to become beautiful, the machine named Jean-Paul. Screenshot by Author. ....	109
Figure 13. <i>Transistor</i> : Red's lament, Transistor's apology, and the reveal that Red has lost her voice. Screenshot by Author. ....	127



Figure 14. <i>Divinity: Original Sin II</i> : Lohse's ethereal lute and staging. When Lohse stamps her foot in time with the music, light explodes around her. Screenshot by Author.....	130
Figure 15. <i>Final Fantasy XV</i> : The ominous narrative foreshadowing of FF XV's opening, and the introduction of Noctis Lucis Caelum. Screenshot by Author.....	139
Figure 16. <i>Final Fantasy XV</i> : The player's introduction to the characters' group dynamics. Screenshot by Author.....	140
Figure 17. <i>The Witcher 3</i> : Map of the Northern Territory, excluding The Duchy of Toussaint further south. Screenshot by Author.....	152
Figure 18. <i>The Witcher 3</i> : A rare glimpse of the sun in Velen. Screenshot by Author.....	155
Figure 19. <i>The Witcher 3</i> : Ard Skellig's environment. Screenshot by Author.....	163
Figure 20. <i>The Witcher 3</i> : The ladies of the wood's silo residence. The images show the hanging oddities around the portrait and the room, and the tapestry of the witches which is made from human hair. Screenshots by Author.....	169
Figure 21. <i>The Witcher 3</i> : Kaer Morhen, the white fortress. Screenshot by Author.....	172
Figure 22. <i>The Witcher 3</i> : The inside and outskirts of Novigrad. Screenshots by Author.....	174
Figure 23. <i>The Witcher 3</i> : Priscilla's poster as entertainer, Priscilla and her lute during performance, and Geralt looking reminiscent during Yennefer's descriptors. Screenshots by the Author.....	179
Figure 24. <i>Animal Crossing: New Horizons</i> : Reddit user ashmush posted an image of their in-game mock wedding ceremony.....	188

## List of Examples

Example 1. <i>Divinity: Original Sin II</i> : Eight bars of 'Sing for Me', note the repeated patterns in the rhythmic sequences. Transcription by Josh Dibble.....	132
Example 2. <i>Final Fantasy XV</i> : The original arpeggiated 'Prelude' theme. Transcription by Author...	141
Example 3. <i>Final Fantasy XV</i> : 'Prelude' theme in 'Stand by Me' with only its rising arpeggio to begin. Transcription by Josh Dibble.....	142
Example 4. <i>The Witcher 3</i> : Opening musical cells of Velen's outskirts, including a short rhythmic pattern on the Lute, drones and small movements on the strings. Transcription by Author.....	156
Example 5. <i>The Witcher 3</i> : Velen's marimba leitmotif, an adapted rhythmic pattern from the lute. Transcription by Author.....	157
Example 6. <i>The Witcher 3</i> : Graphic Transcription of instrumentation and vocal techniques in 'Silver for Monsters...'. Transcription by Author.....	159
Example 7. <i>The Witcher 3</i> : The opening and first verse of 'The Fields of Ard Skellig'. The score includes the melody and rhythm on the viola and lute, and the solo female voice. Transcription by Author.....	165
Example 8. <i>The Witcher 3</i> : 'The Tree Where We Sat Once'. Transcription by Author.....	167
Example 9. <i>The Witcher 3</i> : Vocal bass, bowed gusli, and contrabass drones and melody of 'Ladies of the Wood'. Transcription by Author.....	170

## Abbreviations

ADR – Automatic dialogue replacement

*D&D – Dungeons and Dragons*

DLC – Downloadable content

DM – Dungeon Master

*FF – Final Fantasy*

FMV – Full motion video

*GoT – Game of Thrones*

JRPG – Japanese role-playing game

LARP – Live action role-play

*LoTR – The Lord of the Rings*

*LoTRO - The Lord of the Rings Online*

NPC – Non-playable-character

PC – Personal computer

RPG – Role-playing game

*WoW – World of Warcraft*

## Chapter One: Introduction

The fifth edition of the table-top role-playing game (RPG) *Dungeons and Dragons (D&D)* is defined by its *Player's Handbook* as 'about storytelling in worlds of swords and sorcery. It shares elements with childhood games of make-believe. Like those games, *D&D* is driven by imagination.'<sup>1</sup> Where the virtual brethren of *D&D*, and other *D&D* like table-top RPGs, do not need to acquiesce to creating worlds of swords and sorcery to be an RPG, the influence of this worldbuilding culture is prominent. Worldbuilding an RPG is heavily influenced by *D&D's* extensive lore and building blocks that allow Dungeon Masters (DMs), and players, to create their own game world that can also be moulded through play.

These building blocks are the essential tools of worldbuilding RPGs. In video games, the scaffolding of worldbuilding, created by developers, incorporates most aspects of the game world for the player to engage with and explore, including: environmental features such as terrain and climate; people and cultures; religions; enemies and monsters; playable-characters. The aim of worldbuilding an RPG world is to provide the player with a coherent structural world that appears larger, and with a greater history, than the player can initially sense; like *D&D*, most virtual RPGs still allow the player to mould story and character outcomes through their decisions in the game world. However, worldbuilding a virtual world has its difficulties when the player has lost certain senses, such as touch, taste, and smell. In a role-playing video game, the player has access to a medium in the game world, a moveable character model that the player can move and use to react to in-game stimuli. The player of a game, unlike the audience of a linear medium, thus has access to a mediated sense of touch as they move a character or avatar throughout a virtual world by the haptics of a controller. Because the player can 'be' in a video game world through this medium, worldbuilding can actively affect the immersion of the player; immersion is the player's engagement in the world without questioning its worldbuilding features and coherency. Player immersion in a video game greatly differs from audience immersion in film and television because of this active interactivity with the game world.

The interactivity of a video game causes the game world, and its worldbuilding components, to be reactive to player movements. Where gameplay, visuals, narrative, and in-game non-playable-characters (NPCs) react to the player's control within the game world, so too must the soundscape. A soundscape incorporates all sound in a world whether that is a virtual film or video game world, or the soundscapes of a 'real-world' space. As a key worldbuilding tool, the soundscape creates

---

<sup>1</sup> Wizards RPG Team, *Player's Handbook* (Uxbridge: Wizards of the Coast, 2014), p. 5.

functional audio to inform realistic, and hyper-realistic sounds such as birdsong and musical cues. As the visuals and lore of a game world are built by developers, the composer of a soundscape must identify both environments and characterisations of the world through sound design and music. A video game soundscape usually needs to identify to the player where they are at any given moment, the consequences of their actions on the player-character, and what is happening in the story. This audio feedback on in-game actions can affect the worldbuilding process for the player. Particularly given the loss or redundancy of certain 'real-world' senses, the player makes use of extra audio-visual information to understand and feel immersed in the virtual world. The significance of this role of the soundscape is mirrored when games are marketed to potential buyers. The tag 'great soundtrack' was added as a category for selling games on the online store of the online gaming platform Steam.<sup>2</sup> A notable soundtrack is an important consideration for some players when purchasing a game however, for some players a 'great soundtrack' or soundscape may not be a priority when buying or playing a game.

Regardless of player priority, creating a coherent and reactive soundscape amplifies the worldbuilding process for the player of an RPG. This coherency is often enhanced by, or relies on, the inclusion of voices. Voices in the soundscape can bring humanism and a sense of normality – or reality – to a virtual world. Voices are common in day-to-day living and are an important communication tool across the world. Where voices are included in RPGs, they can draw on this familiarity with the 'real-world' and thus build a coherent, and potentially more immersive, game world. Voices of the soundscape can also create their own ruleset, occurring as language-based meaning and emotional-based information in dialogue or performance. Performances themselves, whether within the world's diegetic or non-diegetic audio space, are accepted in non-musical worlds because of the voice's natural performativity. The versatility of voices in the soundscape, of musical and non-musical game worlds, act to worldbuild coherent environments and identifiable characters. On the one hand, the voice and its language-based meanings can engage the player within familiar cultures and humanisms of a game world. On the other hand, vocalisations can induce representational emotions through guttural sounds, phonetic singing, vocal cries, and other vocal manipulations that are not explicitly related to language. Together, these voices of the soundscape can effectively worldbuild environmental and character characteristics for the player. The inclusion of voices as markers of familiarity and emotional content may drive this worldbuilding and possibly create an increased sense of immersion for the player.

---

<sup>2</sup> Gabe Newell, 'Steam' (Bellevue: Valve Corporation, 2003).

## Methodology

Video games are subjective. The experience and reception of a game, its soundscape, and thus its voices, can vary between players – and also between analysts and researchers. With voices, for example, player reception may vary due to the historical and technological progression through ‘silent’ games, synthesised voices and sound, and the eventual inclusion of recorded audio and voices (much like the issues faced in film reception between the silent and talkie eras). Because of the relatively quick transition between these stages, nostalgia for older games has become a selling point for gamers interested in ‘retro’ or ‘classic’ aesthetics. Thus, some players may dislike voices in the soundscape because they remember a ‘lack’ of recorded voices fondly.

Although it is imperative to note the subjectivity of video games, it is not the intention of my thesis to provide a definitive model of how voices affect the player’s engagement with RPGs. My research provides a critical and theoretical text-based analysis that is offered from my own informed but ultimately subjective opinion. My analysis is informed by my own experience of gameplay through ‘analytical play’, as well as a range of literature on both video games and voices. My research considers literature on the study of voices throughout various disciplines, including film, vocal studies, and psychology. Due to the nature of the voice, and its potential gendering, feminist ideas of the voice have been consulted, especially in relation to the frequent use of female song, performance, and mutism in games instead of male-centric voices. Empirical study would allow different conclusions to be drawn about broader player perspectives on voices. This is beyond the scope of my research, however, and instead the study focuses on the creation of a framework for discussion of voices within the worldbuilding of video games from a critical and analytical standpoint. My methodology thus lends itself to a legacy of video game studies that considers player engagement and immersion.

Like analysing a linear medium, playing a game is important to understand the context of its soundscape. However, due to the interactivity of games, a casual or analytical playthrough can take considerable time, with some games spanning over one-hundred hours of game time. Thus, when researching a video game, it is important to use sources from outside of the game world diegesis. These out-sources refers to reviews, fan criticisms, soundtrack CD liners, documentaries, and more: anything relevant that exists outside of the game itself. Fan and professional criticism of the gaming industry and its products provides another way to engage with video game culture, and for an analyst of the medium, an engagement with both the academic and non-academic discourses around video games is beneficial. Often, gamers post their opinions on video games in online newspaper articles and online forums, and comment on video platforms such as YouTube to debate

their experiences of a soundscape. These outputs by gamers are unusual insights into how others perceive video games and their relevant soundscapes, often without academically assessing the text beforehand. Although this thesis mostly considers an analysis of worldbuilding in the diegesis, deriving from an examination of notable approaches to the soundscape, I will consult out-sources to answer questions relating to multiple player experiences. I therefore examine both scholarly and non-scholarly conversations that occur about video games. This contextualises my analyses of case studies based on both casual and analytical play, from their beginnings through to their conclusions.

## Case Study Scope

As video games are still a relatively new medium, ludomusicology has been focusing on creating literature that documents video game music history and engaging in discourse surrounding particularly popular video games such as *Final Fantasy VII*.<sup>3</sup> Now that ludomusicology is becoming a more accepted part of the musicological canon, video game music literature is now demonstrating a lens of specificity, with an increasing amount of works that consider music and sound in specific genres or video games.

Now that the groundwork of ludomusicology has been set out, the functionality of video games, and their use of audio, can be considered from specific genres and games. Therefore, this thesis focuses on case studies from the RPG genre. The genre's transmedia roots, and their history as table-top games, provides developers with extensive source material that they must loyally adhere to and worldbuild. The transmedia nature of RPGs sees it split into many sub-genres, such as action-RPGs, adventure-RPGs, first person-RPGs, and so on. Although the RPG can be widely split, this thesis focuses on single-player and third-person views where the player fully controls their playable-characters or avatars as though they are observing their movements from over the shoulder. There are some modern video games which use synthesised voices and instrumentation as aesthetical design decisions. These games will be mentioned, as all styles of voices can contribute to the understanding of worldbuilding, but these will be passing case studies that are considered in relation to the ongoing discussion. These passing contributions may not necessarily fit the scope of the thesis, in relation to genre or other visual, audio, or gameplay factors, but still offer insights into voices and vocality in the RPG.

---

<sup>3</sup> Square, *Final Fantasy VII, San Mateo: Sony Computer Entertainment* (PlayStation, Microsoft Windows, iOS, PlayStation 4, Android, Nintendo Switch, Xbox One, 1997).

My study thus focuses on the recorded voice, which brings a technological scope to the thesis. This decision affects the console generation and release date of the case studies: the main case studies considered are all RPGs released during the eighth console generation, from 2012-2020 (minor case study examples range across the spectrum of video games over time). I chose RPGs as the focus of my thesis because of my casual interest towards the genre. My casual playthroughs gave me an average gamers experience of RPG game worlds and adaptive voice and vocalisations. If the voices continued to be a main part of my decision making and emotional state throughout an RPG, I revisited the game for an analytical playthrough in order to assess its effectiveness as a case study. Thus, within the case studies I have chosen for this thesis, I found that the voices emphasised more than just the game world and followed myself through my gameplay actions and narrative decisions. These voices adapted to my personal engagement with the world, and its inhabitants, in order to induce similar emotions to what I would have experienced when making these decisions in reality. Although most modern video games use recorded voices, due to the lack of technological constraints that inhibited previous console generations, not all games use adaptive voices that follow the player's decisions and actions.

The world-wide development and release of video games means that a game may need to be localised, and thus the original language of the game translated. It is more common in modern video games to allow the player to choose the language they play in. However, different languages may affect the understanding of vocal information given to the player. Depending on the language chosen, whether this is the player's original language or a language they must read subtitles to understand, the player's identification of audio information feedback may differ. Therefore, the case studies I use are from the British or European localisations and have been casually and analytically played in English where possible.

## **Overview of the Structure**

The structure of my thesis is a systematic representation of voices. Chapters two and three focus on the literature of games, RPGs, worldbuilding, soundscapes, voices, and video game music. Chapter two begins with the study of games and worldbuilding, examining both virtual and non-virtual games and infamous role-playing games such as *D&D*. Worldbuilding is outlined and categorised through a holistic review of its position in game studies. The chapter also introduces my theoretical model of worldbuilding and soundscape creation, identifying how my research identifies environments and characterisations as worldbuilding tools. Chapter three defines 'voices'. Voice definitions that are

considered within existing film, psychology, and musicological texts have been examined and reorganised. A distinction between ‘voice’ and ‘vocalisations’ has been made so that their implementation in video games and RPGs are clear. I introduce literature that discusses voice in film and television in this chapter as the case studies, and many other video game soundscapes, use composition methods that have been established in other multimedia. As this chapter begins with highlighting where scholarly work has already studied voices, it ends with a discussion on vocal implementations within video games.

Chapters four and five focus on environmental and characterisation worldbuilding and demonstrate how the voice is necessary within the soundscape to create a convincing world. The conclusions made within these chapters are informed through the literature outlined in Chapters two and three. Both chapters introduce case studies that act as examples of how voices are used to worldbuild environments and characterisations. The case studies in each chapter follows the structure of shorter case studies that highlight voices as particular uses in the game world to a larger case study at the end of the chapter that incorporates voices as a significant worldbuilding tool. Finally, these case studies lead into Chapter six, which identifies one greater study that incorporates voices as both environmental and characterisations, as outlined in Chapters four and five.

Chapter four begins my theoretical analysis of voices as a worldbuilding tool, starting with the worldbuilding of in-game environments and NPC cultures. Chapter four considers the outer layers of the worldbuilding process, creating a coherent and interactive world for the player to engage with. Discussions of wider issues, such as the appropriation of cultures through voices, are presented in this chapter. This is due to the relationship between worldbuilding a game environment and the use of real-world vernacular assets to identify various people and culture. The case study of this chapter, *NieR: Automata*,<sup>4</sup> focuses on the use of voices to identify diverse environments and humanistic qualities. Alongside building aural associations with environments, the chapter also focuses on the *NieR: Automata*'s inclusion of voices which engage with adaptive environments and soundscapes, that identify not only where the player is geographically, but where the player is within the narrative and who is in control of it.

Chapter five focuses less on the building of geographical landscapes and cultures within the game world and concentrates on the worldbuilding of characteristics and player identification. The worldbuilding of identification through voices are shown through the joint case studies, *Transistor*<sup>5</sup>

---

<sup>4</sup> Platinum Games, *NieR: Automata*, Tokyo: Square Enix (PlayStation 4, Microsoft Windows, Xbox One, 2017).

<sup>5</sup> Supergiant Games, *Transistor*, San Francisco: Supergiant Games (Microsoft Windows, Linux, OS X, PlayStation 4, iOS, Nintendo Switch, 2014).



and *Divinity: Original Sin II*,<sup>6</sup> which draw particular attention to issues of mutism, sexuality, and gender. The second case study, *Final Fantasy XV*,<sup>7</sup> considers the inclusion of comradery between multiple characters in order to build player identification with the game world. This case study focuses on how voices, in both the diegetic and non-diegetic sound space, envelope the player in these relationships between in-game characters.

Chapter six introduces the main case study, *The Witcher 3: Wild Hunt*.<sup>8</sup> This chapter focuses on how introducing a combination of voices, when worldbuilding both environments and characterisations, can create an engaging soundscape. The case study is also an example of how various uses of voice and worldbuilding aspects, outlined throughout previous chapters, within a single game can create a successful and iconic video game. This substantial case study leads to a final conclusion that summarises the overall findings of this thesis and contextualises its relevance to future research.

---

<sup>6</sup> Larian Studios, *Divinity: Original Sin II*, Tokyo: Bandai Namco Entertainment (Microsoft Windows, PlayStation 4, Xbox One, macOS, Nintendo Switch, 2017).

<sup>7</sup> Square Enix Business Division 2, *Final Fantasy XV*, Tokyo: Square Enix (PlayStation 4, Xbox One, Microsoft Windows, Google Stadia, 2016).

<sup>8</sup> CD Projekt Red, *The Witcher 3: Wild Hunt*, Warsaw: CD Projekt (Microsoft Windows, PlayStation 4, Xbox One, Nintendo Switch, 2015).

## Chapter Two: Game Music and Sound

Game music and sound as a worldbuilding tool is not unique to role-playing video games. Audio is important to the worldbuilding of role-playing games that are non-virtual, such as tabletop RPGs and live action role-play (LARP). José P. Zagal and Sebastian Detering's edited collection of essays considers the role-playing game through a transmedia approach, considering the complicated implications of the diverse term 'RPG'. Many scholars are uncomfortable with defining a 'game' and 'RPG' as rulesets can vary between games; the 'RPG' exists across subgenres (for example, the single-player RPG, the massively multiplayer online RPG, the action RPG, adventure RPG, turn-based RPG, and so on) and encompasses both virtual and physical forms.<sup>1</sup> This issue of defining the 'RPG', because of its transmedia status, means that game developers and players may consider RPGs from their own subjective standpoint. Zagal and Detering argue that the transmedia nature of RPGs is what defines it, as each iteration of an RPG considers its history. These historic, or at least previous, iterations of RPGs provide guidelines to developers with content suggestions, such as environments, gameplay, rulesets, and characters. The history and development of role-playing games is discussed by Jon Peterson in his book *Playing at the World: A History of Simulating Wars, People and Fantastic Adventures, from Chess to Role-Playing Games*. Peterson's extensive analysis of role-playing games, from their wargaming predecessors to the popular fantasy tabletop RPG *Dungeons and Dragons*, defines the essence of the role-playing game as this investment in skills and levelling up of characters:

All the games contain some notion of personal progression, of characters who potentially improve indefinitely, which secures the investment of players in a particular character and game, and moreover invites a strong identification of the player with the character. These qualities set the new genre of role-playing games apart from its parent, wargaming.<sup>2</sup>

Considering the history of RPGs brings insight into the functional components that may define its modern iterations. When researching the use of silence in horror RPGs to increase immersion, Rebecca Roberts considers the RPG as 'rich in narrative capacity, allowing an intricate story to unravel as the game progresses, demonstrating a diverse, dynamic audio range.'<sup>3</sup> Whereas, Karen

---

<sup>1</sup> *Role-Playing Game Studies: A Transmedia Approach*, ed. by Sebastian Deterding and José P. Zagal (Oxford: Routledge, 2018), p. 19.

<sup>2</sup> Jon Peterson, *Playing at the World: A History of Simulating Wars, People and Fantastic Adventures, from Chess to Role-Playing Games* (San Diego: Unreason Press, 2012), p. 592.

<sup>3</sup> Rebecca Roberts, 'Fear of the Unknown: Music and Sound Design in Psychological Horror Games', in *Music in Video Games: Studying Play*, ed. by K.J Donnelly, William Gibbons, and Neil Lerner (London: Routledge, 2014), p. 143.

Cook describes a typical RPG through its practicalities of the player's control of a character 'through a well-defined game world to complete a series of quests, solve puzzles, or achieve a central goal – usually saving the world (or the kingdom, and/or a particular person, frequently a princess).'<sup>4</sup>

Roberts and Cook both provide a definition for RPGs that centre around puzzle solving, quests, and a rich narrative. However, the multiple styles of RPGs in video games offer this same fluidity in rulesets. There are two broad variations of the video game RPG, those that are developed in the west and those developed in the east. Western and eastern RPGs have differing focuses during gameplay. Western RPGs are more likely to focus on the statistics of characters, customisable by the player, with art-styles and graphics that aim for realism; for example, I outline in Chapter five's case study, *Divinity: Original Sin II*, the ability of the player to completely build the characterisations and combat style of their characters, along with their interactions with the game world. The eastern RPG focuses less on player involvement of statistics, and greater on fully realised characters and the creation of empathy between player and character, with art styles focusing on stylised aesthetics. I consider this empathetic involvement in Chapter four's eastern RPG case study, *NieR: Automata*. This case study focuses on the relationships between the player, characters, and game world environment in an eastern RPG style, focusing on the narrative as opposed to character statistics. Winifred Phillips discusses the differing attractions of the western and eastern RPGs to different types of players:

Both Western and Eastern RPGs draw the Conqueror, who enjoys the challenge of completing all the game's objectives and also tends to prefer Elite and Rock musical styles. But the Western RPG will additionally attract the Manager, who enjoys the puzzle-oriented play, while the Eastern RPG will have added appeal to the Participant, who sympathizes with the characters and enjoys working with them to achieve objectives.<sup>5</sup>

Phillips associations of player type to RPG type supports Zagal and Detering's assertions that the RPG is difficult to define. Not only does the RPG need to be associated with its development origins, western and eastern, but RPGs can vary in gameplay style and aesthetics depending on the influence of previous generations of RPGs, such as *D&D*.

My definition of the RPG is a game that must provide the player with some choice or customisation of skills when the character levels up, including the customisation of equipment to better enhance skills. Regardless of the particular gameplay focus of the RPG, as long as skills can be 'levelled up' by

---

<sup>4</sup> Karen M. Cook, 'The Things I Do for Lust...': Humour and Subversion in The Bard's Tale', in *Music in the Role-Playing Game: Heroes & Harmonies*, ed. by William Gibbons and Steven Reale (New York, London: Routledge, 2020), p. 22.

<sup>5</sup> Winifred Phillips, *A Composer's Guide to Game Music* (London: The MIT Press, 2014), p. 88.

the player then the game is considered an RPG. This definition highlights important similarities between western and eastern-style RPGs, even though their priorities differ. Here, I follow Peterson's definition which foregrounds the mechanism of providing the player with a character that they can level up and improve through skills and abilities. RPGs should also build game worlds with lore, landscapes, and a coherent world for both players and characters to exist. Gordon Calleja defines this RPG worldbuilding in the massively multiplayer online role-playing game (MMORPG) *World of Warcraft (WoW)*.<sup>6</sup> Calleja discusses *WoW* as having a 'depth of geography, history, and mythos[...] enough lore to keep those interested in it occupied, without displaying a sense of the depth and long-term evolution characteristic of secondary worlds in fantasy literature or some pen-and-paper role-playing game settings.'<sup>7</sup> Video game RPGs have thus evolved, as Calleja suggests, from table-top RPGs which thus help to provide a greater scale of worldbuilding without overwhelming the player in too much detail and history.

The following chapter serves as an introduction to sound and music in worldbuilding games and RPGs. Various categories of worldbuilding will be discussed in games before focusing on the role-playing video game specifically. Key terminologies imperative to the discussion of RPGs, sound, and worldbuilding will be discussed. This will include the clarification of sound in the diegesis and what is considered worldbuilding the diegesis. These will not consist of lengthy discussion in order to redefine terminologies, but to clarify my understanding of these terms within the RPG space.

## **Games: Worldbuilding and Immersion**

Worldbuilding can be considered as a vital component of an RPG's game design. The worldbuilding provided by authors, dungeon masters, developers, and players historically throughout LARP, tabletop, and video game RPGs has worked to transport and engage players with their game world. Karen Schrier, Evan Torner, and Jessica Hammer consider worldbuilding in role-playing games to create an 'internally consistent world, either an alternate version of our own or otherwise, that has been imagined for the purposes of fictional storytelling.'<sup>8</sup> The plethora of storytelling tools to create a consistent fictional world is broken down by Schrier, Torner, and Hammer into three: 'core canon', 'fringe' / 'semi-canon', and 'metatexts'. 'Core canon' is the world's lore, created or influenced by developers or original authors and exists outside of fan influences. The 'core canon' is a fictional

---

<sup>6</sup> Blizzard Entertainment, *World of Warcraft, Irvine: Blizzard Entertainment* (Microsoft Windows, macOS, 2004).

<sup>7</sup> Gordon Calleja, *In-Game: From Immersion to Incorporation* (London: The MIT Press, 2011), pp. 190–91.

<sup>8</sup> Karen Schrier, Evan Torner, and Jessica Hammer, 'Worldbuilding in Role-Playing Games', in Deterding and Zagal, *Role-Playing Game Studies*, p. 351.

world that exists in its own reality. ‘Fringe’, or ‘semi-canon’, texts considers additional stories and content that has been created by the fan community. ‘Fringe’ texts are additional stories that exist outside the world, created by fans and, although enjoyed by many, may not be accepted as canon. Finally, ‘metatext’ considers extra-diegetic texts that visualise the world in question, stemming from the ‘core canon’ but existing outside of the fictional world. ‘Metatexts’ include virtual or physical representations of the world, such as maps and event timelines, and focus on expanding the understanding of the world to blurring the boundaries between fiction and reality.<sup>9</sup>

These three components to worldbuilding set up the lore of the world, a space where audiences and players can engage in the foundations of that world. Audio can be a part of each component. Core canon suggests audio as an enhancement of the fictional world; a musical theme such as John Williams’ ‘Hedwig’s Theme’ is a core canon theme which signifies the *Harry Potter* universe after its original use in the film *Harry Potter and the Philosopher’s Stone*.<sup>10</sup> External to the world’s diegesis, ‘fringe’ texts can include audio as player’s can revisit the sound of music of a game with the fan community and can create covers or new pieces of music that they attribute to that world. ‘Metatexts’ use audio as a way of worldbuilding the game world by releasing the audio to be consumed outside of the game, i.e. Original Soundtrack CDs; see Table 1. These various worldbuilding components are especially relevant to the lore-heavy, and transmedia, case studies analysed within the later chapters. I show in Chapter six that the transmedia case study, *The Witcher 3: Wild Hunt*, uses these ‘core canon’, ‘fringe’, and ‘metatext’ components to engage with the preceding literature and succeeding television series that the player, or a fan, can consume outside of the game world.

---

<sup>9</sup> Karen Schrier, Evan Torner, and Jessica Hammer, ‘Worldbuilding in Role-Playing Games’, in Deterding and Zagal, *Role-Playing Game Studies*, pp. 357–59.

<sup>10</sup> Chris Columbus, *Harry Potter and the Philosopher’s Stone*, Burbank: Warner Bros. Pictures (Warner Bros. Pictures, 2001).

Core Canon	Fringe Text	Metatext
'Hedwig's Theme' in the prologue of <i>Harry Potter and the Philosopher's Stone</i> .	Symphonic metal cover of 'Hedwig's Theme'. <sup>11</sup>	Harry Potter and The Sorcerer's Stone CD Soundtrack.
'Hedwig's Theme' in the title sequence of the Harry Potter spin-off film <i>Fantastic Beasts and Where to Find Them</i> . <sup>12</sup>	Classic FM's Article: 'Harry Potter soundtrack: 'Hedwig's Theme' and everything to know about the film franchise's magical score.' <sup>13</sup>	<i>Harry Potter and the Philosopher's Stone</i> video game. <sup>14</sup>

Table 1. Examples of John Williams' 'Hedwig's Theme' as core canon, fringe, and metatexts.

Worldbuilding has colloquial significance within gaming communities and amongst game designers and has various manifestations when communicating with players, including top-down and bottom-up process. Top-down worldbuilding, according to Schrier, Torner, and Hammer, is a process created by the developers and designers who create the game world. Top-down worldbuilding ascribes the world and its landscapes, characters, functionality, and audio before the player enters the world; the player can change little about the world within the top-down process. The bottom-up process allows players to help build the game world through rulesets and prompts set by the designers.<sup>15</sup> This worldbuilding process is seen in games such as *Sid Meier's Civilization* series. *Sid Meier's Civilization VI*<sup>16</sup> (*Civilization VI*) allows the player to build a game world through their creation of a fully functioning civilisation. The player builds up their civilisation from an 'Ancient era' (4000 BC). These civilisations vary considerably between real world countries and leaders from Pericles of Greece to Teddy Roosevelt of America, Mvemba a Nzinga of Kongo, and the mythical Gilgamesh of Babylon. The player can shape the landscape of the world, build cities in locations which do not follow the

<sup>11</sup> Lackey, 'Harry Potter - Hedwig's Theme || Symphonic Metal Cover', *YouTube*, 31 October 2017 <[https://www.youtube.com/watch?v=z7fbcNW\\_2pg](https://www.youtube.com/watch?v=z7fbcNW_2pg)> [accessed 9 June 2020].

<sup>12</sup> David Yates, *Fantastic Beasts and Where to Find Them*, *Warner Bros. Pictures Heyday Films* (Warner Bros. Pictures, 2016).

<sup>13</sup> Maddy Shaw Roberts, 'Harry Potter Soundtrack: 'Hedwig's Theme' and Everything to Know about the Film... - Classic FM', *Classic FM*, 12 March 2020 <<https://www.classicfm.com/discover-music/periods-genres/film-tv/harry-potter-soundtrack-hedwig-theme-john-williams/>> [accessed 9 June 2020].

<sup>14</sup> Griptonite Games, KnowWonder, Argonaut Games, Westlake Entertainment, and Warthog Games Limited, *Harry Potter and the Philosopher's Stone*, *Redwood City: Electronic Arts* (Game Boy Color, Game Boy Advance, Microsoft Windows, PlayStation, Mac OS X, GameCube, PlayStation 2. Xbox, 2001).

<sup>15</sup> Karen Schrier, Evan Torner, and Jessica Hammer, 'Worldbuilding in Role-Playing Games', in Deterding and Zagal, *Role-Playing Game Studies*, p. 355.

<sup>16</sup> Firaxis Games, *Sid Meier's Civilization VI*, *Novato: 2K Games* (Microsoft Windows, macOS, Linux, iOS, Nintendo Switch, PlayStation 4, Xbox One, 2016).

civilisation's placement in reality, and can build wonders that are not related to their civilisation. The player can declare war or peace with civilisations that did not exist in the same time period. Alongside the bottom-up worldbuilding by the player, each civilisation has a representative musical accompaniment which evolves between four eras: 'Ancient', 'Medieval', 'Industrial', 'Atomic'. Karen M. Cook identifies that in *Sid Meier's Civilization IV*<sup>17</sup> 'when engaging in diplomacy with another civilization, the terrain soundtrack stops and a theme unique to that civilization's leader plays in the background.'<sup>18</sup> *Civilization VI* uses representations of traditional national or folk songs from each represented civilisation, performed on instrumentation typical for the era and culture of that civilisation. The player creates the civilisation from the bottom-up, and the games differ wildly from player to player type and the civilisations they choose. Richard Bartle highlights different player engagement types through a multi-dimensional model, and his 2005 model categorises different player types in games, which affects styles of gameplay: 'Achievers', 'Explorers', 'Socializers', and 'Killers'.<sup>19</sup> Bartle further categorises these four types in two 'flavours', two variations of gameplay style from the player type, considering them as 'implicit' and 'explicit' playing styles; 'Achiever' includes an implicit 'opportunist' and an explicit 'planner', 'Explorers' includes an implicit 'hacker' and explicit 'scientist', 'Socializers' include implicit 'friends' and explicit 'networkers', and 'Killers' includes implicit 'griefers' and explicit 'politicians'. The considerable variations of gameplay styles and player types can alter bottom-up worldbuilding can be changed depending on the type of player, but also from the immersion in that play style. Arguably, the bottom-up process allows players to consider and act on their preferred play style. As the player engages with the worldbuilding of a game world, their willingness to engage with a certain style of game may increase. Immersion can be an effect of both top-down and bottom-up worldbuilding. Interactive music, like in *Civilization VI*, follows the player in the game world, and thus their play style; frequent movements to combat music can be more common for the player type 'Killers', whilst a soundscape may be more likely to change environmental sounds for the 'Explorers' player type.

Phillips discusses Bateman and Boon's categorisation of players into similar player types, 'Conquerors', 'Managers', 'Wanderers', and 'Participants', and the 'hardcore' and 'casual' gamer categories within those types.<sup>20</sup> Musical preferences amongst player types and categories can affect

---

<sup>17</sup> Firaxis Games, *Sid Meier's Civilization IV*, Novato: 2K Games, Austin: Aspyr (Microsoft Windows, Mac OS X, 2005).

<sup>18</sup> Karen M. Cook, 'Music, History, and Progress in *Sid Meier's Civilization IV*', in Donnelly, Gibbons, and Lerner, *Music in Video Games*, p. 173.

<sup>19</sup> Richard Bartle, 'Virtual Worlds: Why People Play', *Massively Multiplayer Game Development*, 2.2 (2005), 3–18 (p. 3) <[https://www.researchgate.net/publication/308073596\\_Virtual\\_worlds\\_Why\\_people\\_play](https://www.researchgate.net/publication/308073596_Virtual_worlds_Why_people_play)> [accessed 24 October 2018].

<sup>20</sup> Phillips, 'A Composer's Guide to Game Music', pp. 80–81.

the compositional process, as Phillips examines the target market.<sup>21</sup> Bartle describes the alignment of characters to players, as well as player embodiment, through soft role-playing:

Most virtual worlds, however, use soft role-playing. Here, not only does the player change to fit the character, but the character changes to fit the player. The very aim of soft role-playing is to align character and player: to find a 'you' that you like to be.<sup>22</sup>

The player's type can likely change their way of engaging with characters within a game world which can cause issues when worldbuilding. However, even though player types may differ, worldbuilding coherency in a game world can still aid the likelihood of the player becoming immersed and more involved with the game's functionalities.

### ***Immersion***

Immersion is a difficult concept to identify because of its individual relationship with players and different player types. Sarah Lynne Bowman discusses immersion and shared player imaginations in RPGs throughout her chapter 'Immersion and Shared Imagination in Role-Playing Games'. Bowman explains the concept of immersion as an experience which causes players to use 'phrases such as 'losing myself in the game' or 'the character took over'.'<sup>23</sup> Although these phrases are commonly used when discussing immersion in games, Bowman further identifies the issues of defining immersion as players describe immersion differently, which causes debates amongst scholars, and immersion is not something unique to games. In order to redefine and make the term essential, Bowman's chapter provides the reader with six immersion categories: activity, game, environment, narrative, character, and community.<sup>24</sup> Bowman identifies how players can be immersed within each of these categories rather than believing immersion is a single use term. Different player types, as outlined by Bartle, can thus become immersed in different immersion categories depending on their gameplay preferences. Calleja's book *In-Game: from Immersion to Incorporation* also considers player immersion across various tiers of involvement including kinesthetic, spatial, shared, narrative, affective, and ludic involvement. Although sharing only one phrase, 'narrative', Bowman and Calleja's views on immersion cover immersion as something that can be obtained across various categories. Calleja thus aligns immersion with the theory of presence, the sense of being within the

---

<sup>21</sup> Phillips, 'A Composer's Guide to Game Music', pp. 82–83.

<sup>22</sup> Bartle, 'Virtual Worlds', p. 9.

<sup>23</sup> Sarah Lynne Bowman, 'Immersion and Shared Imagination in Role-Playing Games', in Deterding and Zagal, *Role-Playing Game Studies*, p. 379.

<sup>24</sup> Bowman, 'Immersion and Shared Imagination in Role-Playing Games', in Deterding and Zagal, *Role-Playing Game Studies*, p. 382.



fictional world, which can create any of the multitude of senses, such as ‘engagement, perception of realism, addiction, suspension of disbelief, identification with game characters.’<sup>25</sup>

The idea of immersion as something that can be created through varying senses and preferences could mean that distinct types of players can become immersed in some categories but not others; the player of the explorer type could become immersed in the environment of a game, whereas a socializer player type may become more immersed with the characters of the world. When performingLARPs, players are physically acting within the fictional real-world game space. Immersion thus takes a different role compared to mediated environments as players are using their own bodies as mediums to embody characters. J. Tuomas Harviainen, Rafael Bienia, Simon Brind, Michael Hitchens, Yaroslau I. Kot, Esther MacCallum-Stewart, David W. Simkins, Jaakko Stenros, and Ian Sturrock thus define immersion in the LARP as a ‘common experience’ because players are physically there. Although a common experience, Tuomas Harviainen et al. discuss immersion in the LARP as unstable due to the active rulesets interfering with ‘in-character actions’, and thus interfering with immersive states when play must be paused in order for the rules to be cited.<sup>26</sup> This could also occur in table-top RPGs, but the players must only stop physically rolling dice and calculating rolls whilst the DM considers the ruleset. In video games, the game’s system calculates all rules as the player is playing, and even if the game pauses for player decisions the game world is still active.

James Newman considers video games to include three player motivations: immersion, challenge, and interaction. Newman describes immersion as an act of presence and being in the game world, gaining rewards and gratification of *their* work in the game world, rather than outside manipulators of a character.<sup>27</sup> Newman’s three parts of motivation and expectation aligns with Mihaly Csikszentmihalyi’s ideas of flow. Csikszentmihalyi considers flow as the outcome of player enjoyment through rewarding challenges:

We experience it as a unified flowing from one moment to the next, in which we feel in control of our actions, and in which there is little distinction between self and environment; between stimulus and response; or between past, present, and future.<sup>28</sup>

---

<sup>25</sup> Calleja, ‘In-Game’, p. 25.

<sup>26</sup> J. Tuomas Harviainen, Rafael Bienia, Simon Brind, Michael Hitchens, Yaroslau I. Kot, Esther MacCallum-Stewart, David W. Simkins, Jaakko Stenros, and Ian Sturrock, ‘Live-Action Role-Playing Games’, in Deterding and Zagal, *Role-Playing Game Studies*, p. 90.

<sup>27</sup> James Newman, *Videogames*, 2nd edn (London: Routledge, 2013), pp. 15–16.

<sup>28</sup> Mihaly Csikszentmihalyi, *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* (London: Springer, 2014), p. 137.

Csikszentmihalyi's definition of flow is not that the player forgets where they are, but action and awareness are merged where the player is only aware of their actions and not awareness itself.<sup>29</sup> This means that in order to sense flow, one must be immersed and intrinsically involved in a task, and thus Csikszentmihalyi regards immersion as a condition for flow.<sup>30</sup>

The motivation and expectations of the player within a game world can rely on functionalities to create this sense of flow in a game, such as game world interfaces meant for player benefit. Kristine Jørgensen defines immersion as involvement with game world interfaces that aid the player with gameplay. Interfaces assist the player's existence and functionality in a game (much like that of a racket in a game of tennis). Interfaces exist on-screen in a virtual game but are often only seen by the player (or characters do not realise the full functionality of an interface) and give the player quality of life controls in the game world. Considering interfaces as a possibility to draw the player out of immersion, by providing information that is not a natural sense in a world, is refuted by Jørgensen. Jørgensen argues the 'transparency fallacy' by claiming that in the 'real-world' we frequently use tools as a mediatory without losing involvement in the task (i.e. playing an instrument). The 'transparency fallacy' is thus argued by Jørgensen as an outcome of the 'immersion fallacy', where it is believed that players truly transcend between their world and the game world, and that every aspect of a game needs to thus simulate the 'real world'.<sup>31</sup>

Although immersion can primarily be considered as a part of the game world itself, developers can outsource their immersive worldbuilding tools. Extra components such as trailers or gameplay demos, that are used to encourage the player to buy a game, can be used as a lore builder to encourage the player to continue playing a game and (if there is a chance) invest more money into the game and franchise; for example, the Nintendo Switch's Nintendo eShop allows players to download free gameplay demos of certain games, such as *Rayman Legends: Definitive Edition*.<sup>32</sup>

### ***Outsourced Worldbuilding***

Not all worldbuilding occurs within the diegesis. The paratext of a piece of literature is the creation of extra material which surrounds, and relates to, the original text but is not always made by the author. As with the paratext of a piece of literature, a game world's organicism and coherency can be influenced by material created by developers and fans outside of the diegesis. Henry Jenkins

---

<sup>29</sup> Csikszentmihalyi, 'Flow and the Foundations of Positive Psychology', p. 138.

<sup>30</sup> Csikszentmihalyi, 'Flow and the Foundations of Positive Psychology', p. 216.

<sup>31</sup> Kristine Jørgensen, *Gameworld Interfaces* (London: The MIT Press, 2013), pp. 29–31.

<sup>32</sup> Ubisoft Montpellier, *Rayman Legends: Definitive Edition*, Montreuil: Ubisoft (Nintendo Switch, 2017).

discusses the creation of fan texts as ‘textual poaching’, a term for when fans (or anyone outside of the original creator) appropriate the initial text by creating their own material from their readings of the source. However, Jenkins considers textual poaching more as a ‘participatory culture’<sup>33</sup> in which fans are merely participating in a text that they appreciate and enjoy, rather than simply poaching. Jenkins considers the reader as someone who rethinks the meaning of the author’s text and applies it to their own experiences.<sup>34</sup> This poaching, rethinking, rereading, or participatory culture of original texts provides fans and readers with a secondary glance into the textual world.

Where the term ‘textual poaching’ appears as a negative fan behaviour, I agree with Jenkins’ examination that these extra-diegetic texts of video games can be a positive, ‘outsourced’, form of worldbuilding; outsourced worldbuilding can encompass fan, developer, or developer-contracted texts. Although there can be problems with fan made content, as they do not own the rights to the original source material, creative readings of texts nourish and enhance a world’s lore and can engage players further within the universe. Whether or not the developers officially accept or endorse any fan-made content, this content can be consumed and accepted by other players as a part of their own perception of the game world’s lore. I use the term ‘outsourced’ because it can refer to any source of worldbuilding conducted outside of the diegesis and, unlike the term ‘paratext’, can include fan-made (non-developer associated) texts.

Outsourced worldbuilding contains the ‘fringe’/ ‘semi-canon’, and ‘metatexts’ of games and other mediums that relate to that world. Outsourced worldbuilding includes extra game content released by developers or made by fans, such as physical copies of game maps, reviews, fan made Wikias, soundtracks, game guides, streaming of the game, forums, figurines. These outsourced worldbuilding components allow potential, current, or previous players to be involved in the game world without the need for play. The engagement with a game outside of its diegesis means that the game world can be built from afar. The potential of outsourced worldbuilding can keep players immersed in the game world outside of play or immerse potential players who have yet to buy the game. Melanie Fritsch discusses the complexity of fan made outsourced worldbuilding:

Fans do not only watch and silently contemplate the contents they are presented with, but have instead developed a huge range of ‘gaming the systems’ practices. One vivid example is the practice of fanfiction: fans write their own stories, thereby changing the narrative by

---

<sup>33</sup> Henry Jenkins, *Textual Poachers: Television Fans and Participatory Culture*, 2nd edn (London: Taylor & Francis, 2013), p. xxi.

<sup>34</sup> Jenkins, ‘Textual Poachers’, p. 51.

recontextualizing the importance of, or relationships between, characters, adding content to bridge or exploit perceived gaps in the narrative, or adding a totally new storyline.<sup>35</sup>

What Fritsch outlines here is player engagement with a fictional world that has been built consistently and coherently, enough so that players want to further immerse themselves in the world outside what has been set by developers.

Hype is a colloquial term used in the gaming community that refers to an emotional excitement felt across communities. The sensation of hype is often built towards the release of a game, encouraged by developers and publishers, which can pressure individuals to pre-order a game (buying the game before it is released to ensure players get the game on the day of release). The issue of pre-ordering games is the lack of critical reception before playing, and thus if a game does not perform well can lead to a waste of player money. The emotional hype surrounding the release of the video game *No Man's Sky*,<sup>36</sup> a game which boasted a procedurally generated universe which the player can endlessly explore and discover new planets, caused players to buy the game before release and before critics could comment on the game's features. Pre-sales of the game were exceptional because of the hype and sensation surrounding it. Due to this sensation, gamers ignored the lack of critic responses and the developer's vague responses to game journalists' questions regarding game content and expected player experiences. After *No Man's Sky* was released, it became apparent that the gameplay promised by the developers was not realised in game, leading to distributors providing players with full refunds on the game.<sup>37</sup> Although the emotional state of hype, or excitement, and the collection of physical manifestations of a video game is separate from the player physically playing and engaging with a game world, outsourced worldbuilding can immerse the player in a game world before they have entered the world as the player. This can be both beneficial to players and detrimental depending on player handling of hype levels.

The degree of outsourced worldbuilding depends on the classification and finance of a game. 'AAA video games' is a colloquial term given to games that demonstrate production values similar to the

---

<sup>35</sup> Melanie Fritsch, 'It's a-me, Mario!' Playing with Video Game Music', in *Ludomusicology*, ed. by Michel Kamp, Tim Summers, and Mark Sweeney (Sheffield: Equinox Publishing Ltd, 2016), p. 99.

<sup>36</sup> Hello Games, *No Man's Sky*, Guildford: Hello Games (PlayStation 4, Microsoft Windows, Xbox One, 2016).

<sup>37</sup> After *No Man's Sky's* release, most large scale online and physical gaming magazines voiced their views on the dangers of becoming 'hype' about games that have not received any critical pre-release reviews; Tyler Wilde, 'The Anatomy of Hype: How No Man's Sky Became the Best and Worst Game Ever', *PC Gamer*, 15 August 2016 <<https://www.pcgamer.com/uk/the-anatomy-of-hype-how-no-mans-sky-became-the-best-and-worst-game-ever/>> [accessed 7 November 2018]; Jason Schreier, 'The No Man's Sky Hype Dilemma', *Kotaku*, 18 August 2016 <<https://kotaku.com/the-no-mans-sky-hype-dilemma-1785416931>> [accessed 7 November 2018]; Wesley Yin-Poole, 'No Man's Sky Changed the Video Game Hype Train Forever', *Eurogamer*, 27 December 2016 <<https://www.eurogamer.net/articles/2016-12-20-no-mans-sky-changed-the-video-game-hype-train-forever>> [accessed 7 November 2018].

A-movies of modern Hollywood and can refer to both a game itself or the developer and producers. AAA games like *Red Dead Redemption 2*<sup>38</sup> are published by major developers with access to high finances and quick turnarounds of games, from development to release, because of their expectation of staff to work one-hundred-hour weeks. AAA games can usually release games over multiple platforms and afford larger marketing campaigns which bring, arguably, more hype towards releases. AAA producers and developers therefore have the capacity to afford the production of extra merchandise alongside the game, and thus increase the price of certain game bundles. Outside of the AAA market, independent (indie) games, developers and publisher create games with smaller budgets. Usually, indie developers have a smaller staff size, better working conditions which means games take longer to be released, and produce games that are cheaper to make and buy. The capacity here is usually to release a game and there may be none to limited merchandise (such as t-shirts) that may be sold exclusive of players buying the game; when selling on the online Steam platform it is normal to see indie games selling alone or with the music soundtrack. However independent titles are no less anticipated than the AAA. Games such as *The Witcher 3: Wild Hunt* was an anticipated game and could be considered a AAA release except for the fact that CD Projekt Red are an independent publisher and developer. The ‘hype’ and initial pre-ordering ‘greed’<sup>39</sup> seen during the release of certain games is due to the outsourced worldbuilding that occurs before a game is released. Players of table-top RPGs, like *D&D*, can buy extra physical products for their table, such as physical copies of spells, custom made maps, character and monster models, and environmental features and buildings to engage with the diegesis further.

---

<sup>38</sup> Rockstar Studios, *Red Dead Redemption 2*, New York: Rockstar Games (PlayStation 4, Xbox One, Microsoft Windows, Stadia, 2018).

<sup>39</sup> Many critics disagree with the way AAA games are run, crunch periods and 100 hour working weeks for staff, pre-order specials, limited editions, game exclusives, and loot box exploitations are all arguments which are currently important in gaming culture; see Jim Sterling, ‘Demolishing The Excuses Made For ‘AAA’ Publishers’ Exploitative Greed (The Jimquisition)’, *YouTube*, 15 October 2018 <<https://www.youtube.com/watch?v=ycz3U1HDfJo>> [accessed 7 November 2018]; Lanier, Liz. (2018) ‘Red Dead Redemption 2’s 100-Hour Work Weeks Spark Video Game Industry Outrage’, *Variety*, 15 October. Available at: <https://variety.com/2018/gaming/news/red-dead-redemption-2s-100-hour-work-weeks-spark-video-game-industry-outrage-1202980239/> (Accessed: 9 June 2020); Campbell, Colin. (2019) ‘How Fortnite’s success led to months of intense crunch at Epic Games’, *Polygon*, 23 April. Available at: <https://www.polygon.com/2019/4/23/18507750/fortnite-work-crunch-epic-games> (Accessed: 9 June 2020); Makuch, Eddie. (2018) ‘Top Bethesda Devs On Why Video Game Crunch Is Bad’, *GameSpot*, 27 October. Available at: <https://www.gamespot.com/articles/top-bethesda-devs-on-why-video-game-crunch-is-bad/1100-6462838/> (Accessed: 9 June 2020).

## Soundtracks

Soundtracks and OSTs are the metatext of a medium and thus become components of outsourced worldbuilding. Soundtracks can exist as physical copies of a medium's non-diegetic music as well as existing on websites such as YouTube; it is common for soundtracks to also be edited and included on playlists on YouTube, i.e. *Relaxing Video Game Music (with Rain and Thunder Sounds)*,<sup>40</sup> and *1 Hour of Beautiful Video Game Music (Playlist 1)*.<sup>41</sup> Originally, a soundtrack was a literal object that was integrated into a film during post-production. In modern times the soundtrack is associated with the non-diegetic music of a medium. Soundtracks and OST albums contain tracks to the medium which are often edited or re-orchestrated to be easily consumed by buyers. It is also common to have variances between the soundscape and soundtrack. OSTs are often arranged for the album, as opposed to following the non-diegetic music that the player would hear in game.

James Buhler divides the soundtrack outside a film's diegesis into dialogue, sound effects and music, and considers whether a soundtrack can convey a narrative without the visuals.<sup>42</sup> Buhler concludes that whether the soundtrack is successful apart from its image does not affect its success with the cinematics of a film.<sup>43</sup> The soundtrack can thus exist separately here as an outsourced tool for audiences to experience the sounds of the game outside of the imagery, whether it is successful away from the images. Out of Buhler's three divisions of the soundtrack, dialogue and sound effects often do not make it onto the OST due to their position in the diegetic space. Soundtracks can be heard in illogical positions from the initial narrative, played during car journeys, whilst reading literature, and more activities, warping the narrative identifiers within the soundtrack. However, engaging with the narrative identifiers of the world through a soundtrack may transport listeners to the associated visuals and world building. Returning to the consideration of the core canon music of 'Hedwig's Theme' to signify the *Harry Potter* universe. 'Hedwig's Theme' can be consumed by fans and researchers' years after its conception, resuming an affiliation with the 2001-2011 film series.

The discussion of AAA released games and the inclusion of extra promised material on release of the game has been reiterated through the mistreatment of the soundtrack for the AAA release *Doom*

---

<sup>40</sup> Laure Igami, 'Relaxing Video Game Music (with Rain and Thunder Sounds)', *YouTube*, 7 May 2019 <<https://www.youtube.com/watch?v=DI3Bd1rdleE>>.

<sup>41</sup> ImagineBaggins, '1 Hour of Beautiful Video Game Music (Playlist 1)', *YouTube*, 10 January 2017 <<https://www.youtube.com/watch?v=EETFIPEnBh4>>.

<sup>42</sup> James Buhler, 'Analytical and Interpretive Approaches to Film Music (II): Analysing Interactions of Music and Film', in *Film Music: Critical Approaches*, ed. by K.J Donnelly (New York: Continuum, 2001), p. 53.

<sup>43</sup> Buhler, 'Analytical and Interpretive Approaches to Film Music (II)', in 'Donnelly, *Film Music*, p. 58.

*Eternal*.<sup>44</sup> Mat Omblor reported the outcry of players when the digital release of *Doom Eternal*'s soundtrack, already delayed a month after the game's initial release on March 20<sup>th</sup> 2020, was released with errors; the part of the soundtrack that was not mixed by the composer, Mick Gordon, was of poor quality after being shown to have heavy compression on most of the tracks.<sup>45</sup> The frustration from fan bases highlights the importance of hearing the game world within the outsourced soundtrack, and this must be in the quality that players remember from the game.

Players do not need to engage with the game world or its aural components outside of play, however these sources can be used as worldbuilding if players choose to engage, as they are entering the lore of the world even outside of it.

### ***Diegesis Worldbuilding: Game Interfaces and In-Game Audio***

Outsourced worldbuilding constructs a fictional world's concept outside of direct interaction with the source material. Diegesis worldbuilding considers worldbuilding through direct interaction with the world itself, incorporating moving images, audio, narratives, and objects that are contained within spatial boundaries. Dominic Lash explains 'that a consistent and coherent diegesis directly presents a consistent and coherent world – when of course the latter is actually only an inference made by the willing spectator – with the notion of reality itself.'<sup>46</sup> This notion of reality is created by the consistency of a world, which includes gameplay mechanics within a game world. Within film musicology, the discussion of a diegesis considers the inclusion of diegetic and non-diegetic. In film sound, Michel Chion, in his text *Audio-Vision: Sound on Screen*, examines the spatiotemporal diegesis of a film's audio:

Music can swing over from the pit to screen at a moment's notice, without in the least throwing into question the integrity of the diegesis, as a voiceover intervening in the action would. No other auditory element can claim this privilege. Out of time and out of space,

---

<sup>44</sup> Id Software, *Doom Eternal*, *Rockville: Bethesda Softworks* (Microsoft Windows, PlayStation 4, Stadia, Xbox One, Nintendo Switch, 2020).

<sup>45</sup> Mat Omblor, 'Doom Eternal's Bungled Soundtrack Highlights the Mistreatment of Video Game Music', *VG247*, 22 April 2020 <[https://www.vg247.com/2020/04/22/doom-eternal-soundtrack-bad-mixing/?utm\\_medium=social&utm\\_source=twitter](https://www.vg247.com/2020/04/22/doom-eternal-soundtrack-bad-mixing/?utm_medium=social&utm_source=twitter)>.

<sup>46</sup> Dominic Lash, 'Film as Film in Twenty-First Century: The Function and Evaluation of Diegesis and Disruption', *Screen*, 58.2 (2017), 163–79 (pp. 167–68).

music communicates with all times and all spaces of a film, even as it leaves them to their separate and distinct existences.<sup>47</sup>

This communication of time and space to an audience through non-diegetic sound can build the emotional connection between audience and character. Originally outlined by Claudia Gorbman, in her book *Unheard Melodies*,<sup>48</sup> the issue of omitting non-diegetic sound from the diegesis, through its outline as solely narrative identifiers, is explored by a myriad of academics. Ben Winters examines Christian Metz's definition of diegesis, the entirety of a film's narrative space, when understanding sound and visuals. Winters' reviews this definition of Metz's diegesis whilst defining the space of non-diegetic sound:

[...] whether music belongs rightfully in the diegesis depends on whether it is understood as denotative. Although 'non-diegetic' music is widely assumed to be connotative, and to have little to do with denoting objects in narrative space, one of Adorno and Eisler's chief criticisms of Hollywood scoring was precisely music's redundant, almost denotative character.<sup>49</sup>

The denotative and connotative discussion of non-diegetic music has caused contention in what is considered a part of the diegesis. Winters further examines the non-diegetic soundtrack as integral to the narrative of a film or game, even though it is separate from the visuals:

Trying to imagine the opening idol-stealing scenes of *Raiders of the Lost Ark* (Steven Spielberg, 1981) without John Williams's music is, I would suggest, an unnerving experience: we feel the lack in a way that has prompted many film directors to refer to the music in their films as 'an extra character' or, as with Spielberg, to acknowledge music's constitutive role in defining a character<sup>50</sup>

Winters' argument considers non-diegetic as a dictator of meaning within the diegesis, a meaning inferred to the viewer. However, aural diegesis worldbuilding was restructured by Robynn J. Stilwell and James Buhler's introduction of the functionality of sound between the diegetic and non-diegetic, the 'fantastical gap'.<sup>51</sup> The 'fantastical gap' offers a definition for sound that consists between the

---

<sup>47</sup> Michel Chion, *Audio-Vision: Sound on Screen*, ed. by C Gorbman, trans. by Claudia Gorbman (New York: Columbia University Press, 1994), p. 81.

<sup>48</sup> Claudia Gorbman, *Unheard Melodies: Narrative Film Music* (Bloomington: Indiana University Press, 1987).

<sup>49</sup> Ben Winters, 'The Non-Diegetic Fallacy: Film, Music, and Narrative Space', *Music & Letters*, 91.2 (2010), 224–44 (p. 227).

<sup>50</sup> Winters, 'The Non-Diegetic Fallacy', pp. 230–31.

<sup>51</sup> Robynn J. Stilwell, 'The Fantastical Gap between Diegetic and Nondiegetic', in Daniel Goldmark, Lawrence Kramer, and Richard D. Leppert, *Beyond the Soundtrack: Representing Music in Cinema* (London: University of California Press, 2007), pp. 186–87.



diegetic and non-diegetic, sound that may or may not be heard by the characters. The ‘fantastical gap’, and the ever-changing view and audition point of the audience and player, is summarised by Annette Davison as ‘potential to enable far more specific conceptualization of the border crossings that characterize so much film music, but which have thus far been little discussed.’<sup>52</sup> During the *Music and the Moving Image XII* conference in 2019, Chandler Reedler examined the use of the ‘fantastical gap’ during travel and in-car sequences, in the film *Lady Bird* (2017), as an indicative and literal vessel of the protagonist character, Lady Bird. The film uses audio within the fantastical gap to highlight character status to audiences, whilst also being a part of the characters world.<sup>53</sup> For players of a game, the diegesis includes user interfaces (UIs) that often occur within this ‘fantastical gap’; UIs are made for the player’s quality of life and management of the game, but have the possibility of being seen and heard by characters because of the UIs existence within the diegesis. These terms, diegetic, non-diegetic, and the fantastical gap, will be used throughout the thesis to highlight where the sound is emanating from as this may affect which components of the game are being built. In particular, the fantastical gap is used to later analyse a character’s performance in *Divinity: Original Sin II*, within Chapter five, as her seemingly solo performance is accompanied by extra female voices and string instrumentation that is not visible to the player, but is still a part of the characters diegetic performance.

## **Interfaces**

Interface worldbuilding consists of these interactive and technological aspects of games. Linda Pfister and Sabiha Ghellal believe it is unsuitable to translate the filmic discussions of the diegesis to a game world. Pfister and Ghellal discuss the importance of a game’s diegesis to ‘provide information about the game state to the player and give a certain degree of control.’<sup>54</sup> This player control is unlike audience agency when watching a film, and each player’s interaction with a diegesis can be unique. Players have access to user interfaces within the game world which act as virtual interactive on-screen tools. Jørgensen also discusses the issue of using filmic considerations of the diegesis, due to these added game mechanics, interactivity, and player agency in the game world environment:

---

<sup>52</sup> Annette Davison, ‘Reviews of Books: Beyond the Soundtrack: Representing Music in Cinema’, *Music & Letters*, 90.1 (2008), 113–16 (p. 113).

<sup>53</sup> Chandler Reeder, ‘How Music Shapes Houses and the Domestic Space as Vessels Towards a Romanticized Reality in *Lady Bird* (2017)’, in *Music and the Moving Image XII*, 2019.

<sup>54</sup> Linda Pfister and Sabiha Ghellal, ‘Exploring the Influence of Non-Diegetic and Dieget’, in *Proceedings of the 30th Australian Conference on Computer-Human Interaction* (New York, 2018), pp. 490–94 (p. 491).

The application of the idea of a world environment centered around fictional events and storylines to gameworlds may seem unproblematic to adopt because gameworlds can be seen as a kind of fictional world and often feature storylines. However, using diegesis to understand the gameworld is imprecise and potentially confusing because of the term's origin and the fact that film diegeses are qualitatively different from gameworlds.<sup>55</sup>

UIs offer information to the player in formats that help them navigate the diegesis, including on-screen maps, compasses, temperature, wind direction, weather, remaining health, stamina, weapon ammo, or other appropriate information depending on the game and genre; UIs offer information that the player loses by not physically being within the diegesis. UIs in non-virtual worlds, such as table-top RPGS, can include player reference guides, written rulesets, character sheets, dice, and other extra physical resources that allow the player to effectively interact with the world. Without this information players will not know which actions they can or should perform, as seeing information such as the health of a character can inform actions such as recovering health, evading, attacking, or hiding; although, some players enjoy the challenge turning off on-screen UIs where possible. UIs thus worldbuild as a functional tool for the player to engage with the game. Interfaces provide quality of life to the player, such as providing a place for players to 'level up' their characters. UIs do not intend to provide realism in a game but provide a comfortable and entertaining experience in the game world. Jørgensen explains that the player experience of a game is subjective and differs depending on the player type and their play preference in video games, and thus the interface of a game must respond appropriately to each player type.<sup>56</sup> The interface can therefore seem unimportant to worldbuilding (as it is not an integral component of the characters' game world and exists mostly for the players benefit), but the UI can keep the player engaged in the world through its ease of use.

Technological Interfaces can alter the audio and visual experience of a diegesis. These interfaces are the technological barrier to video games, the equipment the player must own to witness the game world, such as a television screen, monitors, a console or personal computer (PC), controllers, keyboard and mouse, and any headphones, speakers, and microphones which are necessary to a game's functionality. A working console or PC, equipped to a television screen or a monitor where necessary, is vital to experiencing virtual game worlds. These interfaces can create barriers to worldbuilding, as the type of PC or console may limit the types of games available to the player. Variations in consoles or the quality of PC components, such as graphic cards, hard-drive space, and motherboards, can affect the quality of a game; if the player has a poor quality graphics card then

---

<sup>55</sup> Kristine Jørgensen, *Gameworld Interfaces* (London: The MIT Press, 2013), p. 65.

<sup>56</sup> Jørgensen, 'Gameworld Interfaces', pp. 33–34.

they may not be able to witness a game in its intended definition. Only after the interface is set up can the player engage with the game as the developer intended.

These variations in player technologies can affect the audio quality and content the player hears, i.e. the differences between listening to audio with or without headphones. Technological interfaces themselves, such as consoles and PCs, can make sounds when functioning (i.e. the cooler fans) that bleed into the game world. William Cheng recounts an anecdotal experience of the survival horror game *Silent Hill*,<sup>57</sup> its industrial soundscape, and his refrigerator:

At one point while playing through *Silent Hill* for the first time, I noticed a hum: low in register, soft, but timbrally distinctive enough to be audible amid the many layers of in-game noises already in the mix. On a notepad, I documented this sound as a drone pitched approximately at D two octaves below middle C. When I turned off the game twenty minutes later, however, I could still hear the noise. My emotional trajectory rapidly followed thus: panic, puzzlement, and finally embarrassment, when I realized that the sound was coming not from the television set at all, but from my refrigerator. It was a noise I had long learned to tune out while going about daily business in my apartment— and yet, when mashed against this game’s audio, it returned with sudden vengeance.<sup>58</sup>

The soundscape of the video game world can be interrupted by the sounds of technology or the general ambience of the player’s surroundings. This may not be completely intentional by a developer or composer but may have been a consideration when choosing aural and visual styles.

## **Soundscapes**

Diegesis considers all elements that emanate from the on-screen world. Aurally, this includes all non-diegetic sound and music, speech from player-characters/avatars, non-playable-characters, sounds of objects, all narrative levels, environmental factors and other visuals, and audio. With the player engaged with the diegesis through various interfaces, the soundscape can begin to aid the worldbuilding process. A soundscape can envelope all sound, including diegetic and non-diegetic music, sound design, voices, silence, and environmental sounds such as birdsong, wind, and footsteps. A soundscape is a part of the diegesis and is not often considered in the outsourced soundtrack in its entirety, due to the complexity of sounds involved. A musical cue that plays within

---

<sup>57</sup> Konami Computer Entertainment Tokyo, *Silent Hill, Tokyo: Konami* (PlayStation, 1999).

<sup>58</sup> William Cheng, *Sound Play: Video Games and the Musical Imagination* (New York: Oxford University Press, 2014), p. 101.

the soundscape may vary from the outsourced soundtrack. What would be a complete track on an album can be altered within the diegesis. The track may be paused or manipulated by composers, directors, developers, or players for the effect of gameplay, visuals, dialogue, or narrative.

Therefore, unlike the soundtrack, the soundscape is more reliant on its visual partner as it must respond to characters, movement, locations, gameplay, and narratives connected to the diegesis.

Alongside this, the sounds of a soundscape are not completely unique to a single medium. Ambient geographical sounds do not become a signifier for a particular fictional world due to their commonality as sound, and thus may struggle without its pairing with visuals; this is not to say soundscapes cannot be outsourced as soundtracks but it is unlikely from a fictional medium world.

David Samuels, Louise Meintjes, Ana Maria Ochoa, and Thomas Porcello examine the original concept of the soundscape, considered by the composer Murray Schafer, as ‘somewhat analogous to landscape insofar as it attempted to contain everything to which the ear was exposed in a given sonic setting.’<sup>59</sup> The idea of the soundscape as analogous to a landscape further cements the soundscape as a pairing with visuals and diegesis. Landscape, soundscapes, the diegesis and narrative, are further discussed by Aimee Mollaghan when reviewing the film-maker Andrew Kötting:

Acoustic ecology ostensibly studies the environment through increased awareness of sound by virtue of an emphasis on the environmental context in which the sound occurs. Similar to the belief of the psychogeographers that the geography of an environment had a psychological effect on the human mind, proponents of acoustic ecology held that humans are affected by the sound of the environment in which they find themselves. Further to this, acoustic ecologists examined the extent to which soundscapes could be shaped by human behaviour.<sup>60</sup>

Acoustic ecology and the shaping of human behaviour have been outlined in research that considers the everyday use and manipulation of humans through chosen and unchosen music. John A. Sloboda considers ‘everyday music’ as a dependant on its locational occurrence and situation of use. Sloboda considers both static and portable music from shops to public transport, CDs and MP3s as everyday music, whereas curated music in a lesson, in the concert hall, and as therapy are not considered

---

<sup>59</sup> David Samuels, Louise Meintjes, Ana Maria Ochoa, and Thomas Porcello, ‘Soundscapes: Toward a Sounded Anthropology’, *Annual Review of Anthropology*, 39 (2010), 329–45 (p. 330) <<https://doi.org/https://doi.org/10.1146/annurev-anthro-022510-132230>>.

<sup>60</sup> Aimee Mollaghan, ‘An Audio-Visual Gallivant: Psychogeographical Soundscapes in the Films of Andrew Kötting’, *The Soundtrack*, 3.2 (2010), 125–30 (p. 126) <[https://doi.org/10.1386/st.3.2.125\\_1](https://doi.org/10.1386/st.3.2.125_1)>.

‘everyday music’.<sup>61</sup> The expectations humans have of ‘everyday music’ in social or personal environments can thus be mirrored by a fictional soundscape. John Connell and Chris Gibson consider the shaping of spaces through these expectations of music:

Geographical space is not an ‘empty stage’ on which aesthetic, economic and cultural battles are contested. Rather, music and space are actively dialectically related. Music shapes spaces, and spaces shape music. In various ways sounds have been used to create spaces and suggest and stimulate patterns of human behaviour in particular locations.<sup>62</sup>

Thus, recognisable geographical soundscapes (i.e. the soundscapes of a city or park area) can be used to the advantage of multimedia by recreating these sounds in the fictional world. Audiences frequently exposed to these everyday soundscapes may then familiarise themselves with the supposed landscape in the diegesis, i.e. the use of birdsong and wind sounds to represent the small rural English village in *Everybody’s Gone to the Rapture*.<sup>63</sup>

## Interacting with Video Game Worlds

The soundscape of a video game thus needs to virtually recreate the sensation of ‘being’, through worldbuilding familiarity with landscape environments and character being. Outlined are the practicalities of music and sound in relation to worldbuilding games, including non-virtual games like *D&D*.

Worldbuilding in the virtual RPG considers variations on tools that have been identified through film scholarship. Terminologies that refer to diegesis worldbuilding in film scholarship become somewhat outdated in relation to video games. The increase of audience agency as player interaction in video games has thus brought forwards new terminologies that enhance previous film scholarship. Jørgensen outlined new terminologies that differentiate the interactivity of video games from the linear mediums of television and film. Jørgensen’s PhD thesis, *‘What are Those Grunts and Growls Over There?’ Computer Game Audio and Player Action* contributes discussions of the diegesis and the inappropriate use of terminologies, such as diegetic and non-diegetic, regarding interactive soundscapes. Jørgensen’s definition of sound as an information space as ‘transdiegetic’:

---

<sup>61</sup> John A. Sloboda, ‘Music in Everyday Life: the Role of Emotions’, in *Handbook of Music and Emotion: Theory, Research, Applications*, ed. by Patrik N. Juslin and John A. Sloboda (Oxford: Oxford University Press, 2010), p. 497.

<sup>62</sup> John Connell and Chris Gibson, *Sound Tracks: Popular Music, Identity and Place* (London: Routledge), p. 192.

<sup>63</sup> The Chinese Room and SCE Santa Monica Studio, *Everybody’s Gone to the Rapture, San Mateo: Sony Computer Entertainment* (PlayStation 4, Microsoft Windows, 2015).

These transdiegetic sounds are central for the comprehension of the positioning of sound in computer game spaces, and work as a bridge between the game world and the player's world. In this respect, these sounds become part of the interface, and provide usability information at the same time as they are stylistically and functionally connected to the game world.<sup>64</sup>

This bridge between the diegesis and player worlds, much like Stilwell's 'fantastical gap', is often unclear whether characters within the diegesis can hear these sounds; is it possible that Mario can hear the rising perfect fourth when he, through player movement, picks up a coin? The building of a relationship between player, character, and the game world as a whole is developed by these reactive sounds. Returning to film scholarship however, Ben Winter's had considered a similar use of reactive sound in film that bridges the gap between audience and characters:

Secondly, intra-diegetic music or sound exists in the film's everyday narrative space and time, and is thus properly thought of as part of the film's fabula: it may be considered to be produced by the characters themselves (either as a result of their physical movements, as with mickey-mousing, as an expression of their emotional state, or as a musical calling-card), or by the geographical space...<sup>65</sup>

Ben Winters' discussion of 'intra-diegetic' sound synergises with Jørgensen's transdiegetic sounds. Although this is towards a filmic purpose, Winters' had considered non-diegetic sound as an outcome of narrative which exists in the same space as the characters, relating it to the gameplay of video games and suggesting 'that diegetic effect is independent of narrative.'<sup>66</sup> This distinction of sound as a narrative outcome within the a character's space defines the need for Jørgensen's bridge between player and character sound world.

Karen Collins, in her seminal 2008 text *Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design*, also refers to yet redefines filmic discourse. Collins considers Chion's 'acousmatic sound', sound that an audience hears within film without seeing its original source<sup>67</sup>, as a point of interactivity and movement in games:

This function, though present in films (even if we cannot force the camera to go in that direction, we mentally look there, argues theorist Michel Chion), is far more pronounced in

---

<sup>64</sup> Kristine Jørgensen, "What Are Those Grunts and Growls Over There?" (Copenhagen University, 2007), p. 76.

<sup>65</sup> Winters, 'The Non-Diegetic Fallacy', p. 237.

<sup>66</sup> Winters, 'The Non-Diegetic Fallacy', p. 228.

<sup>67</sup> Chion, 'Audio-Vision', pp. 71–72.

games, as sound gives the player cues to head in a particular direction or to run the other way, therefore affecting player decision making.<sup>68</sup>

Collins is identifying the ability of players, unlike television and film audiences, to move towards this acousmatic sound. The player has agency from this interaction and thus the inclusion of sounds in the soundscape must be prepared with the player in mind. The decision to make a sound diegetic, non-diegetic, or transdiegetic are affected by the developer's intention for the player to find the sound, or not. This discussion of interactive sound was further researched by Collins when speaking in greater detail about the interactivity of game audio in texts such as in *From Pac-Man to Pop Music: Interactive Audio in Games and New Media*<sup>69</sup> in 2008, and the 2014 *The Oxford Handbook of Interactive Audio*, edited with Bill Kapralos and Holly Tessler.<sup>70</sup>

Bridging these sounds between the player and the virtual world thus requires a practical and technically informed composition processes that allows the soundscapes to react to player movement. Winifred Phillips, in her book *A Composer's Guide to Game Music*,<sup>71</sup> introduces these practicalities of composing for video games by explaining the creative and technical skillsets needed for new composers. A composer for video games such as *God of War*,<sup>72</sup> *LittleBigPlanet 2*,<sup>73</sup> and *Assassin's Creed III: Liberation*,<sup>74</sup> Phillips' discussion of interactivity from a composer's perspective, engaged with film literature, outlines the foundations of becoming a video game composer. Phillips stresses the importance of learning technology that allows interactive composition as she explains that interactive audio 'has the capability to change based on the state of the game and the choices of the player, creating variations that are determined by the player's actions.'<sup>75</sup>

Interacting with the diegesis of a video game is something that can also be found during certain outsourced worldbuilding. *Video Games Live* is a prevalent concert series which orchestrates and performs popular video game soundtracks with a live orchestra. The concert series tours worldwide and encourages its audiences to interact with the music during the performances; participants are encouraged to vocalise, 'shout', 'clap', 'whoop', 'stomp', and physically involve themselves with the

---

<sup>68</sup> Karen Collins, *Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design* (Cambridge, Massachusetts: MIT Press, 2008), p. 130.

<sup>69</sup> Karen Collins, *From Pac-Man to Pop Music: Interactive Audio in Games and New Media* (Aldershot: Ashgate, 2008).

<sup>70</sup> *The Oxford Handbook of Interactive Audio*, ed. by Bill Kapralos, Karen Collins, and Holly Tessler, Oxford Handbooks (New York: Oxford University Press, 2014).

<sup>71</sup> Phillips, 'A Composer's Guide to Game Music'.

<sup>72</sup> SCE Santa Monica Studio, *God of War, San Mateo: Sony Computer Entertainment* (PlayStation 2, PlayStation 3, PlayStation Vita, 2005).

<sup>73</sup> Media Molecule, *LittleBigPlanet 2, San Mateo: Sony Computer Entertainment* (PlayStation 3, 2011).

<sup>74</sup> Ubisoft, *Assassin's Creed III: Liberation, Montreuil: Ubisoft* (PlayStation Vita, 2012).

<sup>75</sup> Phillips, 'A Composer's Guide to Game Music', p. 158.

music. These concert series are a stark contrast to the expectation of participants within the usual concert hall performance. These interactive performances encourage fans of certain game worlds to reengage with their nostalgic memories of gameplay throughout the performance. Scholarly research surrounding the permeation of video game music in the concert hall has blossomed alongside its popularity within the video game community. The 2017 *Ludomusicology* conference included a session titled *In Concert* which included four speakers, all with similar experiences of these video game concerts.<sup>76</sup> Each paper's consideration of video game concerts ranged from discussing immersive components of video game concerts, video game genres popular within concerts, to the analytical evaluation of a concert piece from a singular game such as *Final Fantasy VII*. Beth Hunt at the 2019 *Constructing the Moving Image: Identity and the Soundtrack* study day considered the nostalgic effect of concerts on generations of gamers and how the 'hype' from a particular game in concert may pass down game preferences to future generations.<sup>77</sup> The discussions of interactive outsourced worldbuilding, such as concert series, usually considers nostalgia for past gameplay sessions. This nostalgia is created by a performance of the soundtrack, allowing players to recognise environments, character emotions and developments, and narratives associated with that track; this is partly the reason for the frustration with *Doom Eternal's* poor quality soundtrack mix, as previously discussed.

The concept of interactive performances, or nostalgic experiences from soundtracks, are not restricted to the concert hall. Melanie Frisch's research covers player performance and the creation of music within video games. Also at the 2017 *Ludomusicology* conference, Frisch discussed the game creator *Super Mario Maker*,<sup>78</sup> a game where players are given the tools to create their own Super Mario levels. Players can create their own functional Super Mario levels, with components from the previous Super Mario games, including enemy types, terrain, and audio. The player created levels can be shared online for other players to engage with and complete, providing a plethora of unique Super Mario levels; some players take delight in the challenge of creating almost impossible levels to complete, however in order to publish the level online the creator must first complete the level themselves. Player relationships with music have been exposed in these creations. Recreations of pre-existing music through game sound was introduced by Frisch, including the recreation of

---

<sup>76</sup> The session included Joana Freitas: 'It's Storytelling Through Music': On Video Game Concerts and the Promise of an 'Immersive' and 'Collaborative Experience'; Elizabeth Hunt: 'Video Games Live and the Gamification of the Symphony Orchestra'; James S. Tate: 'The Allure of the 'JRPG Sound' within Video Game Concerts'; Ben Hopgood: 'Evaluating Video Game Music in the Concert Hall: A Study of the Music from *Final Fantasy VII*'.

<sup>77</sup> Beth Hunt, 'Identity & the Audience: Concerts of Moving Image Soundtracks', in *Constructing the Moving Image: Identity and the Soundtrack Study Day*, 2019.

<sup>78</sup> Nintendo EAD, *Super Mario Maker*, Kyoto: Nintendo (Wii U, Nintendo 3DS, 2015).



Queen's 'Don't Stop Me Now' through a complex level design in the game.<sup>79</sup> The voice and music of 'Don't Stop Me Now' was devised through specifically timed transdiegetic sound effects of in-game enemies, puzzle pieces, death sounds, movement noises, and the well-known rising perfect fourth coin sounds. This use of previous diegeses to allow players to be creators can allow interactions with further functionalities of game world, allowing nostalgic interactions with game world and other area of music. The slice of life simulation video game *Animal Crossing: New Horizons*<sup>80</sup> is not a 'musical game', however the inclusion of instruments within the game world has allowed players to recreate music from other video games, such as the RPG *Persona 5*.<sup>81</sup>

The creation of music within video games could be due to the nostalgia effect of these soundtracks. Andra Ivănescu discusses the use of appropriated popular music as nostalgia in video games alongside the inclusion of video game music in concerts, in her book *Popular Music in the Nostalgia Video Game: The Way It Never Sounded*, outlining both player and academic relationships with nostalgia. Ivănescu concerns nostalgia with the 'nostalgia video game', making it clear that her definition of 'nostalgia game' is separate from retro game genres whose aesthetics refer to the past:

Retro games can be nostalgia games when appropriation goes beyond the aesthetic level into a more complex reconfiguration. Nostalgia games, however, are not necessarily retro games, since they can combine imagistic and narrative signifying systems from film and other popular culture without having older video games as referent. In other words, these are distinct but not mutually exclusive categories.<sup>82</sup>

Ivănescu makes the distinction between 'borrowed' and 'appropriated' in her opening chapter, deciding that the use of slightly modified or recontextualised music in video games is more of an appropriation, as borrowing suggests that nothing will be changed and will be given back.<sup>83</sup>

Ivănescu's case studies consider games which use pre-existing, or appropriated, music of the past like the use of Annette Hanshaw's 'Daddy Won't You Please Come Home' recording over the violence in the dystopian 1920s aesthetical style of *BioShock 2*.<sup>84</sup> The re-contextualisation causes the

---

<sup>79</sup> Melanie Fritsch, 'Game-Music-Performance: Ludomusicology Meets Performance Studies', in *Ludomusicology Conference*, 2017.

<sup>80</sup> Nintendo EPD, *Animal Crossing: New Horizons*, Kyoto: Nintendo (Nintendo Switch, 2020).

<sup>81</sup> P-Studio, *Persona 5, Höfen: Deep Silver* (PlayStation 3, PlayStation 4, 2017); Ethan Anderson, 'Animal Crossing Player Puts on a Persona 5 Concert in New Horizons', *Twinfinite*, 13 April 2020 <<https://twinfinite.net/2020/04/star-citizen-videos/>>.

<sup>82</sup> Andra Ivănescu, *Popular Music in the Nostalgia Video Game: The Way It Never Sounded*, Palgrave Studies in Audio-Visual Culture (Cham: Palgrave Macmillan, 2019), p. 17.

<sup>83</sup> Ivănescu, 'Popular Music in the Nostalgia Video Game', p. 9.

<sup>84</sup> 2K Marin, *BioShock 2, Novato: 2K Games* (Microsoft Windows, PlayStation 3, Xbox 360, Mac OS X, Nintendo Switch, 2010).

appropriation of the aesthetic nostalgia as in the intended place for the song was not originally meant to be seen in a violent world.

### ***Player Embodiment***

The interactive and nostalgic qualities of video games and their environments may allow players to engage with the world by identifying customs and traits that they hear in other popular mediums and in 'everyday music'. This in turn may create player embodiment through familiarity of sounds in a game character's space and characterisations.

Embodied player reactions, drawing on Ivănescu's appropriated nostalgia, is explained by Cheng during his analytical play of the post-apocalyptic RPG, *Fallout 3*.<sup>85</sup> Cheng recounts his interaction with a quest that asks the player to detonate an unexploded nuclear bomb in a populated settlement. The player would receive a reward for completing the quest, a significant sum of money from the quest giver Mr. Tenpenny, making the player's life easier throughout the game as the reward can be spent on weapons and supplies. If the player does not detonate the bomb, they will save the settlement and make an enemy of Mr. Tenpenny, making their life more difficult. However, the long-term benefits of disarming the bomb is the future interaction with the population of the settlement. From an outsider's perspective this decision should not be difficult, and saving the settlement means the player can interact with the inhabitants later in the game. The video game decision to detonate the bomb, and take Mr. Tenpenny's money, comes with the preconception that the player's quality of life in the game will be better after taking the money, and not having Mr. Tenpenny as an enemy would make surviving in the post-apocalyptic world easier. Cheng's decision in this narrative diversion was swayed by the inclusion of music in this scene. *Fallout 3* provides the player with a digital radio in their Pip-Boy (an arm bracelet which allows the player to manage their character and inventory). The Pip-Boy cycles through pre-existing pieces of music from the 1950s and 1960s America that the inhabitants of the game have 'recovered' after the nuclear apocalypse; *Fallout 3* uses this pre-existing music in a post-apocalyptic world to 'build a picture of an everlasting America or at least an everlasting American Dream.'<sup>86</sup> These pieces are played repeatedly on the Pip-Boy's radio and continue to play without any player trigger except to turn the radio on and off. When Cheng is given the option to detonate or disarm the nuclear bomb, he recounts that John Philip Sousa's 'Stars and Stripes Forever' happened to be playing through the Pip-Boy in that

---

<sup>85</sup> Bethesda Game Studios, *Fallout 3, Rockville: Bethesda Softworks* (Microsoft Windows, PlayStation 3, Xbox 360, 2008).

<sup>86</sup> Ivănescu, 'Popular Music in the Nostalgia Video Game', p. 123.

moment. Cheng demonstrates the power of the appropriated music in this moment by describing his decision to detonate the bomb:

As I stood before the Big Red Button [...] this music didn't sound like a badge of liberty, democracy and good old-fashioned values. Rather, it drew my attention to Mr. Tenpenny's [the quest giver] embodiment of the Enclave's radical authorities and extremist ideologies. [...] Within a couple of minutes, the piece came to a rousing close, its final seconds punctuated by an emphatic cadence and an ascending octave flourish in the melody.

Then, before I knew it, I was pressing the button.<sup>87</sup>

This musical background was unique to Cheng at this moment as the radio plays music indiscriminately, and other players may have heard something completely different or musical silence during this decision. Cheng may have heard 'Stars and Stripes Forever' within *Fallout 3* before this moment, but its timing during this quest is comically persuasive. The melodic flourish that accompanies Cheng's detonation of the bomb is comparable to the stereotypical and dramatised death sequences of operatic, television, filmic, and video game characters. Cheng's decision was a manipulated, almost subconscious, reaction rather than a calculated ethical decision. Cheng, at this moment, was an embodied player separate from his own ethical directions and thus was unusually influenced by audio.

Michal Grover-Friedlander focuses on the lost voice in *The Phantom of the Opera*, considering the cinematic opera as a representation of a 'quest for the source of the operatic voice' although, 'in fact, *the reason* for the prima donna's death in tragic opera is precisely this doomed quest.'<sup>88</sup> The doom of the diva's quest and obsession of her vocalised death is outlined by Jeongwon Joe in her book *Opera as Soundtrack* as she dedicates a chapter to cinematic deaths and the use of opera alongside the death:

Over the past few decades, the use of operatic music in conjunction with cinematic death, especially ritualized, stylized, and climactic murder scenes, has been notable across diverse genres of popular titles, ranging from comedy, drama, and action to thriller, horror, and science-fiction films.<sup>89</sup>

Grover-Friedlander also comments on this almost inevitable demise of the diva in opera and now in cinema. Tim Summers highlights the use of opera, pre-existing and written for the game, in *Tosca* in

---

<sup>87</sup> Cheng, 'Sound Play', pp. 45–46.

<sup>88</sup> Michal Grover-Friedlander, *Vocal Apparitions: The Attraction of Cinema to Opera* (Princeton: Princeton University Press, 2005), pp. 30–31.

<sup>89</sup> Jeongwon Joe, *Opera as Soundtrack* (Burlington, Vt: Ashgate, 2013), pp. 56–57.

*Hitman: Blood Money*<sup>90</sup> and *The Beggar's Opera* in *Assassin's Creed II*,<sup>91</sup> alongside arias in *Final Fantasy VI*<sup>92</sup> and *Parasite Eve*,<sup>93</sup> and the use of Wagner in the interactive point-and-click game *The Beast Within: A Gabriel Knight Mystery*.<sup>94</sup> Summers considers a variety of games which use opera to either highlight feminine qualities or as an environment for assassinations and death.<sup>95</sup> As well as the joint obsession with death, the overarching 'quest' of opera and games shows an obsession with rebirth. In RPGs, rebirth is a part of the game as when the player-character is defeated, the player can begin playing again from varying points. If the game includes multiple party members, items and magic based spells are often included in the game that 'pick-up' a fallen ally (often described as unconscious, as opposed to dead).<sup>96</sup> These thematic connections between opera and video games are also seen in musical structures, such as character leitmotif, and a focus on narrative storytelling. Summers introduces these musical sutures through Wagner's structural flexibility of musical units within a space:

Deploying small units of object-associated music accommodates the nonlinear game structures. It was for the same reason that Wagner developed his theme-based musical construction for his later operas (sometimes referred to as 'music dramas'). To allow the musical fabric to closely mirror the dramatic action, Wagner rejected the traditional opera musical formal structures in favour of a more flexible approach based on thematic statement and transformation that followed the on-stage drama, rather than pre-existing templates of musical forms. For games, Wagner's model lends itself well to dynamic music systems linked to game events, particularly when the precise timing and order of such events (unlike a film) are difficult to predict.<sup>97</sup>

Wagner's dynamic musical systems lead the interactive and reaction-based game sound to environments and character developments. The embodiment of the voice on the operatic stage, and

---

<sup>90</sup> IO Interactive, *Hitman: Blood Money*, London: Eidos Interactive (Microsoft Windows, PlayStation 2, Xbox, Xbox 360, 2006).

<sup>91</sup> Ubisoft Montreal, *Assassin's Creed II*, Montreal: Ubisoft (PlayStation 3, Xbox 360, Microsoft Windows, Mac OS X, PlayStation 4, Xbox One, 2009).

<sup>92</sup> Square, *Final Fantasy VI*, Tokyo: Square (Super NES, PlayStation, Game Boy Advance, Nintendo Wii, Android, iOS, Microsoft Windows, Super NES Classic Edition, 1994).

<sup>93</sup> Square, *Parasite Eve*, Tokyo: Square Electronic Arts (PlayStation, 1998).

<sup>94</sup> Sierra On-Line, *The Beast Within: A Gabriel Knight Mystery*, Los Angeles: Sierra On-Line (MS-DOS, Macintosh, Microsoft Windows, 1995).

<sup>95</sup> Tim Summers, 'Opera Scenes in Video Games: Hitmen, Divas and Wagner's Werewolves', *Cambridge Opera Journal*, 29.3 (2018), 253–86.

<sup>96</sup> The term unconscious comes from its use in *D&D*. Death is often a permanent fixture in the game, and players of characters who die often choose to create a new character. However, before death in combat, when the player-character's health is reduced to below zero, the player can attempt to resurrect their character from unconsciousness by rolling ten, or above, three times on a d20, icosahedron, dice before they roll below ten three times.

<sup>97</sup> Tim Summers, *Understanding Video Game Music* (Cambridge: Cambridge University Press, 2016), p. 161.

the engagement with opera connects it with player-characters in video games, and thus discussions within ludomusicology. As the musical developments can form the corporeal body when on-stage, the corporeality of video game characters can induce player embodiment as the player takes control of their movements.

### ***Immersion and Embodiment***

As players embody a corporeal on-screen body, through reactive aural components, immersive tendencies can form. Audio's involvement within diegesis worldbuilding, and thus the formation of immersion, is a common discussion point in ludomusicology. As with most discussions of immersion, including those considered in game and film theory, ludomusicology has suffered from disparate and contended definitions. Ludomusicologists, however, have moved to discuss immersion from other viewpoints and implementations, rather than solely focusing on a definition.

Phillips defines immersion, from the viewpoint of the player, as 'the concept of sinking completely within something or being absolutely surrounded by it.'<sup>98</sup> Through a discussion of Paul Cairns and Emily Brown's identification of the three levels of immersion (engagement, engrossment, and total immersion), Phillips has informed masters theses such as my own<sup>99</sup> and Peter Goosey's<sup>100</sup> when formulating models and frameworks that attempt to document the audio steps to immersion. Summers provides an overview of the theoretical texturing of immersive aesthetics, in his seminal text *Understanding Video Game Music*. Summers begins by explaining that as players we crave immersion and 'we long for those times when a game is so utterly absorbing, we lose track of time and our surroundings.'<sup>101</sup> Summers 'musical texturing' acts as signifiers when the visual technology is restricted, providing player involvement in the game world through recognisable musical codes.<sup>102</sup> Summers' suggestion of 'musical texturing' thus shows how music can keep the engagement level of the player even when the visuals may be lacking signifiers.

Whereas Phillips and Summers outline immersion from a generally theoretical standpoint, other scholars consider specific game case studies through their own analytical play. Cheng considers immersion from his own anecdotal experiences of player immersion through player to player

---

<sup>98</sup> Phillips, 'A Composer's Guide to Game Music', p. 37.

<sup>99</sup> Jennifer Smith, 'The Musical Parameters of Immersion and Flow' (unpublished master's thesis, University of Huddersfield, 2016).

<sup>100</sup> Peter Goosey, 'Creating Immersion' (unpublished master's thesis, University of Huddersfield, 2017).

<sup>101</sup> Summers, 'Understanding Video Game Music', p. 58.

<sup>102</sup> Summers, 'Understanding Video Game Music', p. 60.

musical interactions. In Cheng's playthrough of the MMORPG *Lord of the Rings Online*<sup>103</sup> (*LoTRO*), he describes the player's ability to perform their own music in a game world recreation of J.R.R. Tolkien's Middle-earth. Cheng examines the conflict between players when 'period-correct music' is not performed by other players; players take issue with and report this musical disruption on their perception and worldbuilding of the fantasy-medieval styles of Middle-earth. Cheng also considers the challenge of player loyalty to the musical lore of *Lord of the Rings (LoTR)*, set by Howard Shore in the 2001-2003 film series,<sup>104</sup> and their musical preconceptions of a fantasy-medieval environment; *LoTRO* allows its players to musically bottom-up worldbuild, however the inclusion of multiple player types and personalities can create a conflict on what this worldbuilding looks like. The inclusion of sound and silence in an immersive soundscape is considered by William Gibbons, as he considers the use of minimal music and silence in his playthrough of the character-empty environments in *Shadow of the Colossus*.<sup>105</sup> Gibbons argues that although it was previously considered that games needed continuous music to keep immersion, the act of silence in games is often imperative:

Too much music might cheapen players' experiences, or (possibly even worse) remove them from their immersion. Silence provides the negative space to ensure the music's emotional effect, but it is not in itself aesthetically significant. In fact, if anything, silence functions as something of a deterrent here; if players remain in one location too long without advancing the action, the music fades out 'to avoid getting annoying'.<sup>106</sup>

The placement of interactive audio seems important to player immersion and presence, but when forming a game world which is functionally empty of ally's, silence can highlight this sense of isolation to the players.

## **Modelling Worldbuilding: Environments and Characterisations**

As outlined by Cheng in relation to *LoTRO*, player immersion can be affected by the expected aural soundscape of a particular game genre and style (which can include player music and voices). This affects the sounds of both the game environments and the characters that exist within the fictional world. These environments and characterisations are thus a part of the worldbuilding process as their aural components are an integral part to a 'believable' fictional and virtual world.

---

<sup>103</sup> Turbine, *The Lord of the Rings Online*, Chicago: Midway Games (Microsoft Windows, OS X, 2007).

<sup>104</sup> Cheng, 'Sound Play', p. 130.

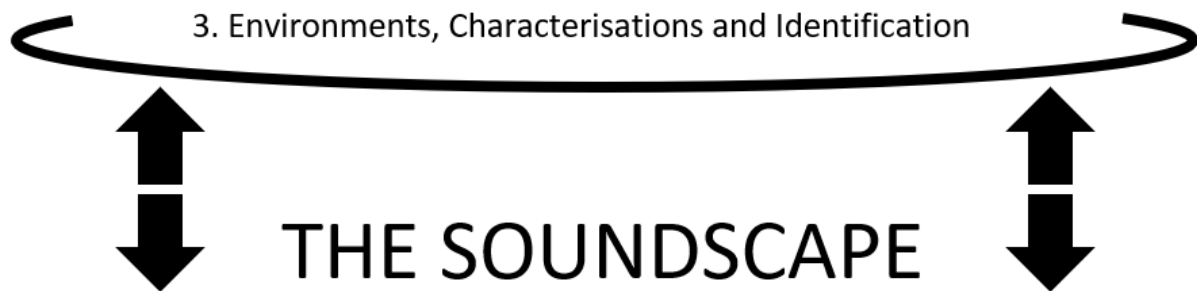
<sup>105</sup> SCE Japan Studio and Team Ico, *Shadow of the Colossus*, San Mateo: Sony Computer Entertainment (PlayStation 2, 2005).

<sup>106</sup> William Gibbons, 'Wandering Tonalities', in Donnelly, Gibbons, and Lerner, *Music in Video Games*, pp. 124–25.

In order to arrive at the worldbuilding elements of environments and characterisations, however, several components of worldbuilding must be established first. I have outlined a model that systematically considers the worldbuilding processes as discussed throughout this chapter. Figure 1 outlines this possible process of worldbuilding, beginning with the initial set up of technology to run the video game world. Diegesis and interfaces are essential to the functionality of a game. A diegesis needs an interface in order to exist, and the interface cannot show the audio, visuals, or gameplay features of a fictional world without an existing diegesis.

# 1. Video Game Diegesis and Interfaces

## 2. Diegesis and Outsourced Worldbuilding



*Figure 1: The model identifies the process in which worldbuilding occurs, eventually influencing or being influenced by the soundscape.*

As diegesis and interfaces are united, the diegesis worldbuilding components can start to build the world when engaged by the player; scholars such as Michel Kamp, however, wonder when the a game diegesis truly begins: when a familiar UI starts, when the game is initially turned on, when the initial menu opens, or when we first see visuals or hear music, for example.<sup>107</sup> Outsourced worldbuilding can begin as soon as audiences have access to the game's content. Therefore, outsourced worldbuilding may have already begun before the player has engaged with the world, through trailers and screengrabs used as marketing material.

The third point of the model finally considers the worldbuilding elements of environments and characterisations. Environmental worldbuilding, which I discuss in greater detail in Chapter four, considers the creation and aural enhancements of the world's landscapes, including aspects such as the terrain, weather, architecture, history, and the time period. Environments can remain the same

<sup>107</sup> Michel Kamp, 'Suture and Peritexts: Music Beyond Gameplay and Diegesis', in Kamp, Summers, and Sweeney, *Ludomusicology*, p. 73.

throughout an entire game or can change as the player progresses through narrative and gameplay; environment challenges in RPGs such as *The Legend of Zelda* series require the player to complete at least four different challenges and isolated narratives over various terrains such as grasslands, mountains, lakes, and deserts. Environmental worldbuilding works to signify where the world is situated, creating a soundscape that mimics 'everyday music' and sound in order to create a 'realistic' fictional world. The player's medium within these environments exist as their player-character or avatar, and any accompanying or non-playable-characters that exist in the world. The characterisations of these characters, further examined in Chapter five, are built through the lore of the world, creating perceived cultures. These characterisations, like the worldbuilding of a landscape, can vary in detail depending on the design decisions and priorities of a dungeon master or developer. Identification with these characteristics heightens the player's relationship and involvement with characters. Identification is included in the model with characterisations, as successful recreations of understandable characterisations can form player identifications with these characters, whether that is empathy or a sense of being. As the game world is built, and environments and characterisations worldbuild and heighten engagement, the soundscape is both the outcome and overall functionality of these components. The soundscape can both incorporate all these sounds within the diegesis, and act as a worldbuilding tool itself.

The model informs chapters four and five of this thesis which discuss environment, characterisation, and identification in relation to the worldbuilding processes of the role-playing video game. The model allows the thesis to scope itself when understanding the place of voices within the worldbuilding process. I will only be considering voices within the diegesis which includes sound and voices diegetic, non-diegetic, and fantastical gap considerations in the soundscape. As shown through Cheng's book, *Sound Play*, although player voices can be a part of the diegesis, their complex inclusion in the video game space requires separate theoretical and empirical study; the changes in perceptions and receptions of sounds in games can differ between solo and multiplayer games, considering the plethora of extra sounds and voices. Alongside player voices, I will be dismissing the sounds of technological interfaces sounds from this study. As each player's gaming set-up is different, so too are the perception of those technological interfaces and thus a theoretical study of these sounds would be difficult without empirical acknowledgements.

Although I am focusing on the diegesis, the thesis will still be informed by outsourced worldbuilding as I consult guides, YouTube comments, and fan-made sites such as Wikias. I am not analysing the effectiveness or importance of outsourced worldbuilding within this thesis (which would provide enough material for another thesis in itself), but the prevalence of fan-made content in the gaming industry provides unique insights into the reception of RPGs. The conclusions drawn from the model



will show the principal uses of the voice in role-playing video game diegeses and how they worldbuild environments and characterisations.

## Chapter Three: Voice, Vocalisations, and Voices in Multimedia

As discussed throughout Chapter two, the soundscape of an RPG is built by audio that exists both within the diegesis and outside of it. Focusing on the diegesis, the worldbuilding processes of environments and characterisations are created through the inclusion of reactive audio. This audio, in modern RPGs, frequently includes diegetic and non-diegetic voices to enhance these worldbuilding qualities through emotional vocalisations, dialogue, and performances. The myriad uses of voice and vocality however are influenced by mediums which have paved the way for sound on-screen, such as film and television. There are thus diverse definitions of the voice, vocalisations, and voices both across moving image musicology and further fields of research. In order to effectively define the voice in a suitable manner for RPGs, I will be drawing on multiple interdisciplinary studies throughout this chapter; these research areas include psychology, anthropology, narratology, semantic studies, and studies of identity and gender, alongside research into film and multimedia.

### Defining Voices

Voices are a part of everyday environments, ranging from natural conversation to vocal performances, vocal cries to singing in the shower. Whether through speech or vocalisations, voices can communicate information to another being by imparting semantic, emotional, or physical sounds; this communication of information is imperative as feedback for game players regarding their actions and reactions within a game.

‘Voice’ and ‘vocality’ are two terminologies which consider varied and defining manifestations of the voice, across a multitude of studies. The term vocality, and the act of vocalising, is explored by Leslie C. Dunn and Nancy Jones in their edited collection *Embodied Voices: Representing Female Vocality in Western Culture*. Dunn and Jones consider vocality as containing meaning, affected not solely through semantic content but through its actions as cultural agencies, sexual autonomies, and expressive freedoms.<sup>1</sup> Voice, on the other hand, considers language, semantic meaning, and the act of using the voice for communication, which Irina Leimbacher explains as ‘an utterance of speech,

---

<sup>1</sup> Leslie C. Dunn and Nancy A. Jones, ‘Introduction’, in *Embodied Voices: Representing Female Vocality in Western Culture*, ed. by Leslie Dunn and Nancy Jones (Cambridge, New York and Melbourne: Press Syndicate, 1994), p. 1.

empowered subjectivity, cinematic authorship, or political agency.<sup>2</sup> Whereas vocality is aligned with syllabic vocalisations, voice is associated with speech and empowerment, and therefore language. These assertions of voice and vocality are useful but are not entirely conclusive independently, voices can be considered with a greater complexity than exclusively language or non-language-based vocalisations. In her book *Queer Voices: Technologies, Vocalities, and the Musical Flaw*, Freya Jarman-Ivens defines the voice through a psycho-analytical study. The project, rather than focusing on quantitative or qualitative methodologies, outlines the voice's permeation into gendered spaces through external recording, production, and internal physiological vocal technologies.<sup>3</sup> When defining the voice Jarman-Ivens identifies two main functionalities, the 'material' and 'immaterial':

[...]there is the immaterial voice, the voice as an abstract potential, existing in the deepest psychological structures but nowhere else[...] In its material form[...], the voice is further to be distinguished from language, and this may be yet another place of paradox, in that the voice both serves and exceeds the semiotics and syntax of the spoken word; it articulates semiotic meaning and, in its bodily nature, offers both another dimension to that meaning and another meaning altogether. The (material) voice can be a mediator between body and language; it gives language meaning, in its inflections, its speed, its accent, its bodiliness, but it is also an object apart from language.<sup>4</sup>

Jarman-Ivens' project offers a link between vocalisations and voice as language. Voice is not confined solely to language and can transcend semantics through bodily gestures, personal vocal inflections, and cultural meanings such as accent. This disproves the constrained differences between voice as language and vocalisations as emotional content, reaffirming the position of voice as a complicated personal signifier. Furthermore, Jarman-Ivens' relationship with the voice, and the spaces 'queer voices' inhabit, is associated to the voice as identification; a listener understands their own voice as their identifier by recognising the differences of another speaker's vocal grain.<sup>5</sup> However, the voice we hear in our head has variations to other's perceptions of our voice, making it difficult to listen to one's own voice on a recording. Jarman-Ivens discusses this notion of the 'mind's ear', the notion of identifying with the 'correct' voice that we hear in our 'mind's eye', rather than what we may identify as 'incorrect'. If voices are identifiers of self, this nuance of 'correct' and

---

<sup>2</sup> Irina Leimbacher, 'Hearing Voice(s): Experiments with Documentary Listening', *Documentary Audibilities*, 39.3 (2017), 292–318 (p. 292).

<sup>3</sup> Freya Jarman-Ivens, *Queer Voices: Technologies, Vocalities, and the Musical Flaw* (Basingstoke: Palgrave Macmillan, 2011), pp. vii–viii.

<sup>4</sup> Jarman-Ivens, 'Queer Voices', p. 4.

<sup>5</sup> Jarman-Ivens, 'Queer Voices', p. 33.

‘incorrect’ voices may cause a juxtaposition between what we identify as our vocal signifiers and what listeners hear.

Listeners may thus identify with manifestations of a voice that is separate to what the speaker deems their identifiers. This can affect greater cultural associations of the voice as Katherine Meizel’s paper ‘A Powerful Voice: Investigating Vocality and Identity’ outlines:

When we investigate vocality, then, attention to cultural meaning is more urgent than ever, so that we hear the voice as integrated with the body, and the embodied voice as a crucial component in how we listen and speak to the world around us, to each other, and to ourselves.<sup>6</sup>

Through the history of vocality and vocal identification of race, culture, and ‘the black voice’, Meizel concludes her discussion of vocalisations as identifiers by considering voice as a way of communicating with the world and commenting on a sense of self.<sup>7</sup> As the voice can interact with the world through cultural signifiers and personal identification, so too can it have a meaningful interaction with the performance space. Susan Rutherford’s chapter, ‘Voices and Singers’, in *The Cambridge Companion to Opera Studies* identifies the voice in the performance space as ‘both timbre (the property of the vocal instrument) and the action (the manipulation of the instrument).’<sup>8</sup> Rutherford considers the voice as its unique, personal, properties alongside the physiological creation of vocal performance. It was Roland Barthes’ seminal essay, ‘The Grain of the Voice’, which initially provided these defining characteristics of voice and vocality as performance, suggesting voices have personal properties which are considered as cultural signifiers and communication. The grain of the voice is explained by Barthes as the body’s materiality.<sup>9</sup> Barthes translates Julia Kristeva’s terms ‘pheno-text’ and ‘geno-text’ into associations of performativity as ‘pheno-song’ and ‘geno-song’;<sup>10</sup> ‘pheno’ refers to the display of characteristics and traits of something, whereas ‘geno’ refers to characteristics formed from through genetic makeup. Barthes’ ‘pheno-song’ consists of ‘everything which, in the performance, is at the service of communication, representation, expression, everything which it is customary to talk about, which forms the tissue of cultural values.’<sup>11</sup> ‘Pheno-song’ therefore covers the identifying characteristics of a sung performance including language, communication, and expression through semantics. The ‘geno-song’ is referred

---

<sup>6</sup> Katherine Meizel, ‘A Powerful Voice: Investigating Vocality and Identity’, *Voice and Speech Review*, 7.1 (2011), 267–74 (p. 272) <<https://doi.org/10.1080/23268263.2011.10739551>>.

<sup>7</sup> Meizel, ‘A Powerful Voice’, p. 269.

<sup>8</sup> Susan Rutherford, ‘Voices and Singers’, in *The Cambridge Companion to Opera Studies*, ed. by Nicholas Till, Cambridge Companions to Music (Cambridge: Cambridge University Press, 2012), p. 117.

<sup>9</sup> Roland Barthes, ‘The Grain of the Voice’, *Image Music Text*, 1977, 179–89 (pp. 181–82).

<sup>10</sup> Barthes, ‘The Grain of the Voice’, pp. 181–82.

<sup>11</sup> Barthes, ‘The Grain of the Voice’, p. 182.

to simply as ‘the volume of the speaking and singing voice’ and ‘the *diction* of the language.’<sup>12</sup> ‘Geno-song’ thereby comments on the genetic makeup of the voice through its stylistic manifestations, such as volume and diction, to create a voice’s characteristics. These definitions again break voices into two aspects, ‘pheno’ for communication and ‘geno’ for signifiers. Sutured, ‘pheno-song’ and ‘geno-song’ can provide both communicative and identifying information to an audience which connects to, but also surpasses, its semantic content. Isabelle Peretz in Patrik N. Juslin and John Sloboda’s *Handbook of Music and Emotion* also considers this separation of communication between tone (grain) and emotional content. Peretz’s chapter, ‘Towards A Neurobiology of Musical Emotions’, identifies through a neurobiological inquiry that vocal meaning and perception ‘can be conveyed by tone of voice in speaking, or by emotional vocalizations such as laughs, cries, and screams.’<sup>13</sup>

When attempting to define voices as clearly as possible, it is worth considering the psychological and cognitive studies which discuss its properties. These research areas conduct empirical studies to further discern the physical creation of the voice and to understand human reactions, and perceptions, towards voices within various scenarios and situations. Key texts such as Diana Deutsch’s edited collection, *The Psychology of Music*, provide a holistic overview of voices within a multitude of manifestations including speech, singing, and vocalisations, focusing on the perception, understanding, and performance of music.<sup>14</sup> Johan Sundberg’s chapter in-particular discusses the ‘Perception of Singing’ through vocal functions like timbre, vibrato, vocal resonances and a listener’s perception of these aspects.<sup>15</sup> The perception of voice and its functionality alongside music, from the perspective of the human mind, is also deliberated by Steven Mithen in their book *The Singing Neanderthals*.<sup>16</sup> Mithen’s anthropological study provides research into the evolution of the voice, and the brain’s function when working with music and language together, language without music, and music without language. The development of research into brain damage has shown how the brain individually works with music and language through a complex system of neuron networks. Aniruddh D. Patel’s chapter, ‘Music and the brain Three links to language’, in the edited collection *Oxford Handbook of Music Psychology* explains that the ‘research into brain damaged patients has been particularly important in demonstrating that music involves a neurocognitive system with

---

<sup>12</sup> Barthes, ‘The Grain of the Voice’, pp. 182–83.

<sup>13</sup> Isabelle Peretz, ‘Towards A Neurobiology of Musical Emotions’, in Juslin and Sloboda, p. 115.

<sup>14</sup> *The Psychology of Music*, ed. by Diana Deutsch, 3rd edn (California: Academic Press, 2012).

<sup>15</sup> Johan Sundberg, ‘Perception of Singing’, in Deutsch, *The Psychology of Music*, chap. 3.

<sup>16</sup> Steven Mithen, *The Singing Neanderthals: The Origins of Music, Language, Mind and Body* (London: Weidenfeld & Nicolson, 2006).

many components.<sup>17</sup> Understanding aphasia and amusia,<sup>18</sup> is imperative to recognising music and language's separation within the brain.

The separation of music and language in the brain is akin to the division of voice and vocalisations, language and emotions. The construction of scientific and empirical studies, therefore, is of a particular interest when examining the creation of the voice and its perceptions. Many of these studies focus on voice as emotion alongside music, such as Escoffier, Zhong, Schirmer, and Qiu. Escoffier et al. test the differences of musical and vocal representations of emotion, processed by the brain.<sup>19</sup> Escoffier et al. conclude that although perceptions of emotional content between individuals may vary, due to existing experiences with certain emotions, emotional expressions created by both voice and music are 'comparable'.<sup>20</sup> This study would suggest that although voice and music can be separated by impairments such as aphasia or amusia, emotions examined across neurobiological studies still intertwine the two. This intertwining of emotional identifiers can connect the performative voice, vocalisations, and voice as language. The separation of specific emotions, however, is not provided on the level of music and voice, but the difference between voice as language and vocalisations. Frühholz, Trost, and Grandjean consider in their article, 'The role of the medial temporal limbic system in processing emotions in voice and music', that non-verbal vocalisations are a signifier of simple emotions, such as happy and sad, whereas socially complex emotions, such as doubt and nostalgia, are created through voice and speech.<sup>21</sup>

These empirical insights into the cognitive crossovers and differences of vocal perceptions are informative to theoretical processes that consider voices within further critical implications; these studies interweave with the scholarly considerations of music and audience perceptions of culture, gender, and identity as discussed by Jarman-Ivens and Meizel. This lays the foundation for wider considerations of the voice and music within mediums such as television, film, and video games. I will be focusing this thesis on a critical analysis of voices in video games, as opposed to empirical study, but these scientific fields of study are pertinent to an underlying understanding of the voice and its manifestations. Research which highlights the voice as a part of the moving image further

---

<sup>17</sup> Aniruddh D. Patel, 'Music and the brain Three links to language', in Ian Cross, Susan Hallam, and Michael Thaut, *The Oxford Handbook of Music Psychology*, *The Oxford Handbook of Music Psychology*, 2008, p. 208 <<https://doi.org/10.1093/oxfordhb/9780199298457.001.0001>>.

<sup>18</sup> See Mithen, 'The Singing Neanderthals', chaps 3–4.

<sup>19</sup> Nicolas Escoffier, Jidan Zhong, Annett Schirmer, and Anqi Qiu, 'Emotional Expressions in Voice and Music: Same Code, Same Effect?', *Human Brain Mapping*, 34.8 (2013), 1796–1810 (p. 1798) <<https://doi.org/10.1002/hbm.22029>>.

<sup>20</sup> Escoffier, Zhong, Schirmer, and Qiu, 'Emotional Expressions in Voice and Music', p. 1807.

<sup>21</sup> Sascha Frühholz, Wiebke Trost, and Didier Grandjean, 'The Role of the Medial Temporal Limbic System in Processing Emotions in Voice and Music', *Progress in Neurobiology*, 123 (2014), 1–17 (p. 3) <<https://doi.org/10.1016/j.pneurobio.2014.09.003>>.

refine the project's scope and relate it to pre-ludomusicological discussions. This inclusion of texts outside of ludomusicology, but within moving image and voice studies, provides the project with a holistic understanding of the foundational analytical considerations of the voice; thus part of this chapter considers the various instances (and absences) of the voice in the moving image and its related academic study.

Where video game and moving image voices are concerned, voices take at least two distinct styles. Voice acts as information, focusing on language that feeds back and communicates with audiences, players, and characters. Vocality, on the other hand, is used to identify realisms such as breathing, movement, and action excursions, alongside what Frühholz, Trost, and Grandjean consider as simple emotions, such as fear. The dearth of literature surrounding voices has pointed towards its breakdown into these two considerations, voice and vocality, and video games mirrors these distinctions. I thus align the term voice with that of language and meaning alongside intonation and accent, not ignoring the body. This incorporates both languages initiated by the player and language which is not triggered by direct player interaction. Player activated language can be seen through the expressions of the protagonist and player-character, Kratos, in the action-adventure game *God of War*.<sup>22</sup> The game, through the player's interactions, communicates with both the player and a semi-playable-character<sup>23</sup> through the consistent use of the semantic word 'Boy'. 'Boy' is used by the Kratos to signal his son Atreus to complete an action such as reading lore markers across the world or helping with an environmental puzzle. As the player presses the action button for Atreus, Atreus almost always responds immediately to the semantic command, communicating to the player that their input has been recognised. Alternatively, language is used in *God of War* outside of player input as Kratos and Atreus talk and respond to each other without the inclusion of 'Boy'. Atreus frequently talks to himself and Kratos during the journey, especially during boat journeys where Atreus convinces Kratos to tell stories. The use of semantic language as both an identifier for player interaction alongside providing extra content, through character agency and conversations, worldbuilds the game by creating the feeling that the game world is larger, older, and more alive and populated than the player can initially sense.

References to the term 'vocality' and the act of 'vocalising' will not consider language but will focus on vocal manifestations which can highlight non-language-based meaning, such as emotions; this

---

<sup>22</sup> Santa Monica Studio, *God of War, San Mateo: Sony Interactive Entertainment* (PlayStation 4, 2018).

<sup>23</sup> The player has some control over the secondary character, Atreus, as they can command the use of their bow and bow abilities, alongside asking them to read runes along the journey, but cannot physically move the character; the player can only command these actions through a single button (square) on the PlayStation 4 controller.

includes the sounds of infant humans, or non-human sounds. The artistic adventure games *Journey*<sup>24</sup> and *ABZÛ*<sup>25</sup> allow players to trigger the vocalisation of their player-character in order to communicate with the world or express enjoyment within the game. Although these vocalisations do not carry semantic values, the player uses these sounds to engage with the world, solve environmental puzzles, and in the case of *Journey* engage with other players when connected to the internet. These vocalisations are triggered entirely by the player through the press of a button. However, some video game vocalisations, such as player-character action exertion sounds, are created by player interaction but not action. The voice of the player-character and protagonist, Link, in *The Legend of Zelda (Zelda)* series clearly outlines vocalisations through general player interactions. The absence of the recorded voice means that the main voices of the games are the vocalisations of Link's movements, through sounds such as 'Hyah', 'Hup', 'Hyup', 'Hai' and further vocalisations. The player only causes these vocalisations by moving Link, rather than through a dedicated vocalisation button on their controller. Therefore, the main expression of vocalicity in the game is created by the relationship between player and player-character. The lack of the recorded voice in most *Zelda* games highlights the few recorded voice lines of Link's fairy companion Navi ('Listen', 'Hey', 'Hello', 'Watch Out', 'Look'), in *The Legend of Zelda: Ocarina of Time*,<sup>26</sup> which directs the player's attention towards Navi's instructions.

Voice and vocalisations offer varying information to listeners, however there are instances where voice and vocalicity convey similar or the same meanings. In this instance I will be using the plural 'voices'. I will also be using this term when there are multiple uses of voice and vocalicity in the same space that cannot or do not need to be deciphered. Most video games use numerous voices when forcing the perceptive of large volumes of beings, whether that is within gatherings or dangerous moments such as the panic of a large herd. The zombie apocalyptic-action adventure game, *The Last of Us*,<sup>27</sup> uses a combination of voice and vocalisations to identify humans and zombies, alongside the action exertion of the player-character, Joel. There are various zombies in the world which exude various vocalisations, such as clicks and screams, to identify to the player the difficulty level of the zombies they interact with. Humans and human enemies are highlighted through their voices and the use of language, mainly American English, as they talk to their allies and goad the player-character into revealing themselves. When fighting human enemies, vocalisations occur through

---

<sup>24</sup> Thatgamecompany, *Journey, San Mateo: Sony Computer Entertainment* (PlayStation 3, PlayStation 4, Microsoft Windows, iOS, 2012).

<sup>25</sup> Giant Squid Studios, *ABZÛ, Milan: 505 Games* (Microsoft Windows, PlayStation 4, Xbox One, Nintendo Switch, 2016).

<sup>26</sup> Nintendo EAD, *The Legend of Zelda: Ocarina of Time, Kyoto: Nintendo* (Nintendo 64, GameCube, iQue Player, 1998).

<sup>27</sup> Naughty Dog, *The Last of Us, San Mateo: Sony Computer Entertainment* (PlayStation 3, PlayStation 4, 2013).



movement, attack, and damage exertions from both the player-character and the human enemies. During most occasions containing zombies, however, voice and vocalisations act together as Joel and his companion, Ellie, communicate through language, independently of player action, whilst zombie vocalisations can be heard.

## **Voices in Multimedia**

The myriad implementations of language-based voice and the vocal manifestations of vocality on-screen are imperative to the worldbuilding process across multimedia. The quality of actor voices can affect the believability of a character and the reality of a diegesis altogether. The synchronisation of voices to their on-screen bodies is imperative to this suspension of disbelief for the audience. However, before voices were introduced on-screen, both film and video games went through a time-period where voices were effectively silenced due to technical constraints. Both film and games included voices as text on-screen during their non-vocal timeframe but as video games already had access to sound, they could at least create the impression of voices through sound chip technologies. However, when voices were implemented on-screen for all multimedia, the demands made of actors and vocal artists increased.

### ***The Demands of Voice Acting***

Motion capture is a recent and popular tool that records an actor's bodily and facial movements which are then imposed onto an on-screen character model, providing realistic movements and facial expressions. Alongside this capture of movement, voices are recorded simultaneously and actively synchronised to character motions. 'Vocal fidelity', or the synchronisation of vocal sound to bodily movements, is important to western films and audiences. A bad synchronisation of the voice acting or dubbing of a character can be distracting to audiences, possibly reducing immersive tendencies. Synchronisation was important during the move from silent-era cinema to 'talkies' in the 1920s and 1930s, and the move from text-box video game dialogue to recorded voices in the early 2000s. Gibbons recounts an anecdotal experience of playing Japanese role-playing games (JRPGs) with a group of friends, in his edited book *Music in the Role-Playing Game: Heroes & Harmonies*. At the time when JRPGs had no voice acting, Gibbons and his group happily acted out the voice lines themselves, reading through the text boxes in turn. This changed when the group played a partially

vocalised remaster of the game *Lunar: Silver Star Story Complete (Lunar)*.<sup>28</sup> In Gibbons' anecdote, the group thus ceases their narration of characters because of this inclusion of the voice, leading Gibbons (through his reading of *The Voice in Cinema*) to draw parallels of his experience with Chion's argument that the voice can restrain an individual's creativity.<sup>29</sup> Many modern video games choose to be a silent or part-talkie game, like *Lunar*, as developers take influence from other aspects of this transitional period, such as the use of retro and pixel aesthetics; this aesthetical choice may be for reasons such as to appeal to nostalgic gamers, to save costs, or because the developer simply enjoys the style. For example, *Octopath Traveler*<sup>30</sup> uses high definition pixel aesthetics and a fully orchestrated soundtrack, combining nostalgia and modern technological advances. Either for stylistic or budget reasonings, *Octopath Traveler* is a part-talkie turn-based RPG which only provides the player with voices during important story-telling scenes; otherwise, the player must read silent textboxes whilst the characters make short vocalisations to signify the beginning of a sentence. Video games are naturally nostalgic for some players due to the quick succession of technological advancements during the 1980s and 1990s. Alongside this, loyalty to a series played during childhood can maintain this nostalgia for players as some game franchises have continued to make games since the 1980s (e.g. *Final Fantasy*). Experiencing games which remind players of the past, such as Gibbons' description of playing *Lunar*, is summarised by Tim Wulf, Nicholas D. Bowman, John A. Velez, and Johannes Breuer:

[...] playing such games may serve as connecting links between adults and their childhood. This observation is critical, because it directly relates to the potential for game players to accrue enough experiences to feel a sense of nostalgia when looking to recreate and reconnect with games from their past.<sup>31</sup>

Wulf, Bowman, Velez, and Breuer's assertion that games can directly link players to the past thus explains the aesthetical choice of some retro style games. This dominant sense of nostalgia thus informs the decisions of some retro games to recreate pre-transitional 'silent-era' voices, attractive to certain types of gamers.

---

<sup>28</sup> Game Arts and Japan Art Media, *Lunar: Silver Star Story Complete*, Tokyo: Kadokawa Shoten (Sega Saturn, PlayStation, Microsoft Windows, iOS, 1996).

<sup>29</sup> William Gibbons, 'Song and the Transition to 'Part-Talkie' Japanese Role-Playing Games', in *Music in the Role-Playing Game: Heroes & Harmonies*, ed. by William Gibbons and Steven Reale (New York, London: Routledge, 2020), pp. 9–10.

<sup>30</sup> Square Enix and Acquire, *Octopath Traveler*, Tokyo: Square Enix (Nintendo Switch, Microsoft Windows, 2018).

<sup>31</sup> Tim Wolf, Nicholas D. Bowman, John A. Velez, and Johannes Breuer, 'Once Upon a Game: Exploring Video Game Nostalgia and Its Impact on Well-Being', *Psychology of Popular Media Culture*, 2018, 1–18 (p. 2) <<https://doi.org/10.1037/ppm0000208>>.

As video games acknowledge their silent history through aesthetic decisions, films consciously acknowledged (but not often recreated) this uncomfortable transition within some film narratives, most notably within the narrative of *Singin' in the Rain*.<sup>32</sup> *Singin' in the Rain* famously brings attention to the dubbing era and the technical transitions which brought about difficult changes to the 'coming of sound', including the introduction of on-screen sound and voices. Alongside a difficult technical transition, silent film actors struggled with adapting their voices to the screen rather than solely focusing on bodiliness. The challenges of technological and aesthetical changes, voice acting, and audience receptions of the sound era are embedded in *Singin' in the Rain's* narrative. The character Lina Lamont, performed by Jean Hagen, was a notable silent film actress famous for her looks and silent era acting. When the sound era begins in the film, Lina lends her voice to the screen but comically cannot point her voice towards the microphone; the sudden inclusion of microphones on set meant that actors needed to adapt to directing their voices towards a microphone, a skill which Lina Lamont clearly does not adopt. Alongside her problematic acting, Lina's voice is not aesthetically pleasing because of its strong New York accent alongside shrill and comically child-like qualities. Since Lina's voice is considered unpleasant, the narrative introduces the character Kathy Selden who is performed by Debbie Reynolds. Kathy Selden is a juxtaposition to Lina as her voice is pleasing for the sound era, but she has little prior screen experience, thus the narrative (focusing solely on vocal talents) allows Kathy to become a new star through the exiling of Lina. When the in-film cinematic *The Dueling Cavalier* has its preview screening, the character Don Lockwood, performed by Gene Kelly, and Lina become out of synch with their voices, triggering laughter from the audience. Lina's voice, however, is the focal point of this judgement due to its unappealing qualities. Thus, when Lina is asked to perform live on-stage, the producers live dub Lina's voice with a performance by Kathy behind the theatre's on-stage curtains. When Kathy is inevitably uncovered as the performer, by other characters who wish for her stardom, there is outrage from the audience that Lina has been dubbed and thus Kathy finds her fame.

The demands on an acting career in the sound era transition is highlighted clearly in *Singin' in the Rain*. The transition to 'talkie' cinema outlined both synchronisation issues and concerns that certain silent-era actor voices would not suit the screen. Video games mirrored this transitional issue. The transition to voice acting from synthesised sounds suffered equally with frequent issues of bad voice acting, writing, voice direction, money constraints, and rushed translations. Although the quality of voice acting can be improved with time and funding, if voices are not synchronised to on-screen characters then suspensions of disbelief can be broken. If a character's ancillary gestures, such as

---

<sup>32</sup> The plot here acts as anecdotal evidence of the issues of synchronisation and placement of microphones when voice and sound first came to screen.

mouth and bodily movements, do not synchronize with voice and vocalisations made, the voice can become disruptive rather than immersive. The non-synchronised voice remains both comical and frustrating, outlined by *Singin' in the Rain*, as it is a clear failing of the technology which should be providing characters with 'realism' through the voice.

### **Dubbing**

When conducting her own reading of *Singin' in the Rain*, Liz Greene speaks of the control over the female voice in Hollywood cinema through the handling of recorded female voices. Greene argues that 'it would be a futile exercise to seek out an 'authentic' voice in the cinema as soundtracks are often made up of numerous takes, re-voiced or dubbed'.<sup>33</sup>

The mapping of voices to character models in video games and animation mirrors this process of re-voicing, dubbing, and the use of automatic dialogue replacement (ADR) in film. ADR is a post-production process that re-records dialogue in a studio if it was unclear during the original recording:

Dialogue editors work to remove the filmmaking from the film. Dolly squeaks, camera noise, crew rustling, and light buzzes must go; otherwise, the magic of the movies is compromised. They help present the actors in their best light, quieting dentures, eliminating belly noises, and sobering slurred syllables. And when the production sound can't be saved, the dialogue editor is involved in the ADR process, that is, the rerecording of voices in the studio to replace problem field recordings or to beef up performances.<sup>34</sup>

This post-recorded dialogue is mapped to the on-screen character to ensure a close synchronous relationship – a 'better' vocal performance – that tricks the audience into believing this was the original audio; this leads Chion to refer to the voice as 'privileged over other soundtrack elements'.<sup>35</sup> As ADR re-records the same actor over their original performance, dubbing records a new voice over the facial movements of an on-screen actor or animated character, often mapped to a different language. The issues faced by actors here is the synchronisation of a new voice over the facial movements of an alien body, and possibly language.

---

<sup>33</sup> Liz Greene, 'Speaking, Singing, Screaming: Controlling the Female Voice in American Cinema', *The Soundtrack*, 2.1 (2009), 63–76 (p. 65) <<https://doi.org/10.1386/st.2.1.63/1>>.

<sup>34</sup> John Purcell, *Dialogue Editing for Motion Pictures : A Guide to the Invisible Art* (Focal Press, 2014), p. 1.

<sup>35</sup> Michel Chion, *The Voice in Cinema*, trans. by Claudia Gorbman (Chichester: Columbia University Press, 1999), p. 81.

Whilst western cinema generally prioritises (the appearance of) synchronicity between ‘voice’ and ‘body’, the dubbing of characters and the desynchronised voice is accepted and celebrated within Bollywood cinema. Budhaditya Chattopadhyay discusses the ‘dubbing era’ of Bollywood from 1950-1990. During this era in Bollywood history, voices were dubbed through ADR in order to sufficiently reduce extra, unwanted, sounds when recording on-location:

The typical audio practices in this specific era of Indian cinema led the spectator-listener to imagine the pro-filmic space or the site of film sound. This was facilitated by a minimal use of ambient sound, but relied heavily on the voice and background musical score, song, and dance sequences along with highly processed sound effects to carry out the visual narrative.<sup>36</sup>

These added processes, including reverb, alongside a lack of ambient location sound would be an issue for western audiences due to this lack of fidelity to the landscape. According to Ajay Gehlawat, however, realism is not the goal in Bollywood film but rather to locate the audience within a romanticised India:

The frequent song and dance numbers, for instance, are ‘unspoken’ in a double sense: narratively (i.e., as departures from the realm of actual discourse) and structurally (i.e., via their use of playback). Additionally, the song and dance sequences can be seen as instances of ‘interior’ stories, as they reflect, often, the ‘fantasies’ of the characters. Furthermore, one would have to note all the shared structural aspects of cinema (cinematic literacy) that are used to convey characters’ ‘inner states’: the point-of-view (POV) shot; mood music; plus, more generally speaking, camera movement and angles (i.e., ‘framing’).<sup>37</sup>

There is thus a juxtaposition in Bollywood cinema between diegetic fidelity and romanticising the on-screen narrative, actors on-screen, and playback singers. During the diegetic performances within Bollywood cinema, playback singers are used to dub the vocal performance of the on-screen actor by mimicking the unique vocal characteristics and identifiers of the diegetic actor voice. Deepali Y Loni and Shaila Subbaraman explore the playback singers of Bollywood and their abilities to ‘modulate their voice according to the category of the song, the situation in the movie and most important according to the actor or actresses to whom they are lending their voice.’<sup>38</sup> The modulation of the playback singers’ voices enhances certain qualities of the actor’s voice to

<sup>36</sup> Budhaditya Chattopadhyay, ‘The Auditory Spectacle: Designing Sound for the ‘Dubbing Era’ of Indian Cinema’, *The New Soundtrack*, 5.1 (2015), 55–68 (p. 65).

<sup>37</sup> Ajay Gehlawat, *Reframing Bollywood: Theories of Popular Hindi Cinema*, *Reframing Bollywood: Theories of Popular Hindi Cinema* (London: SAGE Publications Ltd, 2010), pp. 56–57.

<sup>38</sup> Deepali Y Loni and Shaila Subbaraman, ‘Robust Singer Identification of Indian Playback Singers’, *Audio, Speech, and Music Processing*, 2019.10 (2019), 1–14 (p. 9) <<https://doi.org/10.1186/s13636-019-0153-0>>.

romanticise the musical numbers of Bollywood film. The vocal fluidity between actors and playback singers questions the ownership of the voice as the position of the playback singers is to construct another's vocal identity. The vocal uniqueness of human voices is further examined by Jelena Novak, through Adriana Cavarero's understanding of vocal expressions, who considers voice as a unique identifier of a person's body, due to the distinctiveness of bodily structures.<sup>39</sup>

The act of dubbing can thus bring forward the discussion of vocal ownership. The ownership of the voice is outlined through Ben Macpherson's reading of Mladen Dolar's psychoanalytical text, *A Voice and Nothing More*,<sup>40</sup> outlining Dolar's schema which suggests that neither body nor language can take complete ownership of the voice. Macpherson considers the voice through Dolar's work as 'a kind of dissociative limbo that is neither linguistic nor corporeal and yet which to an extent relied on the properties of both language and the body.'<sup>41</sup> This disassociated voice is attached to character models in animated television, film, and video games. The voice in video games and animated film and television is dubbed to the ancillary gestures of the on-screen character, a secondary body to the original source of the voice. In order for a voice to seemingly be a part of an animated character, the voice must be disembodied from its source and re-embodied with that character. This means the voice has the possibility of becoming a shared signature of both the voice actor and character. Rebecca Coyle identifies this signature voice, or sound in some cases, as the anthropomorphism of the character, 'whether they be the spiders and maggots in Tim Burton's films or the mutating houses and creatures in Hayao Miyazaki's films.'<sup>42</sup>

## **Animation**

As the majority of video games are animated, the need for voice actors with an aptitude for dubbing grew. Dubbing and voice acting in animated television, film, and video games, however, incorporates both professional voice actors alongside all-star casts. The movement towards creating unique character voices is outlined by Jon Fitzgerald and Philip Hayward when discussing the animated film, *The Brave Little Toaster*.<sup>43</sup> Fitzgerald and Hayward examine the movement away from using the popular cartoon voice actors of the time, as the director of *The Brave Little Toaster*, Jerry Rees,

---

<sup>39</sup> Jelena Novak, 'Voices beyond Corporeality: Towards the Prosthetic Body in Opera', *Studies in Musical Theatre*, 6.1 (2012), 73–88 (pp. 75–76) <[https://doi.org/.libaccess.hud.ac.uk/10.1386/smt.6.1.73\\_1](https://doi.org/.libaccess.hud.ac.uk/10.1386/smt.6.1.73_1)>.

<sup>40</sup> Mladen Dolar, *A Voice and Nothing More* (MIT Press, 2006).

<sup>41</sup> Ben Macpherson, 'A Voice and so Much More (or When Bodies Say Things That Words Cannot)', *Studies in Musical Theatre*, 6.1 (2012), 43–57 (p. 45).

<sup>42</sup> Rebecca Coyle, 'Introduction: Audio Motion: Animating (Film) Sound', in *Drawn to Sound: Animation Film Music and Sonicity*, ed. by Rebecca Coyle (London: Equinox Publishing Ltd, 2010), pp. 13–14.

<sup>43</sup> Jerry Rees, *The Brave Little Toaster*, Hyperion Pictures (David Newman, 1987).

considered their voices as too recognisable. Rees wanted to branch out from the usual animation actors to a cast that would create unique vocal identities for each character in the film.<sup>44</sup> This movement created change in the film industry, outlined by Fitzgerald and Hayward as they consider ‘features such as Lasseter’s *Toy Story* and a slew of more recent feature films have used known lead vocal actors to lend depth of characterization and identification (and saleability) to the projects.’<sup>45</sup> Alongside these lead vocal actors, Colleen Montgomery outlines the argument for the use of an all-star cast in *Pixar* animations, as studios wanted to be considered as ‘critically and commercially successful as live-action cinema.’<sup>46</sup> The use of the all-star voice created a recognisable anchor alongside playing a key role in promoting material.<sup>47</sup> As a star voice can sell the initial product (Tom Hanks as the character Woody in *Toy Story*<sup>48</sup>), voices are then recognisable within future products and can provide a recognisable anchor for a single company, such as John Ratzenburger’s continued use in *Pixar*’s feature films. Star appearances themselves are more common in high budget modern day animated cinema.<sup>49</sup> This recognisable anchor, although can aid the outsourced worldbuilding of a fictional world, can recreate the issue Jerry Rees attempted to change of including voices that are not unique to the characters but the actor themselves. When dubbing eastern video games, it is more common for professional voice actors who are experienced with dubbing Japanese animation, like the experience of playback singers. The animation of characters in eastern video games is often mapped to the ancillary movements of the original language. This means that when voice actors approach the characters, they must tailor their vocal performances to the on-screen mouth movements of a different language. Therefore, when all-star casts are chosen for these translations, seen especially in some higher budget or AAA releases, the quality of vocal performances suffer due to a lack of experience.

When casting the all-star voice or the professional voice actor, Sarah Kozloff identifies the physical differences which affect the voice:

The casting of an actor assigns dialogue-as-written to a person, a body, a voice. Different voices— with all their physical individuality and all their markers of age, gender, ethnicity,

---

<sup>44</sup> Jon Fitzgerald and Philip Hayward, ‘Resilient Appliances: Sound, Image and Narrative in *The Brave Little Toaster*’, in Coyle, *Drawn to Sound*, p. 164.

<sup>45</sup> Fitzgerald and Hayward, ‘Resilient Appliances’, in Coyle, *Drawn to Sound*, p. 172.

<sup>46</sup> Colleen Montgomery, ‘Pixarticulation’, *Music, Sound, and the Moving Image*, 10.1 (2016), 1–23 <<https://doi.org/10.3828/msmi.2016.1>>.

<sup>47</sup> Montgomery, ‘Pixarticulation’, p. 5.

<sup>48</sup> John Lasseter, *Toy Story, Walt Disney Pictures Pixar Animation Studios* (Buena Vista Pictures Distribution, 1995).

<sup>49</sup> For example: *Finding Nemo* – Ellen DeGeneres; *Monsters Inc./Emperor’s New Groove* – John Goodman; *The Lion King* – James Earl Jones, Jeremy Irons, Rowan Atkinson; *Frozen* – Seth Rogan, Idina Menzel; *Moana* – Dwayne Johnson.

experience— will give dialogue different nuances[...]The voices of famous movie actors are instantly recognizable, and intimately interwoven with viewers' conceptions of their personae.<sup>50</sup>

As outlined here by Kozloff, the star voice can be instantly recognisable to audiences who are aware of their voice. When all-star voices are used in western video games, their inclusion can characterise characters with identities that match the styles of roles actors are previously associated with. The horror adventure game *Until Dawn*<sup>51</sup> includes the vocal and motion capture talents of Rami Malek as a deviant and trickster personality, whilst *Fable II*<sup>52</sup> uses the voice of Stephen Fry as the deviant yet clever character, Hero Reaver. *Sid Meier's Civilization* series uses popular actors as narrators who introduce the player to their civilisation, any upgrades to their progress, and unique environments such as Mount Kilimanjaro. The voice of Leonard Nimoy in *Civilization IV* and Sean Bean in *Civilization VI* become vocal identifiers to the player of their current progress; the use of star voices is a novelty in the game, but the voices also act as authority by familiarity. Sean Bean's voice over in *Civilization VI* relates to his acting roles in historical and fantasy genre television and film, such as the *Sharpe*<sup>53</sup> television series and *The Lord of the Rings: The Fellowship of the Ring*.<sup>54</sup> These genre based and recognisable voices are affected by an actor's unique vocal 'grain'. Thus, the ownership of the voice prior to its inclusion on-screen can change the assumption of identifying features of the voice by audiences.

### ***Voices as Environment and Characterisations***

Whereas the all-star voice can be used as a recognisable anchor, Gorbman suggests that in modern cinema this 'recognizable voice need no longer prevail.'<sup>55</sup> Releasing the voice from its contract as a recognisable anchor allows the voice to be manipulated by actors such to identify characters, rather than their personas. Gorbman outlines this movement towards vocal manipulation in her paper 'The Master's Voice', drawing on case studies which highlight actors who manipulate their vocal

<sup>50</sup> Sarah Kozloff, *Overhearing Film Dialogue* (London: University of California Press, Ltd., 2000), p. 91.

<sup>51</sup> Supermassive Games, *Until Dawn, San Mateo: Sony Interactive Entertainment* (PlayStation 4, 2015).

<sup>52</sup> Lionhead Studios, *Fable II, Redmond: Microsoft Game Studios* (Xbox 360, 2008).

<sup>53</sup> Tom Clegg, 'Sharpe' (ITV, 1993-1997).

<sup>54</sup> Peter Jackson, *The Lord of the Rings: The Fellowship of the Ring, WingNut Films* (New Line Cinema, 2001).

<sup>55</sup> Claudia Gorbman, 'The Master's Voice', *Film Quarterly*, 68.2 (2014), 8–21 (p. 9).



identifiers for their on-screen characters. Gorbman describes the vocal manipulation of Philip Seymour Hoffman in *The Master*<sup>56</sup> to clearly embody his role as Lancaster Dodd:

Dodd's voice inhabits the full range of cinema's three types of audio—speech, music, and noise—as Hoffman marshals every nuance of intonation, tempo, rhythm, timbre, volume, articulation, eccentric pronunciation, and silence, as well as facial expression and gesture, to constitute this voice, including an inventory of coughs, gasps, grunts, and varieties of audible breaths. The character possesses the voice of authority, speaking with a clear, resonant midcentury American diction.<sup>57</sup>

The manipulation of Hoffman's vocal 'grain' characterises the character as opposed to selling Hoffman's voice. Hoffman's altering of his voice from speech to vocalisations, also highlights the environment and culture which Lancaster Dodd fosters in his religious movement. Gorbman further highlights the character's takeover of the soundscape through a diegetic vocal performance, providing character development within the soundscape through a single change in the voice. Hoffman's character performs a song, in a non-musical film, which highlights the personal narrative of Lancaster Dodd through what Gorbman exclaims as an 'extraordinary moment, because it's so mysterious.'<sup>58</sup> Hoffman's manipulation of his own vocal qualities worldbuilds the fictional filmic environment of Lancaster Dodd, suspending the disbelief of the audience that they are only witnessing an actor's performance, and that Hoffman is Lancaster Dodd.

### ***Gender and identification***

Voices which emphasise the characteristics of a character do so through these three styles of voice highlighted by Gorbman; an identification with a character can be formed through accent and cultural stereotyping, alongside familiar vocalisations, vocal grains, and perceived gendered voices. The gendering of voices can cause stereotyped character roles in multimedia, due to what Heidi Wilkins refers to as the 'noticeable return to binary representations of gender'<sup>59</sup> during the Hollywood Blockbuster era in the 1970s. Whereas gendered roles in multimedia can be problematic, the voice can be argued to have a greater purpose than to solely identify the body's gender. As the voice is disassociated with the body, becoming soundwaves travelling towards a listener, it enters

---

<sup>56</sup> Paul Thomas Anderson, *The Master*, Annapurna Pictures (The Weinstein Company, 2012).

<sup>57</sup> Claudia Gorbman, 'The Master's Voice', p. 10.

<sup>58</sup> Claudia Gorbman, 'The Master's Voice', p. 16.

<sup>59</sup> Heidi Wilkins, *Talkies, Road Movies and Chick Flicks: Gender, Genre and Film Sound in American Cinema* (Edinburgh: Edinburgh University Press, 2016), p. 125.

what Jarman-Ivens identifies as the ‘third space’. Jarman-Ivens argues that gender is not associated with the voice within this ‘third space’, as frequencies and sound waves are not gendered, meaning that the association of gender and identity is actually a perception of the listener.<sup>60</sup> Jarman-Ivens outlines this gendering of voices as such, and identifies the ‘third space’ as a place where voices can be genderless and thereby ‘queer’:

Its bodily origin and destination, and its operations across borders and through borders, and its traversal of the space between bodies, collectively give the voice a physical location in two bodies and in no body at all, and its meaning arguably arises in all three locations too, again bringing its categorizability into question. In gendered terms, then, the voice is a slippery beast and already potentially queer in this way.<sup>61</sup>

Voices as operational throughout bodies and borders, and within the genderless ‘third space’, can allow for vocal manipulations that identify a character body as opposed to a vocal actor. Annette Schlichter further explains the lack of difference between male and female vocals in the English language, as she refers to the scholar Pamela Hendricks who ‘argues that vocalization in itself does not provide enough information to the listener to communicate gender.’<sup>62</sup> This means that within animation, video games, and bodiless voices in multimedia, the bodily source of the voice does not need to match the gender it is attempting to create. It is common in both western and eastern animation and video games to use female voices to represent male characters.<sup>63</sup>

The ‘third space’ becomes more complicated when these extra characters are included between the original body of the voice and the listener. This extra connection in the journey between the body of the voice and the audience can further disrupt audience perceptions of the voice’s gender, and other characteristics (see Figure 2).

---

<sup>60</sup> Jarman-Ivens, ‘Queer Voices’, p. 3.

<sup>61</sup> Jarman-Ivens, ‘Queer Voices’, p. 19.

<sup>62</sup> Annette Schlichter, ‘Do Voices Matter? Vocality, Materiality, Gender Performativity’, *Body & Society*, 17.1 (2011), 31–52 (p. 44) <<http://journals.sagepub.com/doi/10.1177/1357034X10394669>>.

<sup>63</sup> For example, Masako Nozawa as Son Goku in the Japanese animation *Dragon Ball Z*, Mayumi Tanaka as Monkey D. Luffy in the Japanese animation *One Piece*, Veronica Taylor as Ash Ketchum in the English dub of the Japanese animation of *Pokémon*, Nancy Cartwright as Bart Simpson in *The Simpsons*, and Cassandra Lee Morris as Morgana in the English dub of the JRPG *Persona 5*.

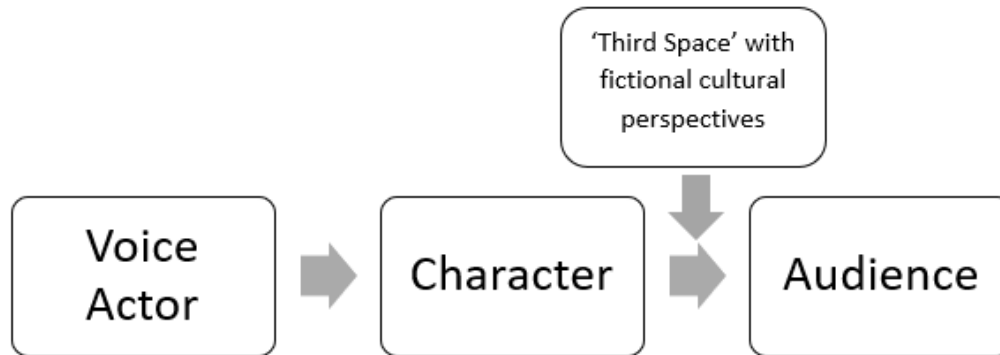


Figure 2: The 'third space' between character and audience, as opposed to speaker and listener.

These culturally prescribed gender roles can also become mistaken due the intermediary of a character, as Linda Fisher examines: 'vocal gendering is the product of a complex interplay among anatomical differences, socialization into culturally prescribed gender roles.'<sup>64</sup> As the voice can be manipulated to use codes that portray a character, so too can it be used to portray the social and cultural backgrounds of that character, rather than the voice actor themselves. This portrayal is discussed in the *Final Fantasy XV* case study in Chapter five. The four male playable-character voices, which are a constant source of vocalised feedback and information for the player in *Final Fantasy XV*, engages with the 'third space' as each voice enhances the character's characteristics, rather than solely identifying their own, and the voice actor's, gender. Therefore, the 'third space' in this situation allows the audience to step back from the voice actor, and vice versa, focusing on the character presented to them. This focus on the character can aid the worldbuilding process of both characterisations and the environmental associations of that character.

### **Silenced voices**

Fisher outlines voice as having a relationship with the feminine, 'from classical and mythological references such as Delphic Oracle, the Homeric Sirens, mermaids, and Echo, to contemporary identifications of the paradigmatic singing voice with the female diva.'<sup>65</sup> Where the voice is intentionally gendered, male and female voices receive varying treatments throughout multimedia. One such handling of voices is its silencing. Silenced voices are portrayed in a multitude of ways

<sup>64</sup> Linda Fisher, 'Feminist Phenomenological Voices', *Continental Philosophy Review*, 43.1 (2010), 83–95 (p. 88).

<sup>65</sup> Fisher, 'Feminist Phenomenological Voices', p. 86.

through mute characters, characters with the ability to either speak or sing but not both, and the disembodiment of voices and thus gestural communication and meaning.

The silencing of a voice is often directed towards female characters in cinema. Silencing women is not unique to multimedia, however, and exists within theatre and literature similarly. Dunn describes the labelling of Shakespeare's Ophelia in *Hamlet* as 'mad' due to her favour to communicate through song and reluctance to speak:

As female is opposed to male and madness to reason, so song in *Hamlet* is opposed to speech - particularly those modes of speech that serve to defend the patriarchal order from the threat represented by Ophelia's 'importunate' self-expression.<sup>66</sup>

This opposition of gender and voice, and thus sanity, can be seen within the treatment of the feminine voices throughout Hollywood cinema. Heather Laing categorises female muteness in classical Hollywood into three levels in her book *The Gendered Score: Music in 1940s Melodrama and the Woman's Film*: the first level of female vocal silence includes 'muteness' and a position where a woman cannot physically speak; the second level considers 'mutism' where the woman psychologically cannot speak but may be able to in the future or past; and the third level considers 'selective mutism' in which a woman can speak but chooses to remain silent.<sup>67</sup> These three levels of silenced female voices offer feminine silence from the perspective of trauma and madness, rather than a woman's agency. Catherine Haworth discusses this convoluted representation of female agency through her reading of the 1940s female gothic film *The Spiral Staircase*.<sup>68</sup> Haworth identifies that although female agency is the goal of these gothic films, the women depicted in these films often lose their agency filled detective roles to an overactive imagination:

The unusual level of agency and textual control found in many female gothic heroines represents a clear threat to the patriarchal depiction of identity more commonly found in classical film, and care is often taken to cast doubt upon the reliability of the heroine's vision and understanding.<sup>69</sup>

Through the ideas of Chion's 'acousmètre', Lisa Coulthard considers the silent female character as 'a silent witness or moral conscience, not necessarily at the centre of the story but operating as a key

---

<sup>66</sup> Leslie C. Dunn, 'Ophelia's songs in Hamlet: music, madness, and the feminine', in Dunn and Nancy Jones, *Embodied Voices*, p. 52.

<sup>67</sup> Heather Laing, *The Gendered Score: Music in 1940s Melodrama and the Woman's Film* (Aldershot: Ashgate, 2007), p. 28.

<sup>68</sup> Robert Siodmak, *The Spiral Staircase*, *Vanguard Films* (RKO Radio Pictures, 1946).

<sup>69</sup> Catherine Haworth, 'Something beneath the Flesh': Music, Gender, and Medical Discourse in the 1940s Female Gothic Film', *Journal of the Society for American Music*, 8.3 (2014), 338–70 (p. 342) <<https://doi.org/10.1017/S1752196314000236>>.

focal point running alongside the main action of the plot. In this regard, the voiceless body and bodiless voice have corresponding powers: angelic or demonic spectrality, ubiquity, terrifying omnipotence and secret knowledge.<sup>70</sup> These varying silencing and mutisms, outlined by Laing and the corresponding powers afforded to these mute women by Coulthard, places the mute female character in almost an ethereal, non-real, space that oversees the narrative but cannot directly effect it. Britta Sjogren's gendered analyses of the female narrator thus places the feminine character in this space that cannot physically touch the narrative, examining the feminine narrator as "seeing' through an 'internal (blind) vision."<sup>71</sup>

This distinction between male and female narrators combines the two genders, as male characters often affect the muteness of the feminine. Haworth reflects on the sonic handling of the *femme fatale*, and the silencing of her personal characteristics, due to the male subjective experience of her 'hormonal' voice:

Women in Hollywood (and in dominant cinema more generally) are defined by their sexuality and their 'lack' of maleness – which indicates whom that sexuality (whether submissive and wifely or untamed and deviant) is presumed to be directed. Music and sound form a crucial part of these representational strategies: as Luke recalls later in *A Woman's Secret*, reflecting upon Susan's performance: 'She just happened to have one of those voices [...] the sort that go down your spine and directly into your heart [...] she just had a voice with hormones.'<sup>72</sup>

This silencing objectifies the female voice as sexual, as Haworth outlines, whether this is of a deviant *femme fatale* or a wife. This interweaving of handled voices between male and female characters, agency and subjectivity, combines the fate of voices between these two genders. The subjective experience of the female voice by the male can render female characters mute in agency as well as physical voice. In Chapter five, I examine and expand on this treatment of female agency and the mute female voice. In the case study, *Transistor*, I examine the loss of agency surrounding the mute female player-character, her performative silencing, and her relationship with a male voice which narrates her past and present actions. Performative silencing however is not only a feminine issue as the two genders intertwine. Whereas female characters are physically silenced through madness or mutism, male vocal silencing often refers to their masculinity and performativity. When discussing

---

<sup>70</sup> Lisa Coulthard, 'The Violence of Silence: Vocal Provocation in the Cinema of Michael Haneke', *Studies in European Cinema*, 9.2–3 (2012), 87–97 (p. 90) <[https://doi.org/10.1386/seci.9.2-3.87\\_1](https://doi.org/10.1386/seci.9.2-3.87_1)>.

<sup>71</sup> Britta Sjogren, *Into the Vortex: Female Voice and Paradox in Film* (Urbana and Chicago: University of Illinois Press, 2006), pp. 21–22.

<sup>72</sup> Catherine Haworth, 'Introduction: Gender, Sexuality, and the Soundtrack', *Music, Sound, and the Moving Image*, 6.2 (2012), 113–35 (pp. 115–16).

the male musician, Laing admits that ‘the musical-emotional characterization of women only works as it does in relation to the musical-emotional characterization of men, and vice versa.’<sup>73</sup> This link between men and women can silence male artistic agency if they are missing a female muse. Alongside this, the effeminate inclusion of musical representations of characters silences masculinity in 1940s film soundscapes as ‘male characters are usually only explicitly represented through music in very particular narrative situations.’<sup>74</sup>

Where male and female voices have vocal agency, films often silence character voices for their privacy, not allowing audiences to hear what has been said. Justin Horton explains that an audiences’ corporeality and sense of realism is defined by voices alongside the visual imagery:

The voice, as we’ve elaborated, is the most privileged component of the soundtrack, and for good reason. In the same way that the mimetic representation of the human figure on-screen maintains an irrational sway over us, the voice, too, carries the uncanny power to assure us that, somewhere, these apparitions on the screen have a corporeal existence.<sup>75</sup>

As audiences are generally omnipotent within the diegesis, creating vocal silence where audiences can see but not hear character voices can form intrigue and frustration for audience members. Horton considers the vocal silence within the 2003 romantic comedy film *Lost in Translation*<sup>76</sup> within his text ‘The Unheard Voice in the Sound Film’. In the conclusive moments of the film, and during the conclusion of a relationship between the characters Bob Harris (Bill Murray) and Charlotte (Scarlett Johansson), vocal silencing is used to allow the two characters a private moment. The voices of Bob and Charlotte are not exactly silenced, but the volume of their voices and dialogue are kept to a whisper; the audience can hear the formation of the dialogue, but not the actual words themselves. This production of the voice allows the audience to become a part of the diegesis with a first-person view, as opposed to their usual omniscience. The audience becomes a part of the busy street as the soundscape loses any non-diegetic sound or music and maintains only natural diegetic sounds. Horton considers this silencing as ‘an unusual and unsettling disruption’.<sup>77</sup> For audiences used to being omniscient in a filmic space, this vocal silence thus could be considered as a deafness of the audience in order to provide the characters a moment of intimacy and privacy. This unusual view ‘normalises’ a moment rather than its usual filmic romanticising, reminding viewers that this interaction is nothing exceptional.

---

<sup>73</sup> Laing, ‘The Gendered Score’, p. 139.

<sup>74</sup> Laing, ‘The Gendered Score’, p. 140.

<sup>75</sup> Justin Horton, ‘The Unheard Voice in the Sound Film’, *Cinema Journal*, 52.4 (2013), 3–24 (p. 23).

<sup>76</sup> Sofia Coppola, *Lost in Translation*, *American Zoetrope, Elemental Films* (Focus Features, Tohokushinsha Film, 2003).

<sup>77</sup> Horton, ‘The Unheard Voice’, p. 4.

Whereas the inclusion of voices can transcend the audience to an omniscient observer, the lack of voices where they should exist can bring the audience into the diegesis. These intimate moments can be jarring for audiences which do not expect or want to be left out of audio feedback. The film *Star Wars: The Last Jedi*<sup>78</sup> contained ten seconds of full vocal, diegetic, and non-diegetic silence.<sup>79</sup> The silence was used to represent both the lack of sound in a vacuum, and to highlight the death of an important character. However, this silence did not translate to audience members who complained about a fault in the audio, thus forcing movie theatres to include warnings of silence.

### ***Narrating voices***

The vocal agency of Ariel, in Disney's 1989 film *The Little Mermaid*,<sup>80</sup> contains her ability to narrate herself through her siren song 'Part of Your World'. After Ariel exchanges her voice for human legs, Coyle and Jon Fitzgerald recognise that Ariel's song becomes an orchestrated rendition of her identifying theme for the remainder of the film. This musical rendition attempts to symbolise Ariel's will, despite the loss of her voice. Even after Ariel regains her voice, and marries Prince Eric in the final scene, Ariel does not sing her siren song 'Part of Your World' again. The silencing of Ariel's voice is outlined by Coyle and Fitzgerald as the loss of her siren song to ensnare the prince.<sup>81</sup> Ariel's ability to ensnare Prince Eric and narrate her personal story is silenced until she regains her voice; the antagonist of the film, the sea witch Ursula, instead ensnares Prince Eric herself by using Ariel's voice instead of her own. After breaking the siren's spell on Prince Eric and 'defeating' Ursula by reclaiming her voice, Ariel thus ensnares the prince by proving that she has a beautiful speaking voice. However, Ariel's siren song is reduced to a non-diegetic chorus rendition rather than sung by Ariel, suggesting that she has given up her life as a mermaid, and thus her own narrative power, in order to marry the prince. This use of non-diegetic vocals to narrate the stories' conclusion, and Ariel's decisions, could be considered as a voice-off, as Sjogren considers the narrative space of the voice-off as existing within its own space:

Sound thus emerges as a parallel process, powerfully structuring the meanings available to the spectator of a film. It is by way of stressing this 'parallel' status of sound, moreover, that

---

<sup>78</sup> Rian Johnson, *Star Wars: The Last Jedi*, Lucasfilm Ltd. (Walt Disney Studios Motion Pictures, 2017).

<sup>79</sup> Sven Raeymaekers, 'Silence, and...action! History and analysis of silence in the Hollywood sound film', paper delivered at the British Forum for Ethnomusicology and Royal Musical Association Research Students' Conference, University of Huddersfield, 4-6 January 2018. Raeymaekers continued to identify diegetic silence, non-diegetic silence, and full silence throughout thirty films from 1927-2016 and how this, through semiotic meaning, can enhance emotional experiences.

<sup>80</sup> Ron Clements and John Musker, *The Little Mermaid*, Buena Vista Pictures (Alan Menken, 1989).

<sup>81</sup> Coyle, *Drawn to Sound*, p. 230.

the term ‘voice-off’ is used throughout this text, rather than the more traditional ‘voice-over.’ I prefer the term ‘off,’ in part, because it registers an independent space. Whereas ‘over’ suggests a top ‘layer’ or cloak of some kind, ‘off’ connotes otherness— a distinctness that moves alongside, ‘elsewhere.’<sup>82</sup>

Voice-overs and voice-offs are common vocalised story-telling tools in literature and multimedia to identify on-screen actions, character developments, and overarching narratives. The voice-off, according to Sjogren, interrupts the filmic world, and audience expectations, by adding an ‘other’ vocal space into a ‘normalised’ filmic world.<sup>83</sup> This type of narrator, that does not appear in the same filmic world, could be what Chion defines as the ‘commentator-acousmètre, he who never shows himself but who has no personal stake in the image.’<sup>84</sup> Sjogren’s separate space for the voice-off can contain this commentator, seen within the choral rendition of Ariel’s siren song, separate from the filmic world and thus from having personal stakes. Jesse Kinne described this voice-off phenomenon with the use of the sung non-diegetic female voice, at the 2019 Music and the Moving Image Conference, as the ‘destiny topos’.<sup>85</sup> The ‘destiny topos’ uses a female voice performing syllabic phrases to narrate a known past, or future destiny of a character, sometimes echoing a character or story leitmotif. Kinne used the finale of the HBO series *Game of Thrones*<sup>86</sup> as a prime example of the ‘destiny topos’ as it narrated the character Daenerys Targaryen as she reached her planned goal and destiny, the iron throne. As Daenerys enters the throne room, the ‘destiny topos’ soprano voice performs a syllabic (‘ah’) rendition of the *Game of Thrones (GoT)* leitmotif, indicating Daenerys is fulfilling her dream. In this scene, the velocity of the sound design is reduced as the non-diegetic voice is highlighted as Daenerys walks towards the throne. These destiny voices are not always solo however, the voice is accompanied in this scene by drones on low string and brass instruments. These instruments transition into the forefront of the music as Daenerys touches the throne, and the voice is silenced having completed its narrative journey. The voice is replaced by a viola which is performed with tension on the strings as it plays the *GoT* leitmotif. This tension symbolises Daenerys’ growing madness throughout the series and her association, by other characters, with her father Aerys II Targaryen (the ‘mad king’); although Daenerys is not mute throughout *GoT*, her agency becomes a classic representation of vocalised ‘madness’ which must be stopped by her ‘sane’ male counterpart.

---

<sup>82</sup> Sjogren, ‘Into the Vortex’, p. 6.

<sup>83</sup> Sjogren, ‘Into the Vortex’, pp. 23–25.

<sup>84</sup> Chion, ‘The Voice in Cinema’, p. 21.

<sup>85</sup> Jesse Kinne, ‘The Solo Female Voice as Destiny Topos in Fantasy Media’, in *Music and the Moving Image XII*, 2019.

<sup>86</sup> David Benioff and D.B. Weiss, *Game of Thrones*, HBO (Warner Bros. Television Distribution, 2011).



Where the voice-off considers a narrative voice that is separated from the diegesis, the voice-over exists within an off-screen extension of a world, whether this is from an over-the-shoulder perspective or an internal 'blind' image of an off-screen character. Voice-overs can also consider Chion's concept of the *acousmêtre*. Narrative *acousmêtres* can exist as a 'complete *acousmêtre*', an off-screen voice whose body has not been seen but has the possibility to be seen, and the 'visualized *acousmêtre*', the off-screen voice of a bodied character that the audience has already seen.<sup>87</sup> Alongside Chion's separation of *acousmêtre* types, Sarah Kozloff categorises the voice-over into first-person narration and third-person narration. When considering defining these types of narrators, Kozloff considers the literary voice-over terminology offered by Gérard Genette. Genette distinguishes voice-overs by characters within the story as 'homodiegetic', whereas voice-overs from off-screen characters are considered 'heterodiegetic'.<sup>88</sup> Kozloff's categorisation of the voice-over highlights the differences in which these voices identify on-screen action alongside audience reactions to different narrators. The homodiegetic, first-person, voice-overs act irregularly as narrators but, unlike the voice-off or third-person voice-over, can be witnessed narrating by the audience:

[...]through showing the character in the act of narrating, through the narrator explicitly telling us that this is his or her story (that is, by the voice providing us a Labovian abstract and/or orientation), or through a combination of conventional indicators such as the camera dolly or zooming in on the character's face and the image dissolving to another scene, often to the accompaniment of dreamy music.<sup>89</sup>

Whereas the first-person voice-over narrator can be seen, the 'heterodiegetic' third-person voice-over is defined by Kozloff as omniscient and off-screen:

[...] these narrators, who speak for (or rather as) the image-maker, are particularly likely to provide guidance concerning what conclusions the viewers should draw. They tend to voice the ideological and/or moral agenda behind the film. Taking advantage of their remove from the story, they are generally heavier on Labovian evaluation than first-person narrators, more likely to use their platform to probe, judge, generalize, and interpret.<sup>90</sup>

This omniscience could be seen as a part of the voice-off, but their connection with the narrative world through their all-knowing of facts and figures connects them to the diegesis even though they

---

<sup>87</sup> Chion, 'The Voice in Cinema', p. 21.

<sup>88</sup> Sarah Kozloff, *Invisible Storytellers: Voice-over Narration in American Fiction Film* (London: University of California Press, Ltd., 1988), pp. 42–43.

<sup>89</sup> Kozloff, 'Invisible Storytellers', p. 45.

<sup>90</sup> Kozloff, 'Invisible Storytellers', p. 80.

are set apart. Although Gennette's model is a useful tool to understanding various voice-over styles in literature, Kozloff questions its unnecessary complexity, and considers whether the design is suitable for film. Gennette's model is convoluted but the distinctions made between the 'homodiegetic' and 'heterodiegetic' narrator characters aids understanding of different narrator voices when identifying environments, characteristics, and personal narratives.

## Voices in Video Games

The intertextuality of voices, as discussed throughout this chapter, and their handling in multimedia bears similarities to their use in video games. The inclusion of interactivity in this medium, however, can change the implementation of voices within the diegesis; where a voice-off in a film teases the audience that they may find the off-screen narrator if they turn their head, in a video game the player can actively turn their character to look for this off-screen voice, bringing it on-screen.

The inclusion of voices in RPGs are still relatively new due to the age of the medium. The inclusion of recorded voices in video games has had a similarly jarring transition period to the transition of film from its silent era to the talkies. However, voices were included in video games before the implementation of recorded voices, as programmers could create simulations of voices with sound-chip and synthesisers. The inclusion of a sound chip in arcade machines meant that synthesised style noises could be created as carriers of information. These bleeps and bloops of a sound chip were interpreted as a voice by players, requiring creativity and imagination in order to hear sounds become words, a deciphering tool now inherent in most gamers. Unfortunately, introducing these character sounds to a game, where sounds were prioritised for music and feedback for player action, caused what Kenneth McAlpine describes as 'note stealing'.<sup>91</sup> When working with limited memory in arcade machines and consoles (specifically here with the Nintendo Entertainment System console), McAlpine explains that different channels of sound had to be prioritised. Thus, when too many channels were played at one time other channels of sound could not play. Composers needed to decide between prioritising reactive game sound or the musical cue, and thus needed to create music which would work whilst dropping in and out of the crowded aural space whenever reactive sounds were implemented.<sup>92</sup> This lack of memory meant that voices were not included in earlier games as the challenge of creating synthesized voices with appropriate vocal inflections also came

---

<sup>91</sup> Kenneth B. McAlpine, *Bits and Pieces: A History of Chiptunes* (New York: Oxford University Press, 2019), p. 28.

<sup>92</sup> McAlpine, 'Bits and Pieces', p. 122.

with an added cost.<sup>93</sup> This ‘note stealing’ however did not stop players from mentally imagining the sounds they heard attached to characters as ‘voices’.

### ***Do Synthesised Voices Sing?***

When performing opera, singers manipulate the opera text through extended vocal techniques and tessituras. The manipulation of words through the act of singing rather than speaking in opera, however, does not dampen the audience’s understanding of what is happening on stage. It is common for audiences to hear an opera outside of a language they know, but the complexity of human’s relationship with communication, and the generation of language, creates a form of musical-communicative listening. Rutherford outlines this reshaping of opera texts, by the vocal manipulation of the singer and their voice, and the discussions of how this effects an audience’s relationship with language and communication:

Music’s ludic transmutations of word are far from a new, alarming intrusion into our lived experience.[...]The sung word might thus be viewed (or heard) not as a degeneration of language, but rather as a sophisticated elaboration of the sounds or ‘dynamic emotional syntax’ that actually generate language.<sup>94</sup>

Laing also connotes this relationship of sung, or manipulated, language with an emotional syntax. Laing identifies that in Hollywood Musicals, such as *Summer Stock*,<sup>95</sup> vocalised musical numbers can formulate the narrative and emotional philosophy of the film, highlighted by the lyrics.<sup>96</sup> If the discussion surrounding opera, film, and musicals considers the singing voice with the ability to permeate possible language barriers, through emotional and sound syntax as opposed to identifying meaning through words, so too can the synthesised voice follow these rules. Rutherford identifies that audiences can understand musical communication through their evolutionary relationship and lived experiences of vocalisations and non-verbal communications of meaning. Where audiences have this lived experience, the player’s creative ability to translate synthesised sounds into a type of language can exist. George Reid has instigated research on the academic considerations of chiptune fandoms and fan identity in their article ‘Chiptune: The Ludomusical Shaping of Identity’. After defining chiptune as a component derived from the historic sounds of the formative years of video

---

<sup>93</sup> See Cheng, ‘Sound Play’, pp. 58–60.

<sup>94</sup> Rutherford, ‘Voices and Singers’, in Till, *The Cambridge Companion to Opera Studies*, p. 123.

<sup>95</sup> Charles Walters, *Summer Stock*, MGM (Loew’s, Inc., 1950).

<sup>96</sup> Heather Laing, ‘Emotion By Numbers: Music, Song and the Musical’, in *Musicals: Hollywood and Beyond*, ed. by Bill Marshall and Robynn Stilwell (Exeter: Intellect Ltd, 2000), p. 5.

games, Reid examines identity and self from heterogeneous subjectivity of chiptune through this paper and their further PhD thesis:

For the context of chiptune fan subjectivity, the amalgam of ‘entities’ referred to here are musical/non-musical, human/non-human ‘actors.’[...] For chiptune, we can understand that these actors may include non-human and non-musical 8-bit or 16-bit timbral qualities. Chiptune technology becomes a nonhuman and non-musical actor in the network; hearing, for example, that a chiptune is composed using a Sega Master System chip also contributes to listening subjectivity formation as an encounter. The listener would encounter the sound-mediating agency of the Master System chip as a non-human actor in the network, allowing for the listener to not only characterise or categorise the chiptune, but also construct subjectivity based upon how they personally relate to this encounter.<sup>97</sup>

Reid’s association of human or non-human, musical or non-musical, actors incorporates the technological creation of chiptune sounds as the synthesised voice. The ability of players to understand the meaning of the synthesised voice thus triggers the identification and sense of self between players and vocal, chiptune characterisations; Reid further examines this as including ‘memories, nostalgia, self-identification, videogaming pleasures and (sub)cultural ‘capital’—all of these elements are triggered and, moreover, become interconnected through chiptune interaction.’<sup>98</sup> Both Reid and Rutherford consider musical parameters as ingrained within these transliterations. Rutherford suggests that ‘communicative musicality’ considers ‘pitch, melody, rhythm, tempo and dynamics’ enabling mental and emotional attunement’.<sup>99</sup> Reid also suggests that ‘melody, rhythm, timbre, harmony, articulation, and technological mediation [...] all have an impact upon the listener’s subjectivity.’<sup>100</sup> These similarities in musical literacy, within both opera and video games, can confirm that where synthesised voices do sing, player comprehension is not far behind.

The opera *Maria and Draco* in *Final Fantasy VI* uses synthesised (chiptune) voices, alongside text boxes and character animations, to portray the vocal performance of an Aria. The character Celes Chere impersonates the opera star Maria to foil the plans of her kidnapper, and thus must perform an Aria on stage. To capture the kidnapper, Celes must embody Maria’s role and perfectly perform the Aria with the help of the player. Before the opera begins, the player has a chance to study the lyrics of the Aria as they must choose the correct lyrical stanza during the performance. If the player chooses an incorrect stanza in the performance, the opera restarts. The player’s interaction with the

---

<sup>97</sup> George Reid, ‘Chiptune: The Ludomusical Shaping of Identity’, *The Computer Games Journal*, 2018, 279–90 (pp. 284–85) <<https://doi.org/https://doi.org/10.1007/s40869-018-0070-y>>.

<sup>98</sup> Reid, ‘Chiptune’, p. 285.

<sup>99</sup> Rutherford, ‘Voices and Singers’, in Till, *The Cambridge Companion to Opera Studies*, p. 123.

<sup>100</sup> Reid, ‘Chiptune’, p. 284.

Aria demands the player's acknowledgement that Celes' voice is a performative one. The strength of this inclusion of the player is outlined by Cheng as he outlines the complexity of the reception of Celes' voice:

Players' fondness for the synthesized aria is sustained by memories of its charming novelty. Nostalgia is the miracle fuel of deep-seated loyalties: it trumps aesthetic conventions with the sheer force of yearning. Like the pixelated graphics, chiptune soundtracks, and hammy voice acting of early games, Celes's voice has since come to be not just vindicated but outright valorized as an artifact of historical design.<sup>101</sup>

The historicism of chiptune sounds of voice permeates fields such as live video game music concerts and re-orchestrated soundtracks. Cheng outlines that because of this nostalgia for the original synthesised version, as discussed above, players find the original synthesised voice of Celes more identifiable and emotional than a live recreation.

Celes' synthesised siren song for the player is seen within modern-retro style games, such as *Undertale*.<sup>102</sup> *Undertale* is a retro pixel aesthetic RPG with the gimmick that the player never needs to 'kill' an enemy, and can always resolve conflict with talking; there is a consequence to every action in *Undertale* which have repercussions such as a reduced number of NPCs to talk too if the player has killed enemies in combat. During combat with Shyren, a shy siren, the player can talk to the enemy and help provide their agency to sing; the mermaid-like monster attempts to hide her singing abilities from the player otherwise. If the player chooses to make the player-character 'hum' and 'conduct' as an action against Shyren's melodic 'attacks', Shyren will begin to sing extracts of the player-character's melody. These extracts escalate until Shyren is comfortable with her singing and 'suddenly, it's a concert'. Shyren's vocal performance consists of synthesised quaver-length notes which shift in pitch. Each performance is the same melody but accelerates as the player-character continues to 'hum' or 'conduct'. Shyren's siren identity is not just a characteristic design as her visualised notation when singing can damage the player-character. As video game technology advanced, voices moved from these synthesised sounds to recorded voices. The original synthesised voices of Pokémon, in the 1998 RPG *Pokémon Yellow*,<sup>103</sup> signified each 151 variations of Pokémon as players understood these synthesised sounds as vocal cries; Pikachu itself had a recorded vocalisation of its cry, however, the remake of *Pokémon Yellow*, in *Pokémon: Let's Go, Pikachu!* and

---

<sup>101</sup> Cheng, 'Sound Play', p. 78.

<sup>102</sup> Toby Fox, *Undertale, Independent: Toby Fox* (Microsoft Windows, Linux, OS X, PlayStation 4, PlayStation Vita, Nintendo Switch, 2015).

<sup>103</sup> Game Freak, *Pokémon Yellow, Kyoto: Nintendo* (Game Boy, Game Boy Color, 2000).

*Let's Go, Eevee!*,<sup>104</sup> sees the inclusion of recorded vocalisations, to identify the Pokémon Eevee alongside Pikachu at a much higher quality, due to advances in technology.

### **Recorded Voices**

The case studies I analyse in chapters four, five, and six all use recorded voices in the diegetic and non-diegetic space in order to establish coherent environmental worldbuilding, and realistic characteristics. Although players understood the aesthetical and character-based meanings of synthesised voices, recorded voices can be used to worldbuild greater overarching environments and cultures of the game world. The unique language created for the original and remake of *Shadow of the Colossus*<sup>105</sup> is portrayed through recorded voices that identify both the human characters and unknown god in a deserted environment. The narrative of *Shadow of the Colossus* follows the player-character named Wander, and his horse Agro. As Wander speaks with the unknown god, Dormin, the language of Wander is repeated back within this alien environment. This use of language makes Dormin appear trustworthy as he also appears alongside the imagery of light that emits from the sky. The coherency of Dormin intentionally blurs the player-character and player's recognition of the god's true identity and agenda. It is not until the end of the game, after the player has defeated all sixteen colossi, that it is explained that Wander has acted in selfishness and travelled to a forsaken land in order to revive an unknown dead female character. As Wander completes his goal set by Dormin, to defeat the colossi, Dormin's evil agenda to inhabit Wander's body is revealed. Wander's slow sacrifice of his body to Dormin is not without a prize, however. When the player has defeated a requisite number of colossi, the game provides the player with the glimpse of a female voice in the same language. This voice occurs during a loading, portal style, screen which symbolises the transportation of an unconscious Wander back to Dormin's stone temple. This respite for the player (as Wander is incapacitated) provides a glimpse towards their goal, the restoration of the woman, through the sounds of female vocalised whispers. Recorded voices cement the themes of life, death, familiarity, otherness, and betrayal throughout the game. Voices worldbuild the isolated environment, as outlined by Gibbons, through silences and the inclusion of a language that the player's do not speak. Alongside this, voices shield the real characterisations of the characters until the end of the game through the familiarity of using the same language across various beings.

---

<sup>104</sup> Game Freak, *Pokémon: Let's Go, Pikachu! And Let's Go, Eevee!*, Tokyo: The Pokémon Company (Nintendo Switch, 2018).

<sup>105</sup> Bluepoint Games, *Shadow of the Colossus, San Mateo: Sony Interactive Entertainment* (PlayStation 4, 2018).

Voice actors and performers must consider the physicality of the on-screen body they are representing in games. Collins explains that ‘our understanding of the world is shaped by our ability to physically interact with it.’<sup>106</sup> In this case, Collins’ understanding of mirror neurons is a gateway to the player’s recognition of on-screen environments, characterisations, emotions, and consequential actions of their character. Collins’ research discusses video game sound and perception with reference to psychological topics. Collins’ paper, ‘Making gamers cry: mirror neurons and embodied interaction with game sound’, focuses on topics such as empathy and emotional relationships between the player and characters. Collins takes an approach to this topic which considers embodied cognition, mirror neurons, anthropomorphism, and the combination of gestural movements and reactive sounds. Mirror neurons are the reactivation of sensory neurons in the brain that activate when we witness an action we have experienced and performed ourselves. Consequently, empathy and other simple and complicated emotions can be tied to these neurons and applied to on-screen characters.<sup>107</sup>

The triggering of mirror neurons helps an audience to believe that a sound is emanating from an on-screen character due to the mental understanding of the world, leading to the success of playback singers and ADR technology. Much like the popularity of playback singers in Bollywood, the vocal talents of voice actors such as Matthew Mercer, Yuri Lowenthal, Ashley Johnson, and Ashly Burch are celebrated similarly by fans, leading to the popularity of voice actor related web shows, such as *Critical Role*.<sup>108</sup> The acknowledgement of professional voice actors for video games, who are often also the voice actors for western cartoons and the dubbing actors for Japanese animation, is due to their convincing voices on a character model. These convincing voices will allow the player’s mirror neurons to understand the emotional signifiers and other meanings of the voice because it synchronises with the on-screen narrative and the expected emotions of the character. High quality voices can surpass the player’s expectations and allow players to grow attached to voices over time, as seen with Celes.

When playing video games, Collins suggests that ‘we mentally re-create (visually and motorically) what we hear, and we hear in terms of intentionality and causality—including emotional intent—and thus we empathize with the originator behind the sound.’<sup>109</sup> Unlike the use of playback singers in Bollywood, a character’s voice actor anthropomorphises the animated model, giving it agency by providing a voice. Player’s prefer continuity in the vocal nuances of their characters as they identify

---

<sup>106</sup> Karen Collins, ‘Making Gamers Cry: Mirror Neurons and Embodied Interaction with Game Sound’, *Proceedings of AudioMostly 2011*, March, 2011, 39–46 (p. 39) <<https://doi.org/10.1145/2095667.2095673>>.

<sup>107</sup> Collins, ‘Making Gamers Cry’, pp. 39–40.

<sup>108</sup> Matthew Mercer, *Critical Role, Twitch, YouTube* (Geek & Sundry, Critical Role Productions LLC, 2015-).

<sup>109</sup> Collins, ‘Making Gamers Cry’, p. 41.

themselves, or other non-playable-characters, with that particular voice actor's character voice. Therefore where possible, voice actors are not changed between dialogue and vocal performances in video games; this fidelity to a voice actor is seen with Ashley Johnson as Ellie 'Through the Valley' *The Last of Us Part II* trailer;<sup>110</sup> Ashley Barrett as Zia 'Build That Wall' *Bastion*<sup>111</sup> and Red (all vocal performances) *Transistor*, Tamaryn Payne as Lohse 'Sing for Me' *Divinity: Original Sin II*.

The consequence of congruous sound to character-model, however, is when it desynchronises from what the player is expecting. Technological problems such as animation plugins not connecting to the character body for bodily movements can disrupt empathetic emotions towards characters. As I watched my friends play *Life is Strange*,<sup>112</sup> I witnessed what was meant to be an emotional conversation go array, in a critical moment of the narrative. The movement plugin for synchronised speech did not connect with the player-character Max, and subsequently her mouth did not move whilst dialogue was playing; although my friends were somewhat engaged with the game through countless hours of gameplay and thus found this frustrating, as an audience member it was amusing. In *Shadow of the Colossus* the only companion the player and player-character have is the horse Agro. When travelling to the final colossi, the player loses Agro as a bridge crumbles and Agro falls to what is assumed his death. Non-diegetic vocalisations perform the melodic identifier of the game in a minor key which highlight the death of the player's companion. This is meant to be an emotionally heightened scene where the player feels remorse for the loss of their horse. However, as I casually watched the streamer Day[9] play through *Shadow of the Colossus*, the animation plugin did not connect between the horse's movements and the bridge. Agro's model runs underneath the bridge, clipping through the ground, before the cutscene takes over and Agro surprisingly still manages to throw the player-character to safety. Day[9] erupts into laughter (about the plugin, not the death of the horse) and what should have been an empathetic and heart-breaking moment becomes hysterical.<sup>113</sup>

---

<sup>110</sup> Naughty Dog, *The Last of Us Part II, San Mateo: Sony Interactive Entertainment* (PlayStation 4, 2020).

<sup>111</sup> Supergiant Games, *Bastion, Burbank: Warner Bros. Interactive Entertainment* (Xbox 360, Microsoft Windows, Linux, OS X, iOS, PlayStation 4, PlayStation Vita, Xbox One, Nintendo Switch, 2011).

<sup>112</sup> Dontnod Entertainment, *Life Is Strange, Tokyo: Square Enix* (Microsoft Windows, PlayStation 3, PlayStation 4, Xbox 360, Xbox One, iOS, Android, 2015).

<sup>113</sup> Day9TV, 'Day[9]'s Day Off - Shadow of the Colossus Remastered P5', *YouTube*, 25 February 2018 <[https://youtu.be/6H\\_mFS6\\_6il?t=790](https://youtu.be/6H_mFS6_6il?t=790)>.



## **Live-Action**

Voice acting, whether used for vocal communications or the vocalisations of a character, is a necessary aspect of most video games because of the animated nature of cinematic and gameplay scenes. Full motion video games (FMV) are the creation of games by recording live-action sequences, allowing the player minimum control and interaction with these 'real' scenes. Some games use FMV within their cinematics, such as *The Quiet Man*,<sup>114</sup> but use animated character models during gameplay so that the player can interact and move within the game world. Fully FMV games, such as *Erica*,<sup>115</sup> use live-action actors to enact scenes which allow the player to inform character decision making through dialogue choices and interaction with objects in the world. There are slight differences between FMVs and interactive movies, however. The interactive genre of television and film combines the interactive technology of video games with linear storytelling. These crossovers, unlike FMV, usually only allow the player to engage in narrative decisions whilst they cannot physically interact with the world. Netflix's *Black Mirror: Bandersnatch*<sup>116</sup> (*Bandersnatch*) shares certain structural elements of FMVs and video games. *Bandersnatch* exhibits branching paths of narrative, each filmed and implemented to the structure beforehand, which switch routes depending on the player's decision. Branching paths and deterministic decision making is not a feature of all video games but is a tool that actively changes films to interactive movies by using the non-linear process of a game. These Choose Your Own Adventure-style films allow the player to change the narrative path and thus the outcomes of supposed character decisions and developments, with several endings and failure states. The excitement of this interactivity with on-screen characters, and the inclusion of multiple endings, saw players play *Bandersnatch* various times to see each conclusion of the narrative. Whereas FMV and interactive movies use live-action actors, video games use an increasingly sophisticated production of animation in cinematics and gameplay which incorporate voice acting. Due to the lack of gameplay, physical action, and movement in the worlds of FMVs and interactive movies by the player, these mediums will not be considered within this thesis, however their current increase in popularity since *Bandersnatch* and *Erica* is notable.

---

<sup>114</sup> Human Head Studios, *The Quiet Man*, *Tokyo: Square Enix* (Microsoft Windows, PlayStation 4, 2018).

<sup>115</sup> Flavourworks, *Erica*, *San Mateo: Sony Interactive Entertainment* (PlayStation 4, 2019).

<sup>116</sup> David Slade, *Black Mirror: Bandersnatch* (Netflix, 2018).

## ***Player Voices***

Although stated that player voices are not the focus of this thesis, they are still an integral part of voices in video games, and thus need discussing. Player voices are included in a game world either through the player's insertion of their own voice when they need to communicate with other players during multiplayer games, or as voice activation of characters within games such as *There Came an Echo*<sup>117</sup> which uses player voice to command the action of playable-characters.

As voices became a norm in video games, and the internet blossomed, online multiplayer games such as *Warcraft*<sup>118</sup> allowed players to connect through servers and voice chats. Voice chat servers were opened for gamers to communicate both internally within the game or on external applications, and thus player voices became a part of some game worlds. This is most common with multiplayer games where teamwork needs to be established between players. However, some single player games like *Nioh*<sup>119</sup> and *Dark Souls*<sup>120</sup> can let players link online and leave text-based messages as guides, or as misguidance, for other players. Other single player games such as *Journey* connect players together for a joint exploration experience, but limits communication to the medium of the playable-character and their ability to press a button for a vocalisation. Although the ability to communicate with fellow players during gameplay can be a positive experience, Cheng's involved study of the online multiplayer game *Team Fortress 2*<sup>121</sup> highlights the issues of online communications.

Vocal timbre comes into question when players communicate online as Cheng considers the gaming communities' reception of player voices. As gaming culture grows, so too do issues of racism and misogyny, outlined clearly by Gamergate in 2014. Gamergate refers to the attack on game developer Zoë Quinn for receiving favourable reviews for her indie title *Depression Quest*.<sup>122</sup> The favourable reviews caused a harassment campaign towards Zoë Quinn as some male gamers did not believe the game was successful, and thus believed Quinn provided sexual favours in return for positive criticism. Gamergate brought to light the misogynistic and racist attitudes of some gaming

---

<sup>117</sup> Iridium Studios, *There Came an Echo*, Los Angeles: Iridium Studios (Microsoft Windows, PlayStation 4, 2015).

<sup>118</sup> Blizzard Entertainment, *Warcraft III: Reign of Chaos*, Irvine: Blizzard Entertainment (Microsoft Windows, Classic Mac OS, Mac OS X, 2002).

<sup>119</sup> Team Ninja, *Nioh*, San Mateo: Sony Interactive Entertainment (PlayStation 4, Microsoft Windows, 2017).

<sup>120</sup> FromSoftware, *Dark Souls*, Tokyo: Namco Bandai Games (Playstation 3, Xbox 360, Microsoft Windows, PlayStation 4, Xbox One, Nintendo Switch, 2011).

<sup>121</sup> Valve, *Team Fortress 2*, Bellevue: Valve Corporation (Microsoft Windows, Xbox 360, PlayStation 3, Mac OS X, Linux, 2007).

<sup>122</sup> The Quinnsspiracy, *Depression Quest*, Independent: The Quinnsspiracy (Browser, Microsoft Windows, OS X, Linux, 2013).

communities, something which Cheng outlines within his experience of *Team Fortress 2*. The publication of *Soundplay* in 2014 crossed paths with Gamergate, which began in the August of 2014, as Cheng speaks of vocal harassments from his analytical play of the game. Vocal harassments online include male players deciding whether another player's voice sounds like a girl or a twelve-year-old boy:

The extent to which disembodied voices evoke accurate human bodies necessarily depends on players' varying capabilities to deploy and decode these voices as timbral, registral, and phonetic indices of appearance, age, sex, sexuality, ethnicity, nationality, class, (dis)ability, and other categories of physiological and cultural identity.<sup>123</sup>

These issues of the voice and the treatment of female players and characters has a backlash on the numbers of women in the gaming industry. This has a direct impact on the number of players who are using their voices in this capacity, especially those who are not the stereotypical white male player.

Voices in video games are slightly different to their inclusion in television and film due to this interactivity with other players and characters. The identification process with characteristics of characters causes issues for voices, as they often fall privy to the preferences of the player and what they want out of a character. However, film musicology literature shows a clear influence on the use of voices in video games. This chapter is not a holistic overview of either voice literature or voices in multimedia, nor is it a complete consideration of all voices in video games. What this chapter does do is highlight the important uses of the voice in film and video games to define, organise, and show the relevant use of voices when worldbuilding a fictional world. The following chapter will identify the use of these voices in identifying environments and worldbuilding people and culture in the RPG.

---

<sup>123</sup> Cheng, 'Sound Play', pp. 143–44.

## Chapter Four: Environmental Worldbuilding

The model outlined in Chapter two introduces environments, characterisations, and player to character identification as components of diegesis and outsourced worldbuilding, which can be created through a soundscape. This chapter will focus on environmental worldbuilding in video games through the inclusion of voices as enhancers of a game world's space, place, and time, alongside the cultural signifiers of the game world's people. The aim of this chapter is to further this understanding of the connection between worldbuilding and player engagement, or immersion, with a virtual game world environment and its people. There are some crossovers between worldbuilding environments and worldbuilding characterisations. Specifically, this can refer to the inclusion of people and culture within this chapter who also have characterisations that can resonate with the player. However, the nuance of people and culture is their unique behaviours compared to the player and player-character. The player, and their character, are often disconnected from the NPCs and cultures of the game world due to their engagement with the completion of their own agendas. The agency of the player, in removing themselves from NPC life, categorises NPCs as tools in creating a seemingly full and complex environment.

The divide between the player-character's goals and the goals of NPCs is important to this distinction between people and culture as environmental worldbuilding tools, rather than to encourage player identification. This distinction occurs through the lack of literal timeframes for the player to complete the main, narrative quest which affects the greater game world and its people. The nature of most games is to allow the player control of their own exploration whilst providing extra, side-quests whilst they navigate the storyline. The player does not need to complete side-quests as they are usually in-place to provide extra content for the game, and to engage the player with other characters and cultures. The inclusion of these side-quests, however, means that the player does not have any time limitation on their completion of the main-quest. This can lead to amusing situations where players can ignore, seemingly, time-sensitive story moments in a game, if they want to explore the world and complete side-quests. The action-adventure game *Yakuza 0*<sup>1</sup> is acclaimed for its narrative complexities, gripping story line, and character developments. *Yakuza 0*'s story is rarely actionless, but the game provides the player with various side-quests and minigames throughout. When an NPC of importance to the player-character, Majima, is kidnapped in the game's main storyline quest, there is a narrative impetus for the player to move swiftly in order to recover the character. However, the player can choose to complete side-quests and minigames instead of

---

<sup>1</sup> Ryu Ga Gotoku Studio, *Yakuza 0, Tokyo: Sega* (PlayStation 4, Microsoft Windows, Xbox One, 2017).

reaching Majima's narrative goal; the player can spend hours of gameplay, and days of game time, playing minigames: running a cabaret club; bowling; performing karaoke; playing Mahjong; practicing baseball at the batting cages; performing a rhythm dancing game at the disco; playing darts; going to the Sega arcade; fishing; and so on. None of these minigames are essential, but the player can ignore, what could be considered, an important part of the storyline to spend time playing games.

The disconnection of the player's priorities to the NPC's requires worldbuilding to separately identify the environments, and NPC worlds, from the player's in-game self. Although sometimes the player-character's wishes may align with the culture and NPCs of the world, their control by the player can remove them from that agency. Voices as environmental worldbuilding tools are often used to represent the people in the environment and their wills, as opposed to the representation of the player-character. The difference between character priorities in a game creates an organicism within the world, an identifier that, although the player is the 'hero' of the story, the game world continues to exist in its own state. The soundscape of a game thus needs to environmentally build this organic world, which occurs through the portrayal of its inhabitants and their culture.

## Spatial Immersion

The composer of the RPG *Persona 5*, Shoji Meguro, apologises to western audiences over the use of the female voice and English language as background music:

Apologies to the Western gamers, but to Japanese gamers, because many Japanese don't speak the language and they don't understand the lyrics, the English lyrics blend in as background music. The lyrics are not distracting to the gamer, so it's convenient to use as background music. If the lyrics were Japanese, the players would probably find the music too distracting from the game. Also, [some things] just sound cooler in English.<sup>2</sup>

The idea of 'cool' music as background sound in video games is not an unusual concept. This music as background sound, although can be used to represent character or story narratives, does not need to be an identifier of anything. This fundamental use of music as a musical background, that does not distract audiences, is introduced by Ben Winters in his paper 'Musical Wallpaper? Towards an appreciation of Non-narrating Music in Film'. Winters discusses the likelihood of film

---

<sup>2</sup> Reggy, 'Glixel: Interview with Composer Meguro and Artist Soejima About the Development of *Persona 5*', *Persona Central*, 8 May 2017 <<https://personacentral.com/glixel-interview-composer-meguro-artist-soejima-persona-5/>> [accessed 4 July 2019].

musicologists to consider all music as a narrator role, presuming ‘music’s almost omniscient ability to impart ‘secret’ narrative information.’<sup>3</sup> Winters himself considers music in multimedia from its non-narrative, atmospheric, components which he argues has been overlooked by film musicology. This composition of a ‘musical wallpaper’ is outlined clearly by Shoji Meguro. Whereas Western audiences and musicologists may have initially attempted to find meaning in the non-diegetic English voices of *Persona 5*, in actuality the music is fulfilling the space as a ‘musical wallpaper’ that does not distract non-English speaking players. Winters’ distinction of atmospheric music is contained to linear multimedia and does not consider the interactivity of video games. Although acting as a ‘musical wallpaper’, atmospheric video game music must still follow the player through the game world and adapt to where they are, acting as environmental information. Environmental information does not fall under Winters’ association of narrative identifiers, but it must still adapt to signify the player’s location. This sound information refers to what Cheng considers in *TF2* as ‘environmental sounds’, sounds that the player cannot manipulate, such as rain, train whistles, and sirens, but signify gameplay environments; these sounds differ from what Cheng considers as player-character enforced sounds such as ‘battle sounds’ (explosions, gunfire, and melee combat) and ‘character sounds’ (death cries, footsteps, and taunts).<sup>4</sup>

Environmental worldbuilding can be constructed through music as environmental information. Whereas immersion considers the involvement of the player in a game world as a whole, spatial immersion considers an involvement with game information that includes geographic locations, historic or future timeframes, terrains, and weather systems of a world. These physical signifiers, whether as background music or gameplay feedback, can immerse the player in the lore of the world. Thus, this can inform engagement and a suspension of disbelief with NPC behaviours and cultures through systems such as class, belief, aesthetics, and location based vocal manifestations such as accent.

The player is active in their spatial immersion as their agency of movement through in-game environments means the soundscape must react to their spatiality. This spatial involvement, or audio worldbuilding of environments, can be witnessed from the player’s point of audition which mirrors the point of view and audition of their character model. Michel Chion’s ‘point of audition’ considers the perception of sound from a given space, alongside an alignment with an on-screen character that can hear the same spatial and subjective sounds.<sup>5</sup> Whereas the point of audition of film and television may be set with its linear storytelling, point of audition can be fluid in video

---

<sup>3</sup> Ben Winters, ‘Musical Wallpaper? Towards an Appreciation of Non-Narrating Music in Film’, *Music, Sound, and the Moving Image*, 6.1 (2012), 39–54 (p. 44) <<https://doi.org/doi:10.3828/msmi.2012.5>>.

<sup>4</sup> Cheng, ‘Sound Play’, p. 147.

<sup>5</sup> Chion, ‘Audio-Vision’, pp. 90–93.

games as sounds follow the player's movement through the game world. In modern three-dimensional video games, technological advancements allow player spatiality. If the player turns their camera towards or away from a sound, i.e. birdsong, the sound directionally pans and adjust its velocity accordingly. It needs to be emphasised, however, that depending on the type and style of video game, the spatiality of sound can either be attached to the player-character model or to the player's camera. In first-person games, the player and their character model are in the same spatial position, the player embodying the character model, and thus the point of audition and point of view is attached to the player-character model. Third-person games, however, put the players point-of-view over the shoulder of the character model and thus sound is heard from the perspective of the player's camera. The point of audition must always be with the player's point of view in order to provide the player with multidimensional audio information in replacement of their natural sense of hearing; where the player's point of view is limited by the direction of the camera, audio can provide information on environmental happenings and inhabitants, such as combat and enemy interactions. Whether in first-person or third-person mode, however, the player and character always hear the same sounds through joint-audition. Some RPGs, such as *Horizon: Zero Dawn* and *The Witcher 3: Wild Hunt* (both third-person games) have features where the player can actively press a button to make the character focus on listening. The sounds gathered from this 'focus' are directly from the apparent point of audition of the character, but the sound's implementation and orientation are still from the perspective of the player. There are exceptions to this interactive point of audition, however. The inclusion of cinematic cut-scenes, a linear filmic scene where the player loses control of their character and becomes an audience member, removes the player from their audio and visual agency. As the player is demoted, characters regain their agency and react to environments with their own visual and audio senses, rather than their usual joint audition; spatiality in a video game thus requires two modes, cinematic and interactive, as environments are built to exist coherently within both.

Whilst discussing immersion within RPGs, Sarah Lynne Bowman identifies the environment as an aspect of spatial immersion, an occurrence of player movement, and involvement, within the environment of a game, whether a virtual RPG, LARP, or a TTRPG.<sup>6</sup> Unlike LARPing which occurs in a physical space, video games and TTRPGs lose certain senses, such as smell, touch, taste and sometimes sight. These environments, however, do not need to mirror our own. Gordon Calleja's discussion of environment as immersion consists of 'the player who assimilates this game world into their gameplay as a metaphorically habitable environment'.<sup>7</sup> The sense of a habitable environment

---

<sup>6</sup> Deterding and Zagal, *Role-Playing Game Studies*, p. 385.

<sup>7</sup> Calleja, 'In-Game', p. 27.

in a video game is created through environmental worldbuilding. There is a fundamental lack of need to survive in a game's environment (except for survival or simulation genres). The player-character of an RPG often does not need to go through the action of eating, drinking, or sleeping in order to survive multiple hours of gameplay, and in-game days. When bodily functionality is included in a game it is often linked to a game mechanic, such as regaining health, rather than to maintain a character's physical and mental health. This allows for environments to be diverse in space, place, and time. The exception to the diversity in environment styles, however, is its cohesiveness. If the visuals and audio do not match to what the player may expect, from subjective experiences, then it can cause issues to spatial immersion; a crude example of this would be the inclusion of winter music, often created by melodies on a glockenspiel, in an environment set in the height of summer. Sound design and environmental ambiance are important in this process. The inclusion of sounds such as footsteps, birdsong, wind sounds, the communication of NPCs, and other natural sounds are needed to create a sense of life, organicism, and functionality in the world.<sup>8</sup>

### ***Traversing Environments***

Travelling through a functional environment and soundscape can actively change the environment type the player is in, from varying climates to infrastructure, cityscapes to wilderness. As players are given access to an expansive world, the developers often provide a means for the player to fast travel in order to revisit locations. Fast travel allows players to instantly move from one environment to another of their choosing. Soundscapes and visuals may need to greatly emphasise each environment to remind the player of where they are when they travel between unconnected locations. Calleja considers spatial involvement as the player's ability to cognitively map their environment, or traversable space.<sup>9</sup> Leaving identifiable landmarks such as buildings and extra geographic context, such as climates, alongside a varying soundscape can orientate the player as they move.

This, often free, movement in RPGs can lead to inconsistencies in the soundscape. If there is a change in the visual environment the soundscape must also represent this change effectively. Interactive audio software tools such as Wwise and FMod allow composers to attach their audio to

---

<sup>8</sup> Simulation games such as *theHunter: Call of the Wild* (Expansive Worlds, *TheHunter: Call of the Wild*, Stockholm: Avalanche Studios Group (Microsoft Windows, PlayStation 4, Xbox One, 2017) do not include music in gameplay. The game focuses on environmental sounds, allowing the player to hear vocalisations of animals and their own noise pollution in this environment, such as footsteps and breathing.

<sup>9</sup> Calleja, 'In-Game', p. 75.



trigger points and thus create reactive sound to player movements. Attaching audio to triggers, whether these trigger points are related to character movements or environmental shifts, allows musical cues to adapt automatically to the game state. Vertical layering is a common compositional tool which allows different sections of audio within a soundscape to turn on and off depending on player action. Vertical layering can alter both soundscapes as a whole and aspects of a single cue, for example increasing the volume of a rhythmic section, reducing a percussion line, or turning on a combat theme. The idea of vertical layering is that every layer of a cue can exist independently whilst being able to work with the cue in its entirety. This adaptiveness of vertical layering allows sound to actively follow player action, rather than acting as a linear soundscape in an interactive medium; this linearity of soundscapes, seen in historic video games that lacked certain technological advancements, can cause repetitiveness and thus player boredom. Although there are diverse compositional practices to combat the indeterminable whereabouts of the player at any given time, vertical layering provides an essential adaptability that can be mapped to trigger points; the *Spyro Reignited Trilogy*<sup>10</sup> turns on the cue of a semiquaver hand percussion when the player moves Spyro quickly, matching the speed of Spyro's movement, to signify the faster movement around the environment.

The player's ability to navigate the game world's environments can vary on a scale between free player movement and determined linear movement, decided by the developers. The state in which the player can move in the environment often changes the implementation of music. The open-world RPG *Final Fantasy XV*<sup>11</sup> and open-world action-adventure game *The Legend of Zelda: Breath of the Wild*<sup>12</sup> allow their players to explore the entirety of the game world with relative freedom. The accompanying music of this freedom of movement is atmospheric and minimalist as the player indeterminately navigates various environments and space. The inclusion of minimalist non-diegetic musical cues, that play indeterminately across the environment, engages the open-ness of the visuals and gameplay. Certain cues are triggered by actions in both games, such as riding a steed, whereas other cues are narratively driven, providing diversity amongst the minimalist environmental soundscapes. This minimalist use of music in the exploration phase of a game, however, relates to Gibbons' acknowledgement that gameplay does not always need musical accompaniment to be immersive, in *Shadow of the Colossus*.<sup>13</sup> The music does not guide the player in any direction and

---

<sup>10</sup> Toys for Bob, *Spyro Reignited Trilogy*, Santa Monica: Activision (Microsoft Windows, PlayStation 4, Xbox One, Nintendo Switch, 2018).

<sup>11</sup> 2.

<sup>12</sup> Nintendo EPD, *The Legend of Zelda: Breath of the Wild*, Kyoto: Nintendo (Nintendo Switch, Wii U, 2017).

<sup>13</sup> Gibbons, 'Wandering Tonalities', in Donnelly, Gibbons, and Lerner, *Music in Video Games*, pp. 124–25.

does not enforce emotions such as haste. Instead, the minimalist music allows the player to travel in any direction they wish allowing physical, visual, and aural freedom.

The opposite of this freedom is when players are met with linear pathways. In games that provide players with relatively free movement, the inclusion of this linearity, in its gameplay, visuals, and audio, is often affected by the story. The action-adventure stealth game *Metal Gear Solid 3: Snake Eater*<sup>14</sup> (*Snake Eater*) uses a linear pathway to signify the player's approach to the end of the game. The player must climb an excessively long ladder before they arrive at their final destination which is used to pace these final scenes, allowing time for the player to reflect on the game and to become mentally ready to fight the final boss; see Figure 3. The ladder climbing sequence in *Snake Eater* is notorious in the gaming community due to its inclusion of an almost two-minute-long song, 'Snake Eater', performed by an acapella female voice. This scene, a source of humour for a lot of gamers, is accompanied by the diegetic sounds of the tunnel and the player-character's feet on the ladder as they climb. At the end of the climb, as the player-character pulls themselves up onto the next floor, the song stops, and the words 'snake eater' are whispered as a vocal conclusion to the game. The soundscape that the player has become used to during the game has been replaced for a spiritual reflective moment within these two minutes.



Figure 3. *Metal Gear Solid 3: The game's excessively long ladder sequence. Screenshot by Author.*

The song is not random, however. 'Snake Eater' is the name of the cue that is played over the opening introduction of the game. The song acts as a *James Bond* parody through its opening cinematics which includes several scenes from the game whilst the names of the developers are

<sup>14</sup> Konami Computer Entertainment Japan, *Metal Gear Solid 3: Snake Eater*, Tokyo: Konami (PlayStation 2, 2004).

shown. The audio reflects this parody further as the female vocals that perform 'Snake Eater' are accompanied by Hollywood strings and a brass band; lyrics such as 'someday you'll feed on a tree frog', whether this was meant to be literal or serious by the composer, helps cement the song as a parody.

The representation of geography or location, that accompanies the diversity of movement in video games, is important as a worldbuilding tool. In RPGs it is common for geographies to change as the player should have access to various locations, such as a palace dungeon, a desert, the ocean, or a grassy plain. Movement between different areas can be considered as a type of migration, especially in games such as *Divinity: Original Sin II* in which the player cannot revisit areas once left. Often, as John Connell and Chris Gibson state, 'the music of migrants typically involved such themes as solitude, homesickness, nostalgia, unemployment, racism (and other hardships) alongside the longing for their converse – kin, community, familiar foods and sights, and a place in society.'<sup>15</sup> This migration of music is heard within the inclusion of leitmotifs in multimedia. The adaptation of soundscapes, as players and audiences migrate through a medium, often still longs for leitmotifs and themes presented in previous scenes and locations. This longing and 'homesickness', outlined by Gibson and Connell, is shown in the migration of sound and acceptance of a new geography in multimedia, through the combination of new area signifiers with a known leitmotif. Connell and Gibson also explain that 'different destinations provided different outcomes and expectations.'<sup>16</sup> In the RPG, new environments often provide different challenges for the player, and thus soundscapes may alter musical styles to reflect these distinct geographies and NPC cultures. Voices are a common tool to emphasise these migrations. Accents are used to highlight differences between cultures and geographic locations, acting as representations of that environment. Regional accents, and the use of vernacular terms, support environmental migrations throughout a world to orientate the audience or player. *GoT* uses regional British accents and language to show where a scene is geographically at any given time. The combined use of Yorkshire, Lancashire, and Northumberland accents are used to portray the north of Westeros, whilst Queen's English accents are used to define the south.

---

<sup>15</sup> Connell and Gibson, 'Sound Tracks', p. 161.

<sup>16</sup> Connell and Gibson, 'Sound Tracks', p. 161.

## Worldbuilding People and Culture

Spatial immersion can be a component of this reactive soundscape. Although voices can aid spatial immersion, voices are not always used within the soundscape, especially when games have budget constraints, or short translation and dubbing times. The English dubbing of the action-JRPG *Ni No Kuni 2: Revenant Kingdom*<sup>17</sup> (*Ni No Kuni 2*) uses professional voice actors but, due to unknown constraints, the game has minimal voice acting throughout; the infrequency of voice acting requires the player to read most other dialogue outside of narrative plot points and short animated sequences. To portray environments with this minimal use of voices, *Ni No Kuni 2* contains stereotypical non-diegetic cues of music to overly signify the culture and people of each environment. *Ni No Kuni 2*'s city environment Goldpaw is a casino-run, stereotypically 'Chinese', gambling city whose inhabitants are humanoid dogs. Goldpaw's inhabitants follow the stereotyping of Chinese culture through their names, which also reflect their dog-like aesthetics (Long Mein, Yip Yip, Min Ti, Li Li). Although accompanied by a western style orchestra, the musical cue for Goldpaw uses gongs alongside the pentatonic scale to portray the gambling city. Whereas the majority of the soundscape does not use non-diegetic voices, when voice acting is implemented *Ni No Kuni 2*, alongside its predecessor *Ni No Kuni: Wrath of the White Witch*,<sup>18</sup> there are high variations of accents to highlight the diversity of the world's inhabitants.

The use of accent and stereotypical identifiers to enhance cultural assets of the spatial environment can impact the wider empathetical scale of the player. Alongside identification with the player-character or avatar, the signifying of in-game cultures can highlight the diversity of characters that exist within a game world, their denominations, societal conducts, and colloquial languages; character backstories, environment types, religious beliefs, accents, class, and societal needs are essential in representing a game's culture and to create the sense of a believable and authentic game world. The reflection of real-world cultures through diverse and colloquial uses of the voice help players to understand an in-game culture. This often requires heavy stereotyping of in-game people to recreate these recognitions of the real-world. Alongside Summers suggestion that music contextualises the cultures of a game world, so too does music have a stereotypical relationship with these identifiers:

---

<sup>17</sup> Level-5, *Ni No Kuni II: Revenant Kingdom*, Tokyo: Bandai Namco Entertainment (Microsoft Windows, PlayStation 4, 2018).

<sup>18</sup> Level-5, *Ni No Kuni: Wrath of the White Witch*, Tokyo: Namco Bandai Games (PlayStation 3, Microsoft Windows, Nintendo Switch, PlayStation 4, 2013).

[...] as part of musical connections that draw upon, and feed back into, musical cultures outside the game, game music has the potential for complicated relationships with actual-world musical cultures and histories (whether by misrepresenting musical histories, or by exoticizing and stereotyping cultures associated with particular musical styles).<sup>19</sup>

This misrepresentation of culture surrounds and interacts with the player and their character through these recognisable, transmedia signifiers.

As outlined by *GoT*, medieval and fantasy genre multimedia often use regional British accents to represent a 'dark ages' style history and culture. Tolkien-inspired fantasy texts such as the *LoTR*, *The Hobbit* film series,<sup>20</sup> *The Witcher* television series,<sup>21</sup> and the *World of Warcraft* video game use these accents to highlight both 'fantasy' and 'medieval' aesthetics. This use of accent to represent an 'other' may be caused by the collective understanding of a misrepresented culture, as highlighted by Summers, which audiences believe to be authentic. The collective understanding of what is the norm for these aesthetical characteristics is outlined by Tanya Krzywinska's discussion of *WoW*:

It can be said that most popular cultural artifacts are reliant on intertextual features for the generation of meaning and recognition. These are in part an outcome of genre production, but within certain genres, such as science fiction or fantasy, these are actively deployed to generate what Roz Kaveney (2005) called a "thick text;" in other words, a text richly populated with various allusions, correspondences, and references.<sup>22</sup>

The intertextual features of genre specific accents often seem to consider stereotyping of cultures for the sake of audience recognition. The act of recognition is also important when texts migrate between unfamiliar landscapes. Thus, the stereotypical use of the 'northern' British accent in fantasy medieval genres, to signify the north of a country, can efficiently translate these aspects to an audience. *Ni No Kuni 2*'s shipyard environment, Capstan-upon-Hull, boasts the Yorkshire regional accent of the UK (likely mimicking Hull accents) and parodies colloquial language through dialogue and the naming of areas such as 'Reight Bright Beacon' (Right Bright Beacon). This stereotypical nature of accents can also be applied to race, for example the use of Scottish accents for Dwarves, cockney accents for Orcs, and west country accents for the Hobbits in *The Lord of the Rings*. The pairing of race to accent, however, is problematic as it may invoke an association with the exotic 'other', as Mark Brill considers:

---

<sup>19</sup> Summers, 'Understanding Video Game Music', p. 101.

<sup>20</sup> Peter Jackson, *The Hobbit (Film Series)*, New Line Cinema & MGM (Warner Bros. Pictures, 2012).

<sup>21</sup> Lauren Schmidt Hissrich, *The Witcher (TV Series)*, *Los Gatos: Netflix* (Netflix, 2019).

<sup>22</sup> Tanya Krzywinska, 'World Creation and Lore: World of Warcraft as Rich Text', in *Digital Culture, Play, and Identity: A World of Warcraft Reader*, ed. by Hilde G. Corneliussen and Jill Walker Rettberg (London: MIT Press, 2008), p. 123.

Another signifier of Otherness is the use of accents. [...] The same dynamic works itself in *Jason and Titans*, where the heroes (cast with American actors) and the gods in Olympus (cast with British actors) are meant to evoke recognition from the audience. Conversely, King Aeëtes' inflections are amorously foreign, and Sokurah's accent appears to be a mixture of Yiddish and Russian – thus, to 1950s audiences, dangerous.<sup>23</sup>

Janet K. Halfyard also notes that British accents are used to signify villainous characters in Hollywood, providing a sign to audiences that these characters are evil, whilst American accents are given to the hero characters.<sup>24</sup> This handing of accents allows directors to create associations with, what Brill examines as, dangerous or hero characters. This inclusion of an exotic other through voices can cause issues as cultures are romanticised, rather than providing historical accuracy, much like the use of playback singers to romanticise India.

Stereotyping voices in video games, however, can highlight differences in the backstories and cultural differences of various characters within the same game space. *RDR2* provides a coherent and consistent aural and visual relationship in a wild-western world, following the outlaw Arthur Morgan and his community. The soundscape of the open-world game consists of reactive audio, controlled by triggers, and fifteen layers of sound that transition between narrative, environmental, and open-world tasks.<sup>25</sup> Environmental and cultural worldbuilding is conducted in the game's camp where the player can rest, purchase and sell goods, fast travel, and engage with NPCs. The camp naturally runs itself and the player may choose to interact with characters naturally, as opposed to interactions only being a part of the main story. When providing a sense of community, NPC culture, and worldbuilding, the game incorporates both conversational dialogue between NPCs and campfire songs, executed without the player's interaction.

---

<sup>23</sup> Mark Brill, 'Fantasy and the Exotic Other: The Films of Ray Harryhausen', in *The Music of Fantasy Cinema*, ed. by Janet K. Halfyard (Sheffield: Equinox Publishing Ltd, 2014), p. 22.

<sup>24</sup> Janet K. Halfyard, 'Superconductors: Music, fantasy and science in *The Sorcerer's Apprentice*', in Halfyard, *The Music of Fantasy Cinema*, p. 225.

<sup>25</sup> Lily Moayeri, 'Woody Jackson and Vox Studios - Mixonline', *MIX*, 11 January 2019 <<https://www.mixonline.com/recording/woody-jackson-and-vox-studios-red-dead-redemption-2>> [accessed 12 April 2019].



Figure 4. *Red Dead Redemption 2*: Dutch's gramophone on the edge of camp. Screenshot by Author.

The gramophone seen in Figure 4, plays pre-existing songs, such as Giuseppe Verdi's 'Libiamo ne' lieti calici', whenever Dutch or his partner Molly wish to hear it. The player may never witness the gramophone playing, depending on when they visit Dutch's tent, but the device's use is not solely for the player or Arthur to hear.

Campfire songs also incorporate pre-existing music, such as the traditional folk song 'Jack of Diamonds', to portray the 19<sup>th</sup> century time-period and the gambling nature of the community's leader. Alternatively, the player may listen to the Mexican camp member, Javier Escuella, perform 'Cielito Lindo' which is assumed by Kenneth Marcus to be written in the 1880s; Marcus is uncertain of the authenticity of 'Cielito Lindo' and its musical style because of the time difference between publication and composition.<sup>26</sup> The song's use in *RDR2*, however, reflects the use of performance and language as an enhancer of recognition between the audience and character cultures. Marcus' uncertainty of authenticity is unlikely to bother players as the song is used to highlight Javier as an exotic 'other' in the otherwise American community.

The use of accents and language as signifiers, as shown from fantasy and medieval themes, is often used to highlight franchises across mediums. Research of film and video game crossovers identifies that franchises aid the building of a game world when the franchise is recognisable. Kevin D.

<sup>26</sup> Kenneth H. Marcus, 'Mexican Folk Music and Theater in Early Twentieth-Century Southern California: The Ramona Pageant and the Mexican Players', *Journal of the Society for American Music*, 9.1 (2015), 26–60 (pp. 38–45) <<https://doi.org/https://doi-org.libaccess.hud.ac.uk/10.1017/S1752196314000534>>.

Impellizeri examines the success of the first-person shooter game *GoldenEye 007*<sup>27</sup> due to the previous success of the James Bond franchise. Impellizeri identifies that although the game had issues with gameplay, due to the limitations of technology at the time, the experience of the game was enjoyable because of the franchise's 'attempt to capitalize on 007 as a pop culture figure';<sup>28</sup> players could experience 'being' the character James Bond, rather than passively engaging with Bond through non-interactive multimedia. Arguably there were two main selling points for *GoldenEye 007* as the game both used the 007 franchise as well as becoming one of the most iconic first-person shooter games of its time. However, Abe Stein and Matthew Weise take umbrage with this. Stein and Weise argue that the game's focus on gameplay, as opposed to the diversity of Bond as a character, took from the player's experience:

[...] video games have offered a comparatively shallow version of the figure. The games to date have emphasized action over the more nuanced characteristics. Three specific Bond characteristics, Bond as detective, Bond as secret-agent, and Bond as playboy, have been variously and diversely represented in video games, shaping player's experiences of the Bond figure.<sup>29</sup>

This is one of the key issues of transmedia franchises. The interactivity of video games often needs developers to focus on functional gameplay above the worldbuilding and cultural acknowledgements of a franchise. Stein and Weise's issue with the game not providing a nuanced character is due to this prioritisation of gameplay over character; movie tie-in games are often created quickly after a film is released in order to 'cash-in' on a franchise rather than spending time on each element of a game. To the defence of the 007 video game adaptations, however, the voices of Sean Connery (*James Bond 007: From Russia with Love*),<sup>30</sup> Daniel Craig, and Judi Dench (*007: Quantum of Solace*),<sup>31</sup> are recorded for the games to engage players with the characters of the franchise.

---

<sup>27</sup> Rare, *GoldenEye 007*, Kyoto: Nintendo (Nintendo 64, 1997).

<sup>28</sup> Kevin D. Impellizeri, 'Use your Joystick, 007: Video Games and the Interactive Bond Experience', in *James Bond in World and Popular Culture: The Films Are Not Enough*, ed. by Robert G. Weiner, B. Lynn Whitfield, and Jack Becker, 2nd edn (Newcastle upon Tyne: Cambridge Scholars Publishing, 2011), p. 8.

<sup>29</sup> Abe Stein and Matthew Weise, 'All Bang Bang, No Kiss Kiss? The Bond Figure and Video Games', in Weiner, Whitfield, and Becker, *James Bond in World and Popular Culture*, p. 25.

<sup>30</sup> EA Redwood Shores, *James Bond 007: From Russia with Love*, Redwood City: Electronic Arts (GameCube, PlayStation 2, Xbox, PlayStation Portable, 2005).

<sup>31</sup> Treyarch, Vicarious Visions, Eurocom, and Beenox, *007: Quantum of Solace*, Santa Monica: Activision (Microsoft Windows, Nintendo DS, PlayStation 2, PlayStation 3, Wii, Xbox 360, 2008).



## ***Gendered Religious Environments***

Religious environments in video games use voices to represent denominations. When the player enters a religious location, i.e. a church, they often meet a soundscape which incorporates chant, performed by non-diegetic male voices:

By contrast, the sacred world is almost always represented musically by vocal music, particularly reverberant monophonic chant within church settings. Countless examples abound in film and television, and few members of the general public would realise that there were any historical problems with such stark aural divisions between sacred and secular spheres, and the diversity of sound that they fail to capture. However, from a narrative perspective, such divisions serve a constructive purpose and play upon well-established recognisable aural cues. In a sense, a fidelity to the codes of filmic media, established over many decades, and the perceptions of what audiences may expect, might be seen to take precedence over historical accuracy.<sup>32</sup>

As Adam Whittaker suggests, even with sounds that are attempting to appropriate real-world traditions, to highlight multimedia diegeses, music still manipulates and appropriates cultures to suit multimedia audiences. The use of Gregorian chant as filmic norms is discussed by John Haines who describes its use in some films as ‘dubious chant’, breaking from its high tradition in order to aid multimedia in defining settings that are non-Christian, but medieval.<sup>33</sup> This idea of ‘dubious chant’ is used throughout video games to portray denominations through vocal cues that exhibit aspects of Gregorian chant. ‘Hymn of the Fayth’ is a unique environmental cue in *Final Fantasy X*<sup>34</sup> that emphasises the religious temple of the game by using sung voices that mimic a Gregorian chant style. The usually synthesised orchestral soundscape adopts male and female voices to identify the unique temple environment, accompanied by hand percussion and high notes on the traditional Japanese string instrument, the Kōkyū. Few instruments in the game are non-synthesised due to the original data constraints of the PlayStation 2, however. This meant that when instruments and voices were recorded, they heavily emphasised the environments. This identification of religion through

---

<sup>32</sup> Adam Whittaker, ‘Musical divisions of the sacred and secular in *The Hunchback of Notre Dame*’, in *Recomposing the Past: Representations of Early Music on Stage and Screen*, ed. by James Cook, Alexander Kolassa, and Adam Whittaker (London: Routledge, 2018), p. 90.

<sup>33</sup> John Haines, *Music in Films on the Middle Ages: Authenticity vs Fantasy* (Oxford: Routledge, 2014), pp. 120–21.

<sup>34</sup> Square Product Development Division 1, *Final Fantasy X, San Mateo: Sony Computer Entertainment* (PlayStation 2, 2002).

aesthetics, and sound, are used alongside devoted NPCs and pilgrims who may have dialogue which illustrates their devotion. Sound and voice are used frequently in expressing and stereotyping religious aspects in games. The home of the templars and British monk figures in *Dragon Quest VIII*<sup>35</sup> is represented by music performed on an organ when the player enters the area, highlighting the church style architecture. The government and religious environment of *Xenoblade Chronicles 2*,<sup>36</sup> the Indoline Praetorium, uses acapella voices in the non-diegetic cue 'Our Eternal Land'. The male monophonic chant of the cue includes soprano female voices which highlights a sacred lyrical text. Melismatic lyrics, such as 'it is in Father's wishes we live' and 'guardian angels, they too will keep us safe', encourage the identification of neo-Christian religious behaviours even though the population are worshiping fictional gods. Although female voices are included in the cues, the voices never crescendo and highlight lyrics without sharing the audio space with male voices.

Religious cultures are frequently stereotyped in RPGs, although the player-character themselves are often on the outside of the religion, and the often-appropriated soundscape is used to worldbuild these religious environments. Lars de Wildt examines the post-apocalyptic religious symbolisms in the video games *Fallout 3* and *Horizon: Zero Dawn*, initially comparing religion and video games as meaning making tools.<sup>37</sup> Wildt identifies the common use of religion in videogames as 'fiction-based', acting as a narrative space to provide a sense of 'real-world' meaning which players can relate too.<sup>38</sup> The gendering of voices in these religious environments aids this sense of 'normality' for the player, as the multimedia use of male voices to signify the neo-Christian is common in video games. When presented with female voices, without the accompaniment of the masculine, religious environments often portray paganistic aesthetics, and the female voices sing outside of Gregorian chant influences. The female goddess, the 'All-Mother', in *Horizon: Zero Dawn* is worshipped by the player-character's tribe named the Nora. The Nora's goddess is essentially a 'false' god as the player-character, Aloy, recognises the goddess's 'womb' as a mechanised door, which the Nora refuse to believe. Aloy is more technologically advanced than her peers as she obtained a mechanical device from the 'Old Ones' (pre-apocalyptic earth) in her childhood years, and thus became accustomed to the aesthetics of mechanical objects. When the 'All-Mother' speaks, Aloy identifies the voice as a computer system through hearing a female voice not too dissimilar to Amazon's Alexa; Wildt identifies that 'it immediately becomes clear for the player and Aloy that All-Mother is nothing more

---

<sup>35</sup> Level-5 and Square Enix, *Dragon Quest VIII*, Tokyo: Square Enix (PlayStation 2, Android, iOS, Nintendo 3DS, 2004).

<sup>36</sup> Monolith Soft, *Xenoblade Chronicles 2*, Kyoto: Nintendo (Nintendo Switch, 2017).

<sup>37</sup> Lars de Wildt, 'Things Greater than Thou': Post-Apocalyptic Religion in Games', *Religions*, 9.6 (2018), 1–20 (p. 1) <<https://doi.org/doi:10.3390/rel9060169>>.

<sup>38</sup> Wildt, 'Things Greater than Thou', p. 2.

than a huge mechanical door controlled by an AI that ‘speaks’ with a female-like computer-generated voice’.<sup>39</sup> As the Nora are considered backwards in the eyes of the game world, this becomes reflected in the eyes of Aloy and the player because of their knowledge of computer systems. The game acknowledges the paganistic ‘backwardness’ of worshipping the ‘All-Mother’ through the inclusion of female song during a cultural and religious ceremony. During this ceremony, the player can walk around, talk to Nora, and attend stories and performances in real time. One performance the player can witness is the song ‘Nora u Norawea’ which includes a call and response of lyrics between female voices with no musical accompaniments, including syllabic sounds and screaming. The symbolism of a ‘primitive’, paganistic and ritual based civilisation which worships the ‘All-Mother’ goddess, is thus sonically symbolised by the female voice. This call and response manner of singing is a stark contrast to the religious performance given for the second religion found in the game. The ‘Sun God’, which is never highlighted as a ‘false’ god, is worshipped in the progressive city of Meridian. The religious performance that represents the ‘Sun God’ is sung by an entirely male chorus of sun priests who use ‘Gregorian chant’, or Haines’ ‘dubious chant’, without the movement and call and response aspects of the feminine ‘All-Mother’ religion. The vocal performances of ‘Song to the Sun’ (complete with ‘Dawning’, ‘Midday’, and ‘Evening’ variations, as the Carja track the sun) portrays a more advanced religion. The neo-medieval, all-male Gregorian chant includes drones alongside a solo male vocalist, and elements of polyphony. Vocalisations, such as the syllabic ‘ah’, is still used in ‘Song to the Sun’ however the vocalisations lack the shouts and screams heard within ‘Nora u Norawea’.

### ***Mythological and Supernatural***

Mythological and supernatural occurrences in the environment of games are not too dissimilar from religious voices and locations. Mythological and supernatural environments are used in games that are abstract in narrative and focus on exploration and puzzle-based gameplay. Voices in these environments are often used as guidance to enhance playability where some players might find an issue with the lack of a strong narrative.

*The Talos Principle*<sup>40</sup> is a first-person, role-playing puzzle video game with an abstract narrative. The game begins in a computer-generated garden environment with references to Christianity’s garden of Eden. It is unknown to the player, as they complete various puzzles and expand the narrative,

---

<sup>39</sup> Wildt, ‘Things Greater than Thou’, p. 11.

<sup>40</sup> Croteam, *The Talos Principle*, Austin: Devolver Digital (Microsoft Windows, Linux, OS X, Android, PlayStation 4, iOS, XBOX One, 2014).

whether the environment they are within is ‘real’, virtual, post-apocalyptic, or a type of heaven or hell. The game consists of multiple gardens which are held within three different temples, each with a unique environment and soundscape. The first temple the player has access to assumes the visuals of the ruins of a Roman garden, the second temple appears as ruins within an Egyptian desert, and the third exists as a medieval Christian monastery; see Figure 5. The soundscape of the game adapts to each environment presented to the player through the use of varying instrumentation between each temple; for example, the Egyptian temple area cues include a Goblet drum (an Egyptian hand percussion) and Egyptian flutes to highlight the Egyptian mythology and desert environment. The composer, Damjan Mravunac, states that the overall aim of the game’s composition was to ‘create more than a dozen mystical, atmospheric, thought-provoking (but not distracting) and not-annoying-when-looped tracks.’<sup>41</sup> Each temple has seven gates, which house the puzzle gardens, that transport the player to different historical and mythological influenced environments. The player must complete the puzzles to unlock temples and complete the game through one of its many endings.

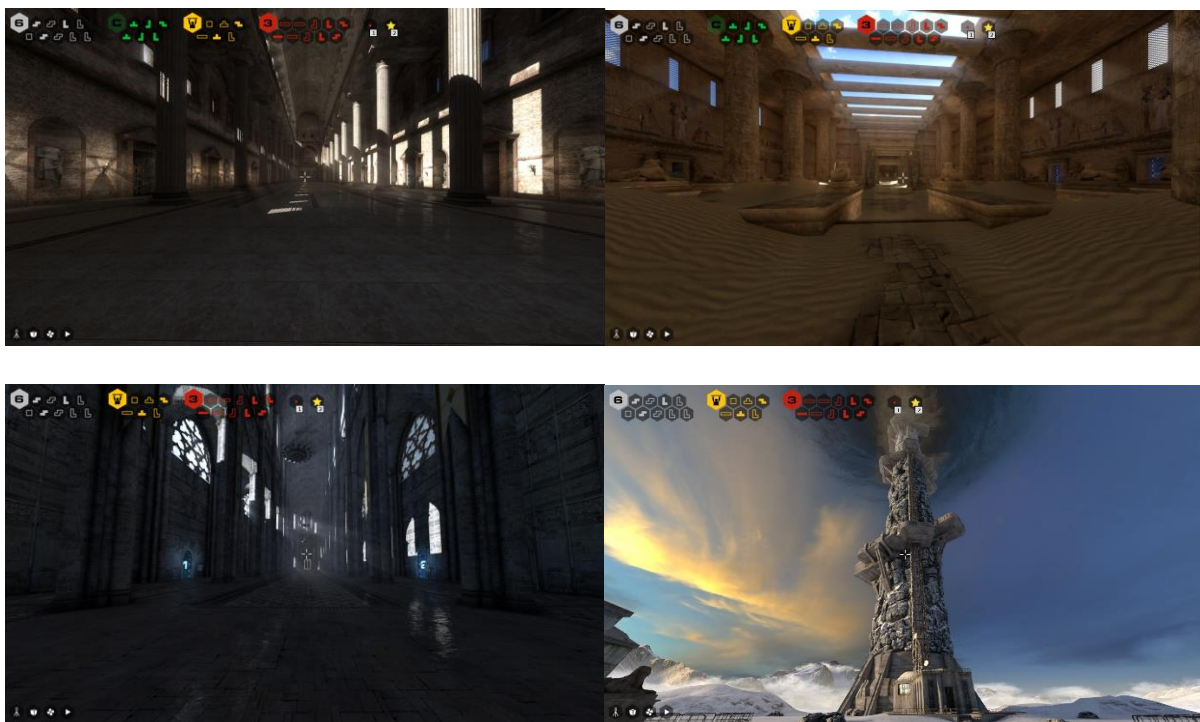


Figure 5. *The Talos Principle*: From left to right – Roman garden ruins temple, Egyptian temple, medieval Christian monastery, and the Tower. Screenshots by Author.

Whilst the temples focus on the historic, the fourth environment includes a tower which invites an industrial and technological environment that disrupts the narrative and environmental continuity;

<sup>41</sup> Thomas Quillfeldt, ‘Interview: Talos Principle Composer on Humanity in the Digital Realm’, *Laced Records*, 11 October 2017 <<https://www.lacedrecords.co/blogs/news/interview-talos-principle-composer-on-humanity-in-the-digital-realm>> [accessed 9 July 2018].

the tower also remains distinct from the temples as the player is forbidden to ascend it. Before the player is coerced to attempt to ascend the tower by the in-game computer, they are guided through the game environments by the non-diegetic voice-over Elohim. Elohim is a male leaning voice that proclaims to be the creator of the world, the multitude of puzzles, the android player-character, and the generations of puzzle-solving androids that existed beforehand. With the player having mostly free movement amongst each environment, the soundscape adopts the ambient and minimalist style of music, discussed previously, to relax the player and provide them with time and space to solve the puzzles. The ambient music also allows Elohim to dominate the sound world when he speaks to the player, providing him with complete agency over the world as a 'god' might have. Elohim's voice, however, does not reach the tower, and thus the entity does not want the player to explore that environment as it reduces his control over the player-character.

Elohim's voice-over follows the player-character's development in the narrative of the game. Alongside the voice of Elohim, sung voices can be heard in the temple areas as short cues that fit with the current environment and mythology. The non-diegetic voices and vocalisations are sung as not to overlap with Elohim's monopoly on the spoken voice. The non-diegetic cue for the first temple, 'Sanctuary', begins with female vocal whispers accompanied by a choir which crescendos on differing notes on an angelic 'ah' syllable, accompanied by bell percussion. The Egyptian temple cue, 'Temple of Sands', uses short cues of a solo syllabic male voice, and the medieval Christian monastery uses a sampled Gregorian choir, that reflect the 'dubious chants' of religious environments, which also accompanies the main menu cue 'Virgo Serena'. The opening cue of the game, 'Welcome to Heaven', emphasises the supernatural and religious symbolisms that the player will be introduced to. The Treble voices accompany the visualisation of this 'heavenly' or 'holy' environment keeps the player audially trapped in the mythological temple and away from the tower.

Dialogue and the spoken voice in the diegesis are key to the worldbuilding of a game's people. As seen with Elohim, his agency as a 'god' character is enforced through his domination within the soundscape. *Everybody's Gone to the Rapture (EGR)* is an artistic first-person adventure game that has an orchestrated soundtrack that dominates the soundscape. Set in a quaint English, 1980s Shropshire, village, the story of *EGR* follows the disappearance of all the village's inhabitants due to a supernatural, apocalyptic, phenomenon. To identify this British village, which has a connection to the church and Christianity, the soundscape includes a strong choral element alongside the orchestra. The game relies on the cue of these syllabic choral voices to identify the religious culture and characteristics of the village; the village thus maintains a level of hope that the disappearances are due to a godly rapture rather than a pandemic. As the player must walk around the village and interact with resident's memories, which remain in the form of a ball of light, the soundscape

remains atmospheric with short vocal and musical cues to stop the environment from becoming stale. The music includes these high treble and soprano voices alongside strings and very few bass instruments. The high frequency of the soundscape provides an ethereal feel which highlights the tranquillity of the village and its relationship with Christian ideologies. Greater uses of the orchestra, alongside male voices in the choral sound, is therefore only implemented to dominate the soundscape when the player is reaching narrative conclusions. Whilst the player is roaming the village and discovering memories, the soundscape of *EGR* follows that of *The Talos Principle* through an ambient and atmospheric soundscape; the soundscape focuses on creating a believable diegesis through diegetic wind sounds and bird song. This diegetic ambience allows the dialogue of the character memories to be at the forefront of the soundscape, providing them agency in the narrative.

### **Adaptive Environments: *NieR: Automata***

Collins identifies that spatial positioning and an inclusion of dynamic sound are important to the interactive audio prompts expected in video games.<sup>42</sup> Adaptive sounds are unique to video games due to this interactivity. Adaptive audio is imperative to environmental worldbuilding as the music follows the spatial environments and situations of a video game, changing audio as the player moves between areas, combat styles, and gameplay interactions.

*NieR: Automata (NieR)* is an action-RPG set in a post-apocalyptic environment where aliens and mechanical beings have captured the earth, and humans are said to have escaped to the moon. The player has access to three player-characters throughout the game, androids 2B, 9S, and A2 who are tasked with saving the earth from the machines. The player must traverse a variety of environments throughout the game to track enemy machines and restore the earth to a habitable state. The environments are diverse in aesthetics ranging from an over-grown city to a desert, and an amusement park to a forest; see Figure 6. The combat and game mechanics of the game are also diverse, flowing between differing styles and acting as Ivănescu's 'nostalgia game' by commenting on gaming pasts:

*NieR: Automata (NieR)* uses a variety of game play styles, tapping into many game genres and moving the player between: different third person camera angles, side scrolling platforms, 3D open world movement, 2D movement, shoot 'em ups, and bullet-hell styles. Therefore, the

---

<sup>42</sup> Collins, 'Game Sound', p. 101.

music must adapt to communicate the game's ever-changing intentions and fluctuating characters' emotions, to the player, during these multiple game play types and interweaving narratives.<sup>43</sup>

The variations in gameplay, environments, and player-characters is supported by the adaptivity of the soundscape. As discussed, creating spatial immersion and worldbuilding an environment is common in games through area music that triggers depending on where the player is, often following the player physically. However, *NieR's* adaptive soundscape goes beyond this as it considers the player's narrative progression alongside their physicality.

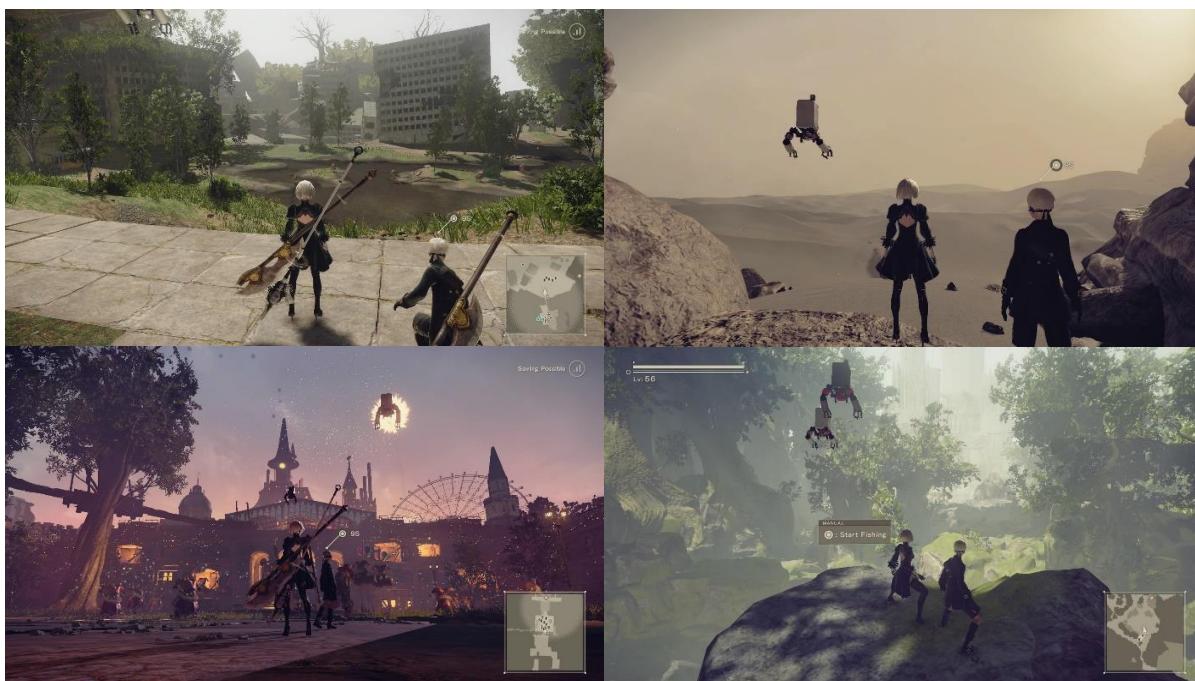


Figure 6. *NieR: Automata*: environments from left to right - City Ruins, Desert, Amusement Park, Forest. Screenshots by Author.

Each non-diegetic musical cue of *NieR* is broken into three variations, each with or without the sung female voice: quiet; medium; dynamic. The quiet variation of a cue includes only atmospheric sounds and accompaniment. The medium variations introduce the leitmotifs and percussion, and the final dynamic variations introduce the cue of music in its entirety, which often includes a countermelody. When the player initially enters an environment, the quiet cue is played to identify the environmental style of the area, allowing players to recognise the real-world associations of the environment and understand where they are. As the player traverses the environment and completes some of its challenges, combat, or narrative, the medium cue is introduced and

<sup>43</sup> Jennifer Smith, 'A Beautiful Song: The Adaptive Music of NieR: Automata', *Medium*, 4 April 2019 <<https://medium.com/the-sound-of-ai/a-beautiful-song-the-adaptive-music-of-nier-automata-4b29e5d7f2f3>>.

establishes the overarching leitmotif of the area. Finally, as the player makes their way towards the final challenge of an area, usually combat with a boss type enemy, the dynamic cue is introduced. With the inclusion of the medium and dynamic adaptations of the cue, especially in the desert environment, voices are also introduced. The non-diegetic sung voices of *NieR* are mostly only female because the majority of the androids, except for 9S, are female. These female voices use an invented language throughout the game, differing from the real languages used in dialogue, with the exception of the English and Japanese versions of the end credit sequence cue 'Weight of the World'.

### ***Voice, Language, and the Adoption of the Human***

In the desert environment specifically, the inclusion of the non-diegetic sung voice in the medium and dynamic cues are used to portray the blurring of moral decisions. Relatively early in the game, the player enters the desert environment only knowing that the machines are enemies of androids and humans alike, and thus they must be destroyed. The quiet cue of the desert's soundscape, 'Memories of Dust,' introduces the player to the environment and the slight change in movement as the player-character runs slower on sand and can slide down dunes; this cue, alongside visuals and haptic gameplay, situates the player within their stereotypical understanding of a desert. The medium cue of 'Memories of Dust' occurs when the player enters the apartment complex. This complex is said to be where humans once lived, but now the apartments have become a part of the desert as they have deteriorated over thousands of years; see Figure 7. The medium cue here includes the female voice in the non-diegetic space as the player is introduced to human properties in the visuals.





Figure 7. *NieR: Automata: Desert Environment – Apartment Complex – Beginning of the medium cue and vocals. Screenshot by Author.*

The composer Keiichi Okabe discusses that his methods in composing for *NieR* was that ‘first, I’ll think about the visuals and the situation, and then put my heart and soul into creating sounds that will make a given scene even richer.’<sup>44</sup> The voice of J’Nique Nicole accompanies the desert cue alongside instrumentation, such as the West African Kora and hand percussion, which portrays a stereotypical hot, desert environment; the environment also has a stronger and harsher timbre to other environmental cues. The voice includes mournful aspects, such as vocalisations and vocal cries, alongside an invented semantic text that highlight the loss of humans in this area. Alongside this, dialogue from the machines is introduced in this area where the player has not heard the machines speak before. The machines act with personal agendas in the apartment complex, running away from the player-character rather than attacking, and speaking short sentences such as ‘You. Not. People.’, ‘Run! Run!’, ‘Hello. How are you?’, and ‘Nice. Weather. Today.’ in robotic voices. As the mournful female voice is performed over this, it starts to become obscure as to whether this voice is representative of the humans, or the human qualities and clear distress of the machines.

Finally, as the player enters the final area of the desert, the environment’s boss-fight arena, the dynamic cue enters with the continuation of the sung voice.<sup>45</sup> The player enters this area to witness a group of peaceful machines that are trying to procreate and look after ‘children’ in an empty crib;

<sup>44</sup> Keiichi Okabe, ‘One Year Later, Composer Keiichi Okabe Looks Back at *NieR:Automata*’, *Playstation.Blog*, 7 March 2018 <<https://blog.us.playstation.com/2018/03/07/one-year-later-composer-keiichi-okabe-looks-back-at-nierautomata/>> [accessed 4 July 2018].

<sup>45</sup> A boss-fight is a colloquial term in video games for an enemy that is fought at the end of a video game’s environment. A boss often exhibits greater strength and some relation to the narrative.

the machines again speak obscure fragmented sentences: ‘Child. Child. Child.’; ‘My. Love. My. Love’; ‘Together. Forever. Together. Forever.’; Figure 8.



Figure 8. *NieR: Automata*: An unexpected peaceful machine in the final Desert environment. Screenshot by Author.

Whilst playing through this scene as 2B, and witnessing peaceful machines that are becoming somewhat human, your partner 9S suggests that ‘They don’t have any feelings.’, ‘They’re just imitating human speech.’, and ‘Let’s take them out.’ As 9S suggests destroying the machines, the female voice of the cue is introduced.

The player cannot choose to not destroy the machines and must start attacking them in order to progress the game. There is no reason, in this moment, to assume that the game is making any comment on the player’s aggressiveness. It is common in most games to progress a narrative through combat, and the player has been told at the beginning of *NieR* that the machines are the antagonists and need destroying. In my playthrough of the game, immediately after 9S told me to destroy the machines I began to attack the otherwise peaceful beings. This can be seen as a questionable moral decision by myself at the time, but because violence is often how games are progressed why should these machines be doing anything but imitating human speech, as 9S says? As I attack the machines immediately after 9S suggests, the mournful female vocal line starts to represent not the humans I was ‘protecting’ but the machines that I was murdering; this sequence made me feel similar to Cheng’s decision to detonate the nuclear bomb in his playthrough of *Fallout 3*.<sup>46</sup>

<sup>46</sup> Cheng, ‘Sound Play’, p. 46.

### ***Amusement Park and the Mad Songstress***

There are variations in the handling of voices in *NieR* however, depending on what is necessary for the environment to display. After the player has completed the desert environment, they are sent to search the amusement park by the android Resistance Camp. The amusement park is as abandoned and ancient as the apartment complex, but this once human establishment has been overrun by peaceful machines; the player is not attacked by any machine in the amusement park unless they attack first, in which all the machines will attack the player en masse. The machines are dressed up as clowns and acrobats, with balloons, confetti, and fireworks being used in ways that resemble a working amusement park. As the player explores the environment, the machines march around the park in clown costumes whilst throwing confetti. Without closer inspection of the ruined park, the somewhat menacing clown costumed machines, shown in Figure 9, appear to have kept this environment functional.



*Figure 9. NieR: Automata: The costume-clad machines of the amusement park environment that appear to be maintaining the environment's original functionality. Screenshot by Author.*

The soundscape of the of the park incorporates another solo female voice in the non-diegetic space, however the song, although still mournful for the loss of humanity, retains the joy of an amusement park. This eery soundscapes incorporates a language made by the performer, Emi Evans, known as chaos language (Nouveau-FR as described in its use in the final cue *Weight of the World*). In a tweet by Evans in 2018, she explains part of the creation process of chaos languages: 'Interesting NieR fact!

Weight of the World Emi version is full of real words from the critically endangered Chamicuro language. In honour of all dying out languages I hoped that at least it'll be preserved forever in this song.<sup>47</sup> The mixture of languages used to create chaos language incorporates both endangered and modern languages. This highlights the loss of human civilisations in *NieR*, but also its replacement by android and machine civilisations as this new language represents the new inhabitants of earth.

This environment differs from the desert, however. Where the player's position in the desert triggered quiet, medium, dynamic, and vocalised cues depending on their progress through the environment, these cues of music are attached to the amusement park itself. When the player first enters the environment, the soundscape plays the quiet amusement park cue. However, after the player has crossed the threshold of the park, the dynamic cue plays with the sung female voice. This cue, instead of focusing on the player's whereabouts, is triggered by the centre of the park, a Germanic-Disneyland style town environment. If the player wanders to the outskirts of this town, or down the side streets, the voice is turned off and the dynamic cue is reduced to a medium cue. Clearly, the non-diegetic music is attached to the park.

The game provides no answer to this disconnection between the player and soundscape in the initial town area. The reaction of the player-character 9S to this environment, and the peaceful machines, is initially no different to his reaction in the desert. The player's reason for traversing the amusement park is to track a signal from other androids. In order to get to this signal, the player must destroy an 'Amusement Park Goliath Tank'. As the tank appears, however, various machine voices begin to show humanistic behaviours by saying 'Let's dance! Let's dance!', 'Let's play! Let's play!', and most notably 'Let's sing! Let's sing!', as the decorated tank fires confetti from its cannon. Even after a high frequency, childlike voice exclaims 'Let's play! Play! Play!' 9S reacts no differently than in the desert. 9S exclaims that the machines are heavily armed and manipulates the player into attacking by saying 'We'll regret it later if we let them escape, so let's take 'em out!' As the player begins to attack the tank, the peaceful machines turn from firing colourful balloons into the air to releasing energy orbs that damage the player-character. As the tank is destroyed, and the player nears their goal of finding the android signals, the player-characters fall through the roof of an old theatre and trigger the area's final boss. In this scenario, the reason for the disconnection of soundscape to player is revealed.

The theatre is inhabited by a large machine known, in the English localisation, as Simone. Without using language, the theatre's curtains are drawn as Simone turns to meet 2B and 9S and projects

---

<sup>47</sup> @emirevans, 'Emi Evans', *Twitter*, 2 April 2018  
<<https://twitter.com/emirevans/status/980820204274855936?s=20>>.

two vocalisations, thus beginning combat. The theatre and Simone have the upper hand against the androids in both physical combat and sonicity. As the player-characters enter this arena, they are taken off-guard to find the signal was being sent by androids that appear lifeless, stripped of their human faces, and bound to Simone's theatrical outfit. Even 9S is uneasy in this situation and does not speak his usual lines of 'taking out' the machine, and instead wavers in his confidence by speaking, in a panicked voice, 'Our records don't say anything about a machine like this!'



Figure 10. *NieR: Automata*: Simone's plea to the game world to become 'more beautiful'. Screenshot by Author.

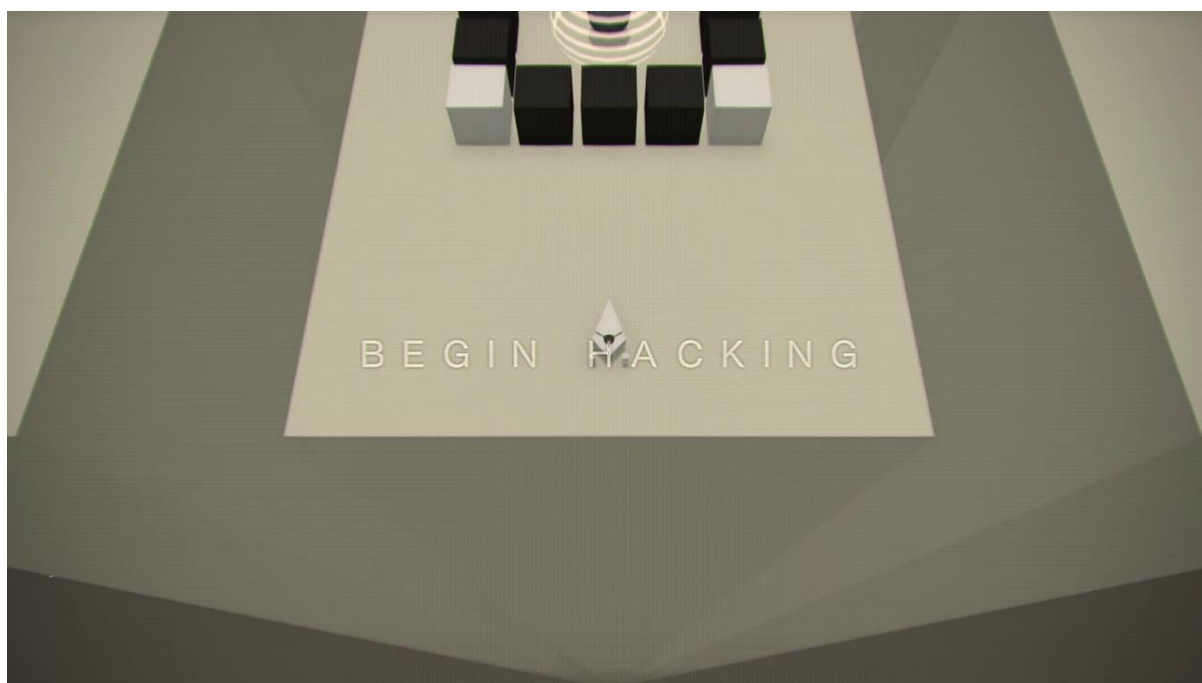
Simone's theatrical character controls both her own voices and the voices of the soundscape. Simone's initial vocalisations, two phonetic 'ahs' on a single note, turn into a mechanised scream that fades out and becomes part of the chorus of female voices who sing in the non-diegetic space. The speed in which the player-characters fall into the theatre and combat begins gives the upper hand to Simone and her voices. After Simone's diegetic scream, the combat cue 'A Beautiful Song' signifies Simone's desperation in her attack, and the confusion and fear of a defending 2B and 9S, through the inclusion of rhythmically fast violin lines and timpani accompaniment.

Unlike the desert, or the amusement park, the theatre has control over multiple voices. 'A Beautiful Song' incorporates the voices of both female soloists of *NieR*, J'Nique Nicole (desert environment vocalist) and Emi Evans, alongside the chorus. The non-diegetic voices perform in chaos language and are fully controlled by Simone. When Simone takes an interval from combat, replacing herself



on stage, the non-diegetic cue stops. In this interval, which occurs after the player has dealt a certain amount of damage to Simone, Simone begins her demand, or plea, to be beautiful.

Simone's diegetic vocal plea in English, seen in Figure 10, removes the player agency further from the narrative of the world. Simone's words 'I MUST BE BEAUTIFUL' are more mechanicalized than her initial 'ah' vocalisations at the beginning of the fight, signalling her fatigue and frustrations. This time, the vocalisations do not become a scream as Simone screams the word 'beautiful' at a higher frequency than before, possibly inducing terror (or at least discomfort) with players. It is not only Simone's desperation and attacks that confront the player-characters. Simone's siren song, her screams, effect the androids by attempting to hack into their systems. Thankfully for the player, they can counteract this hacking by playing a retro, bullet-hell-style minigame, as shown in Figure 11.



*Figure 11. NieR: Automata: The retro bullet-hell style minigame presented to the player as they are 'hacked' by Simone's siren song. Screenshot by Author.*

The soundscape is also distinct in this hacking state. The shoot 'em up minigame adapts the musical cue to a chip-tune variation with no voices. The disruption Simone's voice creates to the visuals, audio, and gameplay through her song shows her desperation to become beautiful.

The tragedy of Simone is the reason for her siren song. Simone does not attack the player because of the machine's orders to destroy humans and androids. Simone attacks 2B and 9S because they are androids. As 2B strikes a final blow to Simone one of her memories, and reasoning to become beautiful, is shown to the player; see Figure 12.



Figure 12. *NieR: Automata*: Simone's reason for attempting to become beautiful, the machine named Jean-Paul. Screenshot by Author.

Simone's need to become beautiful is her adoption of the human emotion love. *NieR* is a game of multiple routes and playthroughs. In the second playthrough, the player still fights Simone and destroys her. However, instead of just a glimpse of Simone's reason to be beautiful, her backstory is uncovered. Simone is in fact a cannibal and the way she becomes 'beautiful' is by consuming the bodies and oil of both machines and androids. Simone uses the corpses of the machines and androids in order to accessorise herself, but this need to become beautiful is not because of her own agency. Simone's tragedy is her need to become beautiful in order to catch the eye of the machine Jean-Paul, which never happens.

The soundscape of *NieR* clearly adapts to the player's movements throughout different environments, and the quiet, medium, and dynamic variations of each cue focus on different areas within each location. When voices are added to the cue, the music becomes an accompaniment as the voice acts a countermelody rather than acting in unison with other instrumentation. However, when the player begins to lose their agency in the game world, as machines become sentient, they lose their grasp on the soundscape and associations with the voice. The game continues to focus its voices as identifiers of the machines after the desert and amusement park environments. The voices' implementation in the soundscape, and the use of chaos language, over English and Japanese, exhibits the game's narrative focus on the machine and android civilisations, rather than humans.

The forest kingdom is the last environment in the game, shown through its command of multiple voices and the machine's ability to use English dialogue to communicate between themselves. The forest kingdom cue uses multiple choral voices, as opposed to the desert's solo female voice, with higher, nasal, pitching alongside Tubular Bells, a Shakuhachi (Japanese flute), and hand percussion for the dense forest; Keiichi Okabe states that the choir sounds in the cue are not a recorded choir but are separate recorded performances from the development team.<sup>48</sup> Voices confront the player with moral dilemmas throughout *NieR*. The game tricks the player into believing that their actions, destroying machines, are righteous but eventually the soundscape, and its voices, abandon the aggressive behaviours of the player.

The soundscape that worldbuilds environments incorporates voices that focus on the people and cultures of the game world. Voices do not identify the player's sense of self but works towards orientating the player. Without environmental worldbuilding as separate to characterisations and identifications, the game world would not be able to effectively create a nuanced and organic world; the player would not be faced with difficult questions regarding their approach to people and cultures as the diversity of these NPCs would not have space to be identified. Environmental worldbuilding is thus separate from identifying the character in order to create a coherent, believable, and organic game world. However, because of environmental worldbuilding's aim to create an organic world that is understandable from the player's experiences, it often uses stereotypical associations of people and cultures. Stereotyping is not contained to one type of worldbuilding, however, and this problematic use of sound and voice as an identifier is further discussed in Chapter five.

---

<sup>48</sup> Marcelus Castle Rain, 'Keiichi Okabe & Takahisa Taura Interview (NieR Automata - Pax East 2017)', *YouTube*, 16 April 2017 <[https://youtu.be/\\_CmYrYhtvGs?t=1m53s](https://youtu.be/_CmYrYhtvGs?t=1m53s)> [accessed 4 July 2018].



## Chapter Five: Characterisations and Identifications

Chapter four outlined the worldbuilding process of environments and cultures through diverse types of voices and their technological implementations into the game world. As the player engages with these environments and cultures, they may also communicate with NPCs that populate the world through their playable-character or avatar. This chapter continues to offer readings of voices as worldbuilding tools in RPGs but instead focuses on player identifications with specific characters. The characters involved in forming the player's sense of game-world identity can range from solely the player-character to all characters in the game, and the level of engagement with these characters can also differ depending on several factors. Firstly, the technical decisions made by the developers can change the amount of communication that can occur between player and characters, ranging from nothing to a freedom of interaction. Secondly, it is common in RPGs for NPCs to react differently to the player-character depending on decisions the player makes throughout the game. For example, if the player unsheathes a weapon in a town area then it is likely that NPCs will vocalise scared or aggressive behaviours towards the player-character. Finally, characters may react to the player-character without any direct interaction, making comments towards them in environments as the player moves through an NPC's field of view.

### The Process of Identification

Voice and communication are strong identifiers of characters as each voice conducts a unique vocal grain. The inclusion of player-character voices can feedback information to the player, including gameplay signifiers such as low-health, and can forge an identification between the player and their playable-character. However, the player's identification with their playable-character, and their characteristics, is not an immediate positive or negative reaction but a process of engagement. The process of identification reacts to the player's involvement with various characterisations of a character's model, including their environmental worldbuilders outlined in Chapter four (accent, class, culture). Where the player can identify with some aspects of self in a character, including Frühholz, Trost, and Grandjean's 'complex emotions' outlined in Chapter three (i.e. love, pride, perseverance), they begin to engage further with the game world. Experiencing identifiable features in a game may occur through several factors: the player's witnessing of character actions during cut-scenes (where the player may have lost control of their character); the fashion and general aesthetics of a character including gender, clothing, and makeup; character backstories and

subsequent relationships and personalities; and a character's personal development over a game's narrative. Voice and vocalisations may develop identifications within this process as they worldbuild character lives. Not every characteristic of a character, however, will necessarily aid the player's identification process. Each player may identify with all, some, or none of a character's characteristics.

Anahid Kassabian considers the identification process as an opening of multiple pathways to identification, intertwined with musical meaning-making for an audience.<sup>1</sup> The issue with branching pathways to identification theoretically locks the player on one particular pathway and forces the other ways to identification closed. I would therefore suggest that the identification process exists on a scale. Neither does identification seem random in the characteristics the player will engage with, nor is it as simple as the player identifying with all aspects of a character, or a single pathway. It is likely that the player identifies with different aspects of a character during various moments of gameplay. As character development is common in most mediums and literature, the player may identify more and less with characters depending on their actions. By this standard, identification can be fluid because the player can identify with certain characteristics in one moment of the game but less so in another. The fluidity of the process encourages this idea of identification on a scale as opposed to pathways. If the player engages with some characteristics, they may find themselves more involved with other aspects of a character as the game continues.

In 1996, Simon Frith examined the discourse of music and identity and its creation within popular music. Frith's research considered the concept and creation of identity in relation to the postmodern, positioning this as a dynamic process which can be formed by the self:

Postmodernism, that is to say, is taken to describe a 'crisis' of signification systems: how can we now tell the difference between the 'real' and the 'simulated'? The postmodern problem is the threat to our sense of place - hence the mapping metaphors, the use of terms like depth and surface. What is underplayed in such discussions is the problem of process - not the positioning of the subject as such, but our experience of the movement between positions.<sup>2</sup>

There are certain parallels between Frith's criticism of postmodernism, its threat to our sense of place, and the goal of escapism for some players. As players may grow attached to characters, they may escape to an identity separate from reality. However, Daniel Muriel and Garry Crawford

---

<sup>1</sup> Anahid Kassabian, *Hearing Film* (London: Routledge, 2001), pp. 2–8.

<sup>2</sup> Simon Frith, 'Music and Identity', in *Questions of Cultural Identity*, ed. by Stuart Hall and Paul du Gay (Los Angeles: SAGE Publications, 1996), p. 109.

challenge the isolated consideration of escapism and suggest that video games also achieve a space for learning empathetic behaviours:

Escapism is indeed an important part of why people play video games, but it is not the only cause or consequence of video gaming. We would suggest that, far from escaping from reality, video games can also connect us with (other aspects of) reality in surprising and unexpected ways. For instance, video games can help us put ourselves in the shoes of others and provide new experiences. Video games work then as mediation devices between players and reality, which could potentially encourage players to empathize with different, even extreme, situations – such as, civilians in a context of war in *This War of Mine* (11 bit studios, 2014), parents of a boy with cancer in *That Dragon, Cancer* (Numinous Games, 2016), migrants trying to pass through a border post in *Papers, Please* (Pope, 2013), or groups living through a scenario of social catastrophe in *The Walking Dead* (Telltale Games, 2012).<sup>3</sup>

Muriel and Crawford consider empathetic responses to in-game characters as a learning experience and emotional training towards some real-world scenarios. On the other hand, if the player can reflect their personal, real-world experiences with an in-game character, this emotional training may intensify player relationships with the game world. Escapism is not necessarily impossible here, but the inclusion of real-world possibilities within a game may change how the player engages with a character, dealing with an issue as though they were in the real-world by using their own experiences. These determinations of actions, and variations of escapism, can affect identifications. The difference between total escapism and using real-world experiences, to react to a game world, solidifies identification as a scaled process; the player may move between these two reactions to game-world issues throughout a single game, or gameplay session.

## Reflections of Self

Kaja Silverman's *The Acoustic Mirror* expanded feminist theories of voice in cinema as a signifier of an audience's personal identity:

[...] it is given the imaginary power to place not only sounds but meaning in the here and now. In other words, it is understood as closing the gap between signifier and signified. Even more important, at least within the context of this discussion, western metaphysics has

---

<sup>3</sup> Daniel Muriel and Garry Crawford, *Video Games As Culture: Considering the Role and Importance of Video Games in Contemporary Society* (London: Routledge, 2018), p. 115.

fostered the illusion that speech is able to express the speaker's inner essence, that it is 'part' of him or her.<sup>4</sup>

This idea of an inner essence encourages empathetic responses between the player and character. Witnessing one's sense of self in a character may induce empathetic emotions and change the player's gameplay motives when regarding that character, as outlined.

Before we can consider processes of reflecting self in the RPG, however, the player's movable in-game character model must be defined, as there are two variations which can directly affect the identification process of the player; the player-character and the player avatar. Crystle Martin explains that the RPG space offers 'the player autonomy in choosing a character's identity, making decisions on what activities and goals to pursue, and making volitional choices.'<sup>5</sup> However, contrary to Martin's explanation, the character's identity and the number of decisions the player can make on behalf of their character, often depends on the type of player-model. When provided with an RPG avatar, the player is given a blank character model that may lack their own back stories, personality, moral codes, and physical attributes. Players can usually greatly customise the avatars to reflect attributes that the player wishes. This customisation occurs in RPGs such as *Divinity: Original Sin*;<sup>6</sup> *The Elder Scrolls V: Skyrim*;<sup>7</sup> *Fallout 4*;<sup>8</sup> *Xenoblade Chronicles X*;<sup>9</sup> *Pillars of Eternity*;<sup>10</sup> and *Disco Elysium*.<sup>11</sup> Avatars provide players with a greater scope of decision making which can include moral choices that can change the narrative ending of certain games i.e., neutral, good, evil, or more complex endings: such as *Bastion*. Avatars, however, are often not outfitted with complete voice acting. The player may be able to choose a character's vocal style that changes the grain of voice lines or vocalisations, such as in *Divinity: Original Sin* which allows the player to choose between voice types such as 'Wizard Voice', 'Rogue Voice', or 'Warrior Voice'. On the other hand, the avatar can also be completely silent, which is often a side effect of having complete customisation. *Xenoblade Chronicles X* supports the silent avatar, allowing full customisation over appearance and gender but only providing the player with a few silent, text-box dialogue choices throughout. With

---

<sup>4</sup> Kaja Silverman, *The Acoustic Mirror: The Female Voice in Psychoanalysis and Cinema* (Bloomington: Indiana University Press, 1988), p. 43.

<sup>5</sup> Crystle Martin, 'Video Games, Identity, and the Constellation of Information', *Bulletin of Science, Technology & Society*, 32.5 (2012), 384–92 (p. 384).

<sup>6</sup> Larian Studios, *Divinity: Original Sin*, Tokyo: Bandai Namco Entertainment (Microsoft Windows, PlayStation 4, Xbox One, Linux, OS X, 2014).

<sup>7</sup> Bethesda Game Studios, *The Elder Scrolls V: Skyrim, Rockville: Bethesda Softworks* (Microsoft Windows, PlayStation 3, Xbox 360, PlayStation 4, Xbox One, Nintendo Switch, 2011).

<sup>8</sup> Bethesda Game Studios, *Fallout 4, Rockville: Bethesda Softworks* (Microsoft Windows, PlayStation 4, Xbox One, 2015).

<sup>9</sup> Monolith Soft, *Xenoblade Chronicles X, Kyoto: Nintendo* (Wii U, 2015).

<sup>10</sup> Obsidian Entertainment, *Pillars of Eternity, Stockholm: Paradox Interactive & Austin: Versus Evil* (Microsoft Windows, Linux, OS X, PlayStation 4, Xbox One, Nintendo Switch, 2015).

<sup>11</sup> ZA/UM, *Disco Elysium, Lutsk: ZA/UM* (Microsoft Windows, PlayStation 4, Xbox One, 2019).

the player avatar, the lack of voice allows the player to include their personal identity, whether that is through the actual choice of vocal style (*The Sims 4*<sup>12</sup>) or being able to imprint a personal identity over a silent avatar.

The player-character often has fixed attributes, accompanied by backstories and associations within the game, with the player entering an ongoing narrative and life. The nature of RPGs means that the player will have some control over certain equipment, attributes, outfits, and moral decisions but overall the character has their own personality and goals set by the developers; this occurs in RPGs such as *Divinity: Original Sin II*; *NieR: Automata*; *Paper Mario: The Thousand-Year Door*;<sup>13</sup> *Final Fantasy XII: The Zodiac Age*;<sup>14</sup> and *Dragon Quest XI: Echoes of an Elusive Age*.<sup>15</sup> Because player-characters have their own agency, outside of the player's control of their movement, they are more likely to be fully voice acted. This inclusion of complete voice acting means that the player-character has vocalised characterisations that the player cannot change. Depending on the genre of the game, or its age and technical constraints, the player-character can also be a silent protagonist, as in *Bastion* or *The Legend of Zelda: Majora's Mask*.<sup>16</sup> The silent protagonist does not represent an avatar, and the player has no more control over physical appearances or overarching goals than a voiced player-character. Silent protagonists remain player-characters because of their set attributes and goals. The silent player-character is never assumed in the game's canon to be 'mute', and other characters and NPCs act as though the protagonist is not silent, but the player cannot hear their character's vocal responses.

Alongside the difference in character type, the development style of the RPG can change the type of character the player receives. William Gibbons' book chapter, 'Music, Genre, and Nationality in the Postmillennial Fantasy Role-Playing Game', outlines the differences between creating a sense of self in a game world within the Western RPG and Japanese RPG (JRPG):

Avatars, common in Western RPGs, are the products of players' imaginations. They are typically customizable in appearance and skills, and often have very little (if any) backstory provided in the game itself, allowing for maximum customization. Many players, in fact, choose to make their avatars look as much like their own appearance as possible, to project themselves into the game. [...] JRPGs seldom, if ever, allow for the creation of avatars.

---

<sup>12</sup> Maxis, *The Sims 4, Redwood City: Electronic Arts* (Microsoft Windows, macOS, PlayStation 4, Xbox One, 2014).

<sup>13</sup> Intelligent Systems, *Paper Mario: The Thousand-Year Door, Kyoto: Nintendo* (GameCube, 2004).

<sup>14</sup> Square Enix, *Final Fantasy XII: The Zodiac Age, Tokyo: Square Enix* (PlayStation 4, Microsoft Windows, Xbox One, Nintendo Switch, 2017).

<sup>15</sup> Square Enix, *Dragon Quest XI: Echoes of an Elusive Age, Tokyo: Square Enix* (PlayStation 4, Microsoft Windows, Nintendo Switch, 2018).

<sup>16</sup> Nintendo EAD, *The Legend of Zelda: Majora's Mask, Kyoto: Nintendo* (Nintendo 64, GameCube, 2000).

Instead, protagonists are ‘agents’ - characters with their own clear identities and motivations separate from those of the player [...] The joy of inhabiting an avatar is projecting your own subjectivity into a virtual space; the pleasure of an agent is experiencing a narrative through the subjectivity of another.<sup>17</sup>

The difference between Western RPGs and JRPGs challenges the usual conception of an identification process. As Gibbons suggests, it is more likely that Western RPGs provide avatars, and player customisation, whereas the JRPG offers pre-set player-characters. The variance in development between both RPG styles creates large scale and subtle gameplay changes: the game’s length; freedom of customisation; reliance on experience gained through combat; levelling up systems; and the size of the party and the number of characters players can control. Although the genre remains an RPG, it is common for players to prefer one RPG style over another because of the vast differences in gameplay. JRPGs tend to be longer games, with story lengths accumulating between sixty and one-hundred hours of gameplay, whereas the Western RPG usually does not surpass forty to sixty hours. Most importantly it is the difference between character types, party dynamics, and thus the use of voices that define the two styles. Gibbons’ analysis of the variations between Western RPG and JRPG player-characters, and avatars, outlines the challenges of researching the identification process within RPGs. To understand the variances in identification between RPGs is not the purpose of this chapter, but it is worth noting that the difference in style, alongside the differences between characters and avatars, can make a difference to player empathy, identification, and enjoyment of the video game.

Gibbons uses the term ‘agents’ in his analysis, where I prefer the term ‘player-character’. ‘Agents’ suggest that the agency of a character lies with the player, rather than themselves, and they are the agents of the player’s goals. However, with pre-set characters, it is more likely that the player is enacting the wishes of the character and the game, helping the character level-up and succeed in combat to complete their story. Thus, I prefer the player-character as it suggests an equal partnership where the character is receiving aid from the player. Due to this partnership, it may be more difficult for the player to personally identify with set attributes and goals, unlike the avatar, therefore the developers must build empathic behaviours within the identity of the player-character.

---

<sup>17</sup> William Gibbons, ‘Music, Genre, and Nationality in the Postmillennial Fantasy Role-Playing Game’, in *The Routledge Companion to Screen Music and Sound*, ed. by Miguel Mera, Ronald Sadoff, and Ben Winters (Oxon: Routledge, 2017), p. 421 <<https://doi.org/10.4324/9781315681047>>.

Zach Waggoner makes the distinction between agent and avatar by eliminating certain video game genres as they do not provide the player with true customisable avatars.<sup>18</sup> Waggoner's project focuses on the conscious and subconscious identifications between the player and their playable avatar in role playing games:

[...] even as the participants' 'separateness' from their avatars disappeared for some diegetic decisions, at other times the distinctions between real-world and virtual world identities were clearly demarcated. However, exactly when these distinctions might occur was unpredictable and/or unconscious.<sup>19</sup>

Waggoner discusses (as do Muriel and Crawford) the fluidity between players identifying as a character to the player using real-world sensibilities, moving between empathy with an avatar to becoming the avatar. Regarding the identification process in *WoW*, Ragnhild Tronstad breaks identification into two variations, 'Empathic Identity' and 'Sameness Identity':

There are different ways of understanding "identification." On the one hand, identification with one's character may be understood as the player entering a state where he or she has an experience of "being" the character. On the other hand, identification may be understood as experiencing what the character experiences, but without the feeling of being identical to it - that is, with a consciousness of the character as an entity other than ourselves, but with which we can identify.<sup>20</sup>

Both Waggoner and Tronstad speak of the avatar, as the playable-character in *WoW* has the freedom of customisation and lack of voices attributed to the avatar. Waggoner discusses a fluidity in identification, whilst Tronstad considers the differentiation between the player 'being' or 'experiencing' their avatar in *WoW*. The variations between the player-character and avatar, however, changes the strength of identification. I would thus affix Tronstad's 'being' and 'experiencing' to either the player-character or avatar. I suggest that an empathic identification is more likely to occur with the player-character as the player may not reflect the character's attributes. The 'being', and sameness identity, may be more likely to be achieved with an avatar because the player has greater customisation.

The issue of the player-character is that their, often, unchanging features may stop the player from seeing a reflection of themselves on-screen. Michael D. Hanus and Ted M. Dickinson explored the

---

<sup>18</sup> Zach Waggoner, *My Avatar, My Self: Identity in Video Role-Playing Games* (Jefferson: McFarland & Company, Incorporated Publishers, 2009), p. 12.

<sup>19</sup> Waggoner, 'My Avatar, My Self', p. 160.

<sup>20</sup> Ragnhild Tronstad, 'Character Identification in World of Warcraft: The Relationship between Capacity and Appearance', in Corneliussen and Rettberg, *Digital Culture, Play, and Identity*, p. 251.

resistance to female protagonists in video games, examining the post-crash period of the video game market in the 1980s. Hanus and Dickinson identify that the shift towards targeting video game advertisement towards males subsequently made publishers believe that the demographic of video games (young male players) would not respond to a female protagonist.<sup>21</sup> To ascertain whether this was the case, an empirical study was conducted to measure player and non-player reactions to a video game protagonist's sex. The study provided subjects with text-descriptions of different video games whilst modifying the protagonists' original sex (note the games were not played by the participants).<sup>22</sup> Alongside this, Hanus and Dickson also measured a subject's overall video game play time in a week, anticipated enjoyment of a game, their purchase intention of that game, their similarity to the character shown, and the desirability of a character through an identification scale. The conclusions of the study showed that players do not need to identify with all characteristics of a character, including sex, but can choose what they want to identify with:

Game players may be aware of the unique process of video game identification wherein the line between the player and the character blends. Players can select aspects of the character that they wish to incorporate into their own self-concept, and as a result can attempt new behaviors or feel differently than they could experience outside of the game.<sup>23</sup>

Therefore, the shift to appease the resistance against female protagonists in video games may have been ill-informed. The player's experience of a protagonist of the opposite sex is unlikely to damage identification as they can select other characterisations to identify with. In film, Robynn J. Stilwell explains that 'experiencing a strong identification with a character in the film places us in another's subject position, creating an emotionally empathetic response.'<sup>24</sup> Catherine Haworth also suggests, in her article 'Introduction: Gender, Sexuality, and the Soundtrack', that where sex may be fixed or changeable depending on the type of character offered, sexuality and gender may not be:

Despite the visible masculinisation and heteronormativity of much gamer culture, the player's gender or sexuality is of course not fixed (a feature mirrored in many games that allow a choice of male, female, androgynous, or otherworldly avatars for players to control).

---

<sup>21</sup> Michael D Hanus and Ted M Dickinson, 'The (Faulty) Assumption That Male Players Prefer Male Characters: How Character Desirability and Likability Influence Video Game Purchase Intentions and Enjoyment', *Psychology of Popular Media Culture*, 8.4 (2019), 395–401 (p. 396) <<https://doi.org/10.1037/ppm000191>>.

<sup>22</sup> Hanus and Dickinson, 'The (Faulty) Assumption That Male Players Prefer', p. 399.

<sup>23</sup> Hanus and Dickinson, 'The (Faulty) Assumption That Male Players Prefer', p. 399.

<sup>24</sup> Robynn J. Stilwell, 'Sound and Empathy: Subjectivity, Gender and the Cinematic Soundscape', in Donnelly, *Film Music*, p. 173.



Thus, like other audiovisual texts, video games must ‘work’ for both male and female consumers, regardless of any intended or perceived gendering of the game itself.<sup>25</sup>

In video games, the empathetic response to a character grows during the multitude of gameplay hours as the player supports and moves their character towards their goals. As the player spends time with a character’s voice, either through dialogue or associated song, they become accustomed to the grain of the character’s voice and thus the accompanying emotions. As Collins suggests ‘we mentally mimic the voice in our own body’,<sup>26</sup> as well as physiologically on some occasions. Therefore, emotions in a character voice may be passed to the player as they mentally mimic these vocalised emotions. This empathetic mimicking of the voice may be structured by perceived ‘male’ or ‘female’ aspects of the voice, however. The vocal grain of a character (i.e. high/low, soft/stern) may direct the player to associating a gender to the character’s characteristics. The perception of vocal genders by listeners, outlined in Chapter three through Jarman-Ivens’s ‘third-space’,<sup>27</sup> may cause issues between the characterisations that developers want to build and the player’s association of these vocal grains.

Kassabian makes the distinction between two styles of soundtracks when forming a process of identification between audiences and film. These two distinctions include the ‘composed’ score, which is created specifically for the film, and the ‘compiled’ score, which uses pre-existing songs.<sup>28</sup> *Fallout: New Vegas*<sup>29</sup> (*New Vegas*), is an RPG set in a post-apocalyptic America which provides the player with a Pip Boy, that acts as an interface for the player alongside being a radio transmitter. The Pip Boy broadcasts an array of existing mid-20<sup>th</sup> century songs, such as ‘Orange Colored Sky’ by Nat King Cole and ‘Blue Moon’ by Billie Holiday, in order to situate the created culture of the *Fallout* series in *New Vegas*. The use of Compilation soundtracks in the *Fallout* series, are to highlight the lack of music developed by the inhabitants that exist in the disturbing, and monster-inhabited, postapocalyptic America.<sup>30</sup> When discussing *Fallout 3*, Cheng identifies the excuses made by the in-game NPCs over the lack of new music:

A few NPCs even offer explicit rationalizations for the wasteland’s dearth of new music.

Agatha, for instance, explains that all of her station’s music is recorded because it would be

---

<sup>25</sup> Haworth, ‘Gender, Sexuality, and the Soundtrack’, p. 125.

<sup>26</sup> Collins, ‘Making Gamers Cry’, p. 44.

<sup>27</sup> Jarman-Ivens, ‘Queer Voices’, p. 3.

<sup>28</sup> Kassabian, ‘Hearing Film’, p. 2.

<sup>29</sup> Obsidian Entertainment, *Fallout: New Vegas, Rockville: Bethesda Softworks* (Microsoft Windows, PlayStation 3, Xbox 360, 2010).

<sup>30</sup> Helen Diggle, at the 2017 Ludomusicology conference, examined the idolisation of Elvis Presley by NPCs in *Fallout: New Vegas*. No Elvis Presley songs exist within the game, yet the identities of some NPCs are represented by the name ‘King’ and adopt his movements and phrases, such as ‘uh, huh’, as gospel.

too tiring for her to play live day and night. ‘That way,’ she quips, ‘you can hear the music anytime you want, and this old woman can get some well-needed beauty sleep!’ Three Dog similarly insists that GNR’s programming is repetitive because he hasn’t managed to scavenge more than a few music records in playable condition.<sup>31</sup>

The compiled score is discussed by Kassabian as offering affiliating identifications with music between characters on-screen and the audience.<sup>32</sup> *Guardians of the Galaxy*<sup>33</sup> uses a compilation soundtrack which acts as the personal soundscape of the main character, Peter Quill. Peter Quill’s identity crisis, demonstrated through his listening to popular music, can create audience empathy with the character as they can refer to their own experiences of the music.

### **Player Attributes: Representations of Identity**

Representations of identity can be problematic. The reflection of characteristics such as race and gender are often used in video games and can reflect on real-world issues if handled well. Gender in multimedia is frequently represented through the power dynamic of the male and female sex, and the association of bodiliness, masculinity, femininity, and other sexualised stereotypes. This is influenced by classic Hollywood productions. When actresses performed with their bodies and voices, their acting style needed to be aligned with the expectations of their character. Kristen Pullen refers to this acting style as ‘naturalism’, a notion where an actor forms a physiological relationship with their character in order to form realistic emotions.<sup>34</sup> The erosion of actress agency is discussed by Pullen because the naturalist acting style in performances meant there was a laboured attempt to develop nuanced, consistent, believable characters and personae through their voice, gesture, stance, and actions.<sup>35</sup> Thus issues arose as the feminine, or what was considered as an appropriate body type, was important during classical Hollywood. Pullen highlights particularly the issues surrounding the athlete and actress Esther Williams, who starred in MGM’s ‘aquamusicals’; there was a constant concern over Esther William’s trademark ability to sing and perform at the same time with her athletic, yet feminine body.<sup>36</sup> The classic Hollywood *femme fatale* trope for female actresses also meant that the focus was on the actress’ body when she performed. *Femme fatale*

---

<sup>31</sup> Cheng, ‘Sound Play’, p. 35.

<sup>32</sup> Kassabian, ‘Hearing Film’, p. 3.

<sup>33</sup> James Gunn, *Guardians of the Galaxy, Burbank: Marvel Studios* (Walt Disney Studios Motion Pictures, 2014).

<sup>34</sup> Kristen Pullen, *Like a Natural Woman: Spectacular Female Performance in Classical Hollywood* (New Brunswick, New Jersey, and London: Rutgers University Press, 2014), p. 4.

<sup>35</sup> Pullen, ‘Like a Natural Woman’, p. 4.

<sup>36</sup> Pullen, ‘Like a Natural Woman’, pp. 62–63.

characters performed risqué jazz pieces to highlight their dangerousness as women whilst cinematography focused on an actress' body, as seen in MGM's 'aquamusicals', during their vocal performances. This use of the female body is examined by Catherine Haworth in her analysis of the film *A Woman's Secret*.<sup>37</sup> Haworth recognises that through music, and a vocal performance by the character Susan, Susan undergoes a sexualised and deviant transformation to a *femme fatale* because of her on-screen jazz performance; the cinematography enforces this musical transition by focusing on her body during her vocal performance of stereotypically non-white music.<sup>38</sup>

As vocal performance highlights the body of the female performer, it can also render her mute. Mutism within female characters is something already used in multimedia, opera, and literature, as seen in the character Ophelia's performances in Shakespeare's *Hamlet*. As highlighted in Chapter three, Dunn explains that Ophelia inappropriately sings during situations of social discourse. This act of singing, rather than speaking 'functions as a highly theoretical sign of Ophelia's estrangement from 'normal' social discourse, as well as from her 'normal' self'.<sup>39</sup> Although not muted completely, Ophelia's estrangement from social interactions because of her singing deems her 'mad'. On-screen mutism in video games, however, considers more than solely a gender issue and identification. Player-characters and avatars sometimes lack a voice actor, rendering them 'mute', for reasons such as budget constraints, aesthetic choices, or to allow the player to imprint their own personality onto the character model. When a character is considered mute by the game world, rather than solely a silent protagonist, it often becomes a gendered point. Female madness and mutism in games follows the gamut of mute female characters in film, often rendered silent because of an ordeal or outside influence. The character Garnet in *Final Fantasy IX*<sup>40</sup> sings and hums the game's leitmotif, 'Melodies of Life', throughout the game. Garnet also uses her singing voice as her combative agency as her vocalisations summon allied monsters into combat. However, as heir to the Kingdom of Alexandria, Garnet is unable to protect her kingdom from destruction and thus falls mute due to her loss. This muteness both affects the inclusion of the sung leitmotif in the game, and her summoning powers. The player does not hear Garnet's song again until the game's final scene when Garnet is reunited with the protagonist Zidane, and finds her voice, suggesting 'Melodies of Life' becomes their song. Garnet's vocalised identity is lost to the player until this moment. It is only until she is reunited with the male protagonist that Garnet is freed from her mute form, unable to undo her madness herself. Because the leitmotif of *Final Fantasy IX* is attached to Garnet, and eventually the relationship of

---

<sup>37</sup> Nicholas Ray, *A Woman's Secret* (RKO Radio Pictures, 1949).

<sup>38</sup> Haworth, 'Gender, Sexuality, and the Soundtrack', p. 115.

<sup>39</sup> Dunn and Nancy Jones, *Embodied Voices*, p. 51.

<sup>40</sup> Square, *Final Fantasy IX, Tokyo: Square* (PlayStation, iOS, Android, Microsoft Windows, PlayStation 4, Nintendo Switch, Xbox One, 2000).

Garnet and Zidane, the player may identify with Garnet and Zidane's characterisations whenever they hear the game's leitmotif within and outside of the game.

Identifications with characteristics can change depending on the player's ability to have an input on the aesthetics of their playable-character. Casey Hart recognises that character models can represent the player's vision of self. This vision of self can be created through a myriad of player choices, whether physical attributes, dialogue choices, and moral decisions.<sup>41</sup> The player is not locked into creating an avatar that fully represents their appearance. This means that players can, as Hart explains, blur the lines of identity and self through the medium of video games. Hart's study concludes that individual players experiment with the physical and psychological attributes of a character, rather than outright projections of self:

Specifically, the data informs the discussion concerning how humans use avatars as vehicles by which to express themselves or experiment with alternative self-concepts. The data seems to suggest that it is much more common for individuals to use avatars as facilitators for experimentation than for vehicles for direct projection. Even when compared to a subject's perception of what is ideal, in most cases it does not appear that individuals are uniformly likely to create avatars that represent their own personality characteristics.<sup>42</sup>

However, whether it is more likely that some players experiment with their character models, some players will still project a sense of self onto the player-character. Some player-characters, as seen in *God of War* and *Horizon: Zero Dawn*, cannot be physically changed but the identification with these characters is about following their personal growths and stories. Seeing ones-self on-screen is not only about physical looks but also the reflection of morality; in these games diversity exists in the environment and NPCs of the game world, rather than with a customisable character.

### ***Issues of Characterisation and Representation***

Although a game is arguably fictional to some extent, real-world issues of sexuality, race, class, and appropriation can often be problematic when characterising people. Where bodiliness, sexuality, and gender are considered as reasons to mute female characters, race and class are often used to identify the 'native', 'other', or 'primitive' trope of a game's people. Voices' ability to create audible

---

<sup>41</sup> Casey Hart, 'Getting Into the Game: An Examination of Player Personality Projection in Videogame Avatars', *Game Studies*, 17.2 (2017) <<http://gamestudies.org/1702/articles/hart>>.

<sup>42</sup> Hart, 'Getting Into the Game'.

characteristics, outlined by Meizel and ‘the black voice’ in Chapter three, can aid these problematic and systemic audio and visual characteristics as cultural identifiers.

*Horizon: Zero Dawn* uses the Native American terminology, ‘Brave’, to represent the warriors of the in-game people, the Nora. Throughout the game, the player-character is frequently told by other civilisations that the Nora are a primitive and backwards people. This appropriation of a Native American word, and its use in referring to a primitive people, is problematic especially because the developers (Guerrilla Games) are primarily European; part of the issue here is the balance of power between European nations and Native Americans in our history. *Horizon: Zero Dawn* accentuates the differences between the ‘primitive’ Nora people and the more advanced Carja civilisation, who aid the player-character on her journey. As seen in Chapter four, the difference between the two civilisations is shown through religious performances. The dedicated religious Sun Priests can always be heard performing a version of ‘Song to the Sun’ when the player visits Meridian, whereas ‘Nora u Norawea’ is limited to an in-game festival and cannot be witnessed once the player has advanced the story. In an interview, the composers Joris de Man and The Flight spoke about scoring for the environment and people by performing contemporary instruments in a style that represents musical illiteracy:

We definitely tied music to certain locations; for example, Lucas would remind us not to use certain types of sound, such as avoiding metallic sounds in the Nora region, as they tend to use wood and animal skin, but I also tried to underpin the experience and the breadth of the landscape – if it was a snowy area, I might choose some filtered noise-based pads to give it a wintery feel, for example.

We also had different tracks for night-time; at night the mechanical creatures are more dominant, and so the musical textures would be more electronical, with the organic elements more pushed to the back.<sup>43</sup>

The appropriation of cultures is rife within the musical and aesthetic choices within video games. Ivănescu’s discussion of *Mafia III*<sup>44</sup> highlights the handling of racial issues in some video games, referring to the use of nostalgic material, such as *film noir* and gangster film and television, for inspiration. In order to study the handling of race in video games, Ivănescu recognises the differences between the radio collection of music, chosen to accompany the black player-characters,

---

<sup>43</sup> Campbell Simpson, ‘We Talked To The People That Made Horizon Zero Dawn Sound So Good | Kotaku Australia’, 20 April 2017.

<sup>44</sup> Hangar 13, *Mafia III, Novato: 2K Games* (Microsoft Windows, PlayStation 4, Xbox One, macOS, 2016).

in *Mafia III* and *Grand Theft Auto: San Andreas*<sup>45</sup> (*San Andreas*). In both games, and usually in most *Grand Theft Auto* games, the player's car has access to music through an in-game radio. In *San Andreas*, the radio stations curated for the player include genre-specific stations such as hip-hop and rap, causing problematic identifications between players and their character CJ:

[...] white middle-class teenagers may assume that CJ is more likely to listen to rap and hip-hop based on similar characters they know from other media. *San Andreas* is again contradictory here: music becomes an aspect of players' identification with a character and understanding of an aspect of American culture, represented here in an exaggerated form; but if the players are not themselves black, is this an exercise in empathy or, as some scholars have argued, is this a problematic embodiment of the other that borders on 'high-tech blackface'?<sup>46</sup>

The argument Ivănescu brings forward here is the player's ability to empathise with a character whose race, and racial experiences, does not correspond to their own. As we have seen through Hanus and Dickinson's study, players can identify with varying characteristics of their playable-character. Therefore, it is not necessarily likely that all players appropriate another race's culture through identifying with a character. Unlike *San Andreas*, however, *Mafia III* focuses on using appropriated music from the 1960s as part of the radio's repertoire, rather than focusing on specific genres. Where players may assume *San Andreas*' CJ should listen to hip-hop or rap because of his race, *Mafia III* provides a radio that represents a developing character in the 1960s. It could therefore be considered that the identification process with pre-made player-characters can be heavily influenced by the stereotyping of characterisations. The player may bring their own stereotypes to the game-world from their own reality, but the developers can lead the player against stereotyping by using appropriate music within the game, as seen with the radio of *Mafia III*.

In the action-adventure game, *Red Dead Redemption 2*, character identities and culture in 1890s America are highlighted through the game's infrequent use of the sung voice. As the player exists in the game world, the identities and cultures of Spanish, American, Caribbean, and Native American people are shown through a diverse cast of characters; although, the protagonists are white, male characters whilst the diversity is saved largely for the supporting character roles. The soundscape especially highlights Native American culture as the player aids them in quests and may travel to a reserve during the final third of the game. During a narrative low point for the player-character, Arthur Morgan, the player must ride back to their outlaw camp after returning from being

---

<sup>45</sup> Rockstar North, *Grand Theft Auto: San Andreas*, New York: Rockstar Games (PlayStation 2, Microsoft Windows, Xbox, Mac OS X, iOS, Android, Windows Phone, Fire OS, Xbox 360, PlayStation 3, 2004).

<sup>46</sup> Ivănescu, 'Popular Music in the Nostalgia Video Game'.

shipwrecked on a (presumably) Caribbean island named Guarma. As the player directs Arthur Morgan's horse back to camp, the cue 'May I? Stand Unshaken', performed by D'Angelo, begins to play in the non-diegetic space. The song includes strong Native American representations in order to highlight Arthur Morgan's journey of repentance during this narrative sequence. The backing vocals that accompany D'Angelo's performance use vocal techniques that Bruno Nettl describes as an identifier of Native American music:

A tense, harsh, raucous manner, and their melodies cascade down a series of terraces or steps, rarely moving upward. To us their music sounds wild. The stressed tones are accented violently, and on the long notes the singers continue the rhythmic pulsations so that the music never comes to a rest. The melodies have large ranges with the singers beginning high, sometimes in a falsetto voice, and descending to a growling depth.<sup>47</sup>

The use of these stressed tones and heavy accents create notable tension in the cue, rather than a celebration of Arthur's return from Guarma. The song acts as a reflective moment for Arthur and the player. Before being shipwrecked on Guarma, the player's quest was to succeed in a bank heist so that the camp would finally have enough money to escape their fate as outlaws. As the narrative leads up to the bank heist, the camp and the player are told on multiple occasions by the camp leader, and Arthur's confidant, Dutch that this is the last job they need to do before escaping to the island paradise of Tahiti. Once a trusted leader, Dutch's failure to lead the player through a successful bank heist causes the death of several camp members and Arthur's closest friend Lenny. This failure is not processed until Arthur returns from Guama, and thus this reflective time is significant.

When the cue is triggered, the interface of the game is removed, and the screen's peripherals are blurred out to focus the player on Arthur in the centre of the screen. The diegetic audio is lowered in the soundscape so that 'May I? Stand Unshaken' is prominent. The drones sung by the backing vocals include vibrato, similarly to how Nettl describes the rhythmic pulsations of Native American song, which stops the music from finding any natural rest. D'Angelo's voice remains in lower registers and dynamics throughout the song, almost speech-singing underneath the backing vocalisations, acting as a narrative voice-over. The semantic meaning attached to the lyrics directly examine the narrative until this moment, singing of standing unshaken 'amidst a crashing world'. 'Crashing world' refers to the slow estrangement between Arthur and Dutch, as both the player and character begin to realise that Dutch is descending into madness. The male backing voices, only sing

---

<sup>47</sup> Bruno Nettl, *Folk Music in the United States: An Introduction*, 3rd edn (Detroit: Wayne State University Press, 1976), p. 58.

the lyrics 'May I? Stand unshaken' before returning to the vocalised drones of phonetic 'ooh' and 'ah' in the verses. This highlights the solidarity between D'Angelo and Arthur at this moment, providing an impetus for reflection on Dutch, the camp, and Arthur's own morality; the player can choose whether to make high or low honour decisions throughout the game, to help or hinder NPCs, which can change some narrative moments and manipulate how NPCs interact with Arthur.

### ***Muteness and Madness: In Transistor and Divinity Original Sin II***

Worldbuilding characterisations in characters that cannot be customised to physically represent the player is often done so through the inclusion of voice. However, where voices are muted for the player-character, the game must use the remaining soundscape to characterise the character.

The silencing of the vocal performer and composer Red is a key narrative identifier in the action-RPG *Transistor*. *Transistor* is set in a futuristic environment which uses voice and adaptive music to build the game's environment and character identities. When the game begins, the futuristic city environment, Cloudbank, has been evacuated and partially destroyed by an ongoing attack by machines called the Process. The Process were controlled by a group of Cloudbank's elite officials, the Camerata, but the group have since lost control of the machines. *Transistor* is a game void of NPCs. The game refers to characters who used to exist in cloudbank before the Process attacked, but if the player finds any physical resemblance of NPCs they are usually not living. The only voices that the player hears throughout the game exist in two forms: pre-existing songs heard in the fantastical gap inside the mind of Red and the dialogue of an unknown man's voice from inside the sword named Transistor. The muteness of Red holds a narrative secret from the player, leaving them unaware of why Red has lost her voice. The inclusion of these two voices, the sung voice and dialogue, unravel the secret that Red holds.



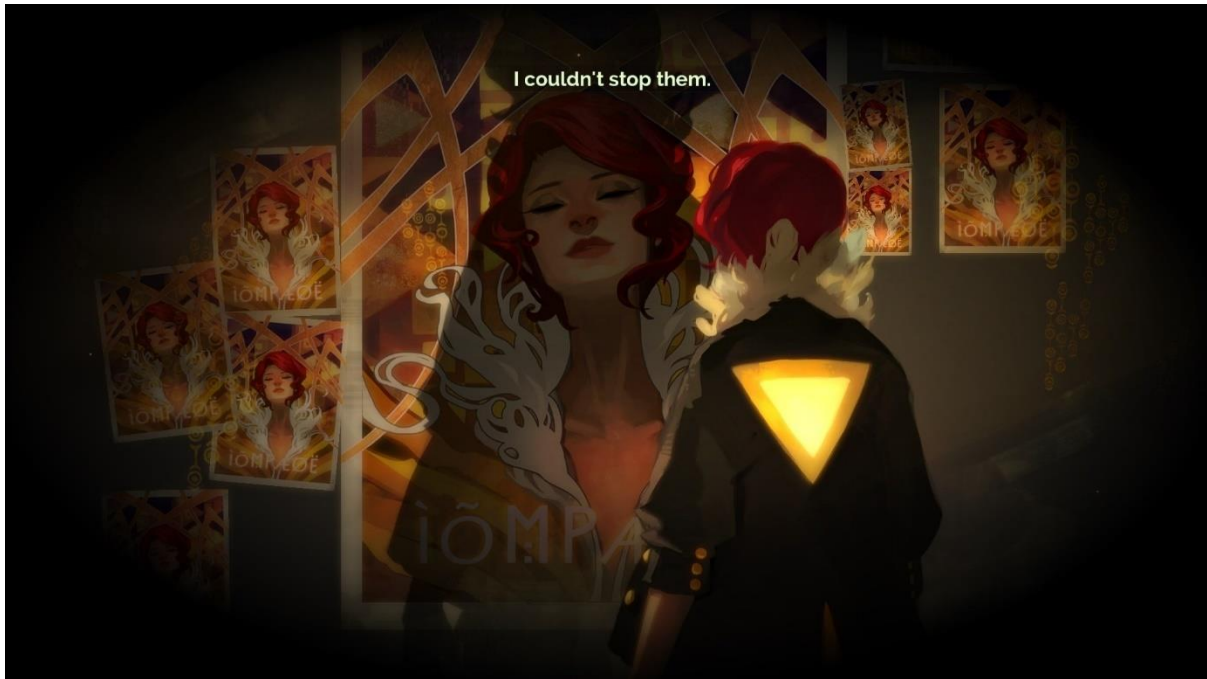


Figure 13. Transistor: Red's lament, Transistor's apology, and the reveal that Red has lost her voice. Screenshot by Author.

The first action of the player triggers dialogue from an unknown male voice - 'Hey Red...We're not going to get away with this are we...' - after which triggers a musical cue. As Red pulls a sword from a male body, the sword continues to speak to Red, stating 'Together again', suggesting Red and the voice have known each other before the beginning of the game. As the player moves through the game, Red's muteness is not immediately apparent because there is no lack of voice due to the sword's dialogue. The player does not hear Red's voice until she is confronted with her performative past through posters in Cloudbank; see Figure 13. As the player enters the area with the posters, the non-diegetic cue, 'The Spine', begins with a riff on the guitar alongside synthesised drums and a synth drone. As Red's performative voice is introduced to the cue, and she stares at her past fame, it becomes uncertain as to whether the voice is only a part of the non-diegetic space or is occurring in Red's mind.

The player is given infinite amount of time to come to terms with Red's muteness. The player can let Red reminisce her past life for as long as they wish. Although Transistor attempts to turn Red away from her image, with dialogue such as 'C'mon just go.', he understands Red's need to silently lament, and leaves Red alone with a final comment of 'Just...Yeah'.'; the player is given impetus to turn Red around, and continue the game, as the shadow of an enemy appears beside Red. Until Red reminisces about her voice, which the player hears in the non-diegetic space, and the sword comments on her stolen voice, it is not revealed that Red is more than a silent protagonist. The

sword's voice speaks on behalf of Red and the player's actions, the transistor uses the man's voice for storytelling, directional, and combat purposes:

Narrative	Directional	Combat
'...I'm so sorry Red.'	'C'mon just go.'	'Nice...!'
'They took your voice.'	'Let's have a look downstairs.'	'Get out of here.'
'Poll's a little out of date.'	'Wait want to check the Channel back there?'	'Gross.'
'Think I know where we are.'	'Unmarked alley... East of the bay...'	'OK get ready.'
'...When I first saw you up on that stage back there it was like...everyone loved you.'	'Sea Monster's really the only choice here.'	'I guess we're done...'

Table 2. *Transistor*: Narrative, directional, and combat dialogue from the *Transistor*.

Table 2 highlights the varying use of the voice as dialogue within *Transistor*. The sword's directions provide aural narrative, directional, and combat related feedback for the player, alongside characterising Red's character throughout the game. The significance of these divisions is their inclusion in what would be an otherwise voiceless world.

Although Red is physically mute, the player can still hear Red hum to herself in personally reflective moments. When the player presses the dedicated hum button, Red takes the sword into an embrace and a spotlight highlights her as though she is performing for the player. Player directed humming is an insight into the past life of Red and her stolen characteristics and identifiers. During the combat of *Transistor*, the player can choose two types of combat style and swap between them; one combat style is real-time action based, whilst the other is a turn-based mode called Turn(), which allows the player to freeze time. During Turn(), enemies are frozen and the player has time to plan a certain number of combat movements before exiting Turn() and allowing Red to action the movements

rapidly. When Turn() is activated, the adaptive music of *Transistor* interacts with the player's decision to freeze combat. When the player enters Turn(), certain instruments are turned off, such as percussion and synths, to provide a space with greater open frequencies; I identify in my Master's thesis that 'Turn() gives a sense of calm for the player in an otherwise stressful moment of the game, pausing the in-game time whilst the player queues a number of movements and attacks.'<sup>48</sup> As other instruments are turned off in this space, Red's humming vocalisation is turned on, almost suggesting that the combat mode is within Red's mind and the space is for herself to relax and think.

Red's physical muteness defines her relationship with the sword. In an otherwise empty game world, the player still has characters to engage and identify with because of this vocalised connection. However, Red's muteness is extreme as she has lost both her ability to identify herself and her agency as a performer. Performative mutism itself is not a complete silencing of a character. The character may be mute in any of Laing's three levels of muteness (physically, psychologically, or selective mutism), as outlined in Chapter three, but does not need to lose their ability to vocalise their identity. In *Transistor*, Red suffers from performative mutism as a side-effect of her stolen voice and thus cannot continue as a performer. The performative player-character, Lohse, in the RPG *Divinity: Original Sin II*, only loses her ability to perform and sing, rather her ability to vocalise entirely.

*Divinity: Original Sin II* provides the player with a choice between selecting the player-character with a pre-set identity, personality, background, and goals, or making the player avatar. The player has access to controlling up to four characters throughout the game and thus can engage with zero to four of the six pre-set player-characters available to them. Even if the player chooses to play one of the six pre-set characters, they do not have to engage with their goals or backstories.<sup>49</sup> If the player engages with the pre-set character Lohse, they will find that she is a well-known performer in the game world. In the character selection screen, Lohse's introduction to the player exclaims that she is an entertainer 'beloved and celebrated by all' across the game world Rivelon. As well as her introduction as a performer, the game introduces a demon parasite that resides within Lohse. The game is always transparent that Lohse's will is frequently taken over by a demon, but it is not until Lohse is offered to play a lute that the player realises that the demon stops Lohse from performing. When the player triggers the action for Lohse to play the lute, the demon forces Lohse to smash the lute on the ground, stopping her from performing, as the demon hates music. This renders Lohse performatively mute as she cannot continue her entertainer role. Like Red, Lohse can still fight against enemies within the game, following a goal to defeat the demon residing inside her. Finally,

---

<sup>48</sup> Smith, 'The Musical Parameters of Immersion and Flow', p. 42.

<sup>49</sup> Although, the player is likely to engage with the characters at least on the first playthrough of the game.

when the player tracks down the demon, named The Doctor, within the final environmental stage of Arx, they are able to initiate combat against him in order to release Lohse from his grasp. When the player defeats The Doctor, Lohse is freed of her performative mutism, unlike Red, and the game rewards the player with Lohse's performance of 'Sing for Me'.

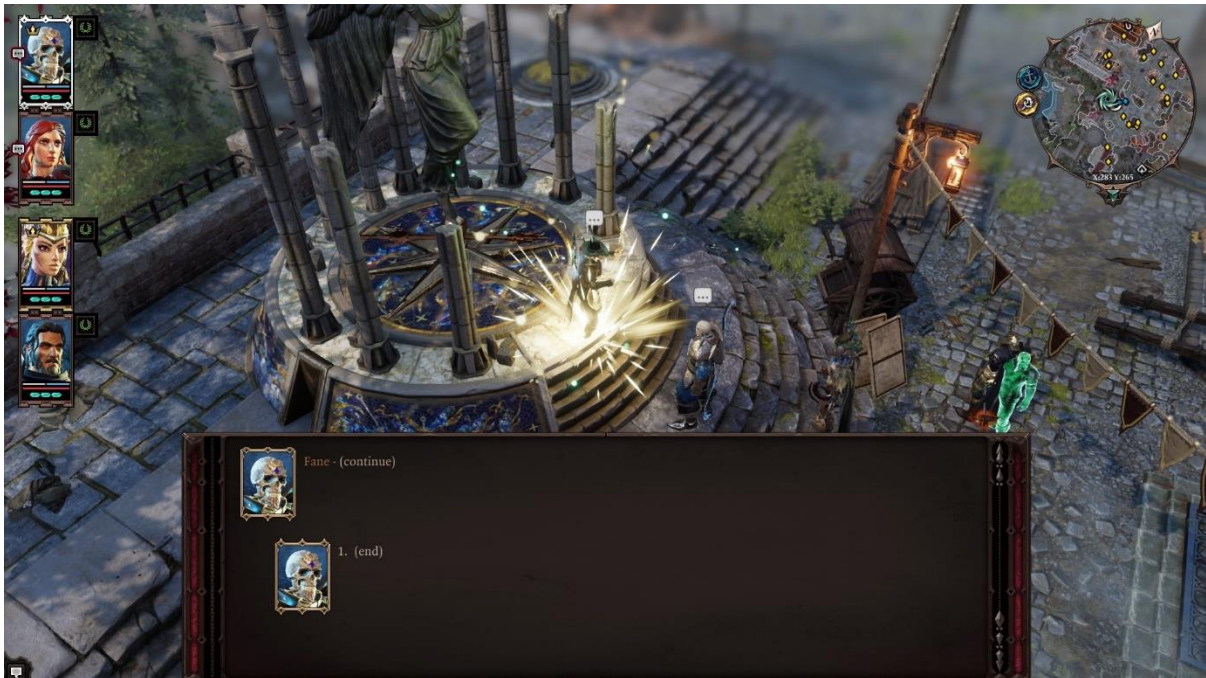


Figure 14. *Divinity: Original Sin II*: Lohse's ethereal lute and staging. When Lohse stamps her foot in time with the music, light explodes around her. Screenshot by Author.

'Sing for Me' is a popular song within the medieval environment, sporting repetitive four-line stanzas within the structure of a popular song, rather than a historically accurate lute and vocal performance. Lohse's lute is ethereal and seemingly made of light, as shown in Figure 14. When the player unlocks the song, they unlock the ability to ask Lohse to perform anywhere outside of combat. Whilst Lohse performs with her ethereal lute, the ground around her is enveloped in light, reacting to her swaying and foot stomping in time with the music, creating an ethereal stage around her. Lohse is the *femme fatale* of her game world. Her deviance against both her performative mutism and medieval aesthetics of the game enforces her characterisations within the story. The use of Lohse's singing voice, and the lyrics of the song, both identify her independence and highlight the player's involvement and possible identification with Lohse as a character. The lyrics for 'Sing for Me' identify a possible relationship between player and character, outlined in Table 3.

<b>Verse Lyrics</b>	<b>Meaning</b>
<i>Sway with me, we'll make them scream.</i>	Call to player, past combat.
<i>Dance with me, we'll make them bleed.</i>	Call to player, past combat.
<i>Sing for me, I'll sing along.</i>	Call to the player to take control of Lohse's actions.
<i>Sing for me, oh sing for me.</i>	Call to the player, the player must try and regain Lohse's voice.

<b>Chorus Lyrics</b>	<b>Meaning</b>
<i>Listen, I'm calling you.</i>	Call to player, breaking the fourth wall.
<i>Listen, you do know me.</i>	Call to player, the relationship built with Lohse over the game.
<i>Listen, swing and roll me.</i>	Player moving the character, the roll of virtual dice for game interaction success and failure.
<i>Listen, I'm calling you.</i>	Call to player, breaking the fourth wall.

Table 3. Divinity: Original Sin II: Lyrical meanings in the verse and chorus' of 'Sing for Me'.

The deviant genre and structure of the popular song style, from the fantasy genre and orchestrated soundscape, is a celebration of Lohse's now independent characterisation. The player can also use the song as their own celebration and may ask Lohse to perform the song after future success in the game. The song's lyrics make comments on certain actions, such as combat; 'sway with me, we'll make them scream, dance with me, we'll make them bleed'. The repeated use of the words 'sing for me' can refer to Lohse's performative mutism and her acknowledgement that her performative

personality has been muted and her actions to retrieve her performative independence have been guided by the player.

Lohse repeats four lute chords of A minor, C, G, and D alongside a melodic riff repeated with each stanza. Example 1 shows the repeated rhythmic pattern within each stanza, highlighting the formulaic nature of the piece and placing it within this popular song style. Although Lohse is seemingly performing by herself, as she stands alone on her stage, the cue includes drones on low string instruments and extra female vocalisations. This inclusion of extra voices and instrumentations places the song in the fantastical gap as it is unclear where these extra sounds are coming from.

8 D Am

A. Sing for me, Oh, sing for me.

10 Am

A. Sway with me, we'll make them scream.

11 C G

A. Dance with me, we'll make them bleed. Sing for me, I'll sing a- long.

13 D Am Am

A. Sing for me, Oh, sing for me. List- en, do you hear it?

Example 1. Divinity: Original Sin II: Eight bars of 'Sing for Me', note the repeated patterns in the rhythmic sequences.

Transcription by Josh Dibble.

Lohse's mutism is selective, only occurring if the player asks her to play the lute, before defeating The Doctor. The implications mutism has on her identity means that her characterisations are complicated throughout the game. If the player chooses any of the other pre-set characters, their 'god' characters frequently tell them to kill Lohse as soon as possible because of the demon that

inhabits her. Lohse is doomed within the game world because characters constantly tell the player to destroy her as though she cannot be helped.<sup>50</sup> Lohse's song is a thank-you to the player for aiding her in a game world that had given up on her free will and identity. It is not necessarily because of Lohse's gender that she became mute, but her transformation into deviant through her song suggests a relationship between her mutism and classical Hollywood tropes of silencing female characters for emphasis on madness; although Lohse is not 'mad', the habitation and guidance of the demon means that certain characters treat her as 'mad'.

### **Combat and Comradery: *Final Fantasy XV***

This chapter has focused on the various instances of characterisation worldbuilding within RPGs, and the identification that occurs with the player through this worldbuilding. Successful worldbuilding of a game's people, and the player's character or avatar, is imperative to player engagement with the world. The identity of the player-character, however, is not only identified by their own voices. The inclusion of extra characters and voices, and their relationships, can actively portray personal narratives and identifiers.

*Final Fantasy XV* is an action RPG belonging to the Japanese RPG game series *Final Fantasy*. The *Final Fantasy* (*FF*) series is known for its emphasis on turn-based combat systems, long and complex plots, and an array of engaging characters for the player to control. The *FF* series is most notable for its composer until the twelfth iteration, Nobuo Uematsu. Uematsu's music is famous among fans and scholars alike, with a myriad of academic papers considering the music and sounds of *Final Fantasy VII*. The traditional use of looped cues are still seen in recent video games due to the compositional style's popularity. When discussing the musical differences that are seen between the Uematsu-composed pieces for JRPGs and compositions by western composers such as Jeremy Soul, Gibbons identifies that the use of repetitive loops have not changed in more recent JRPGs:

In JRPGs, the looped cues seldom change based on player input: the 'town' music sounds the same regardless of what activities the player is undertaking in that location, and 'battle' cues typically do not change based on how close the player is to winning or losing. In many recent

---

<sup>50</sup> The overarching plot of *Divinity: Original Sin II* is that the player-character or avatar is attempting to ascend to godhood after finding out they are 'godwoken' at the beginning of the game. The god of each race within the world (i.e., elves, dwarves, humans) is fighting with the other gods, and further enemies, for their chosen character to ascend. If the player allies with the other 'godwoken', the gods seem perplexed but focus their death-related comments on Lohse.

Western RPGs, however, composers seem fixated on creating music that closely reflects the player's moment-to-moment situation.<sup>51</sup>

The music of most *FF* games uses musical loops which represent player-characters, or 'agents', through the inclusion of leitmotifs. The significance of the musical loop and theme is its memorable qualities as players hear the loop continuously whilst playing the game. This creates a significance to the music, if the player remembers the game fondly, causing identifications with the video game even outside of playing; Phillips identifies the excitement of audiences in concerts which perform either the orchestral or popular cover versions of video game music.<sup>52</sup>

Character emphasis through leitmotif solidifies their appearance and reminds the player of their previous associations, goals, and backstory. *Final Fantasy XV (FF XV)* is no exception to this rule. *FF XV* is unique as a main-line *Final Fantasy* game in a gamut of ways. Although the *FF* series has evolved its combat systems over generations, *FF XV* is the first of the series to include an action-based combat system and a static party of four characters (this is excluding the MMORPGs *Final Fantasy XI* and *Final Fantasy XIV*, which are not single-player games). Traditionally, *FF* games I-X have incorporated turn-based combat which allows the player to strategically plan their combat attacks and defences, taking turns to attack enemies (like a complex game of Chess). Recent *FF* games, such as *Final Fantasy XII-XIII* have provided flow to this combat system using real-time, turn-based combat mechanics. Real-time, turn-based combat systems still incorporates turn-by-turn action, but the player must wait until their characters are ready to take an action, giving combat a sense of flow and action to it. Usually, the enemies in a turn-based combat system stand in one place and do not move, directing their attack at one character, whereas action-based combat gives the player control over spatial character movement during combat. The move to a fully action-based combat system in *FF XV* means that the player can attack as often as they like during any real-time combat, which means that the enemy can attack as often as they wish. This difference in combat symbolises the message of the game's goal, during each boot up of the game: a message appears before the main menu claiming that the game is 'A Final Fantasy for Fans and First-Timers'. The developers appear to have wanted to create a game that is inclusive for all. The combat system is simpler because the player only needs to hold one button for attack and one button for defence, however seasoned players have access to other combinations of actions such as warping to other points in the map, weapon skills, magic, blocks and parries. The reception of this change was mixed between fans who enjoyed the older combat system and believed that *FF* games should continue as turn-based, players

---

<sup>51</sup> William Gibbons, 'Music, Genre, and Nationality in the Postmillennial Fantasy Role-Playing Game', in Mera, Sadoff, and Winters, *The Routledge Companion to Screen Music and Sound*, p. 420.

<sup>52</sup> Phillips, 'A Composer's Guide to Game Music', pp. 56–57.



who enjoy action-based combat, and players who enjoy the game(s) regardless of combat mechanics, see Table 4. The new combat system, however, has created a unique space for voices in the combat of *FF XV*.

Positive Comments	Neutral Comments	Negative Comments
<p>Love it because that archaic system of turn-based fighting needed to die a long time ago, along with the other systems Square tried to tinker around before like in FFVIII. (ghost_cat, 2017)</p>	<p>It's not turn based so...if that's what we're judging it on, sure, it's bad. If you take time to get used to it and accept that it's not like the old games then it's pretty fun though. (MezZa, 2017)</p>	<p>I'm about 15 hours in and I don't love the combat much, it's too sluggish compared to any real action game and too shallow to be really tactical. Now that things don't die as fast it's a bit better, but still feels like you just need to keep your distance and spam the companion attacks when you can (when fighting bigger enemies. especiall when there is more than a couple). I hope I will get more options as the game progresses, so far it feels weird. (Quipido, 2017)</p>
<p>You mean best combat in the series right? Because it's excellent. Also comparing real time action to ATB is sort of a weird comparison to make. (kishinfoulux, 2017)</p>	<p>At least it doesn't drag like most of the menu based Final Fantasy games. I just replayed VI and oh boy is it a slog to get through the combat in that one. (Hunkulese, 2017)</p>	<p>I'm really not liking the combat so far, 5 hours in. Hold square and sometimes press other buttons while trying to wrangle the lock-on and camera. Also, they squandered the potential of the warp. I feel like they should have embraced the combat of character action games, made something slower and methodical, or thought of something radically different. What's here just isn't fun or</p>

		interesting. (wildpomme, 2017)
--	--	--------------------------------

Table 4. *Final Fantasy XV: Positive, neutral, and negative comments towards FF XV's combat system. These comments were made on a GiantBomb forum relating to FF XV. All typos are made by the users.*<sup>53</sup>

The developers have incorporated several collaborative combat mechanics between the four player-characters. If the player can position Noctis' attacks behind or on the side an enemy, Noctis will automatically commit to a 'blindside' or 'side-strike' combat action which incorporates another character. The game provides players with combat advantages if they focus on co-operation, as any action that uses more than one-character increases damage done to enemies. The game also provides the player with extra vocal lines that are only triggered during these moments, such as the chance of Gladiolus shouting 'You forget about us? Tell us what to do!'. In these moments, the vocal lines of the characters remind the player of certain combat mechanics, such as reminding the player to incorporate the other characters into the fight. Alongside this, the characters talk to each other in seemingly useless ways during combat: 'Yeuch. Not a fan of mushy desserts', or 'This is gonna take a while', using noticeably colloquial phrases such as 'mushy' and 'gonna'. The significance of these phrases is not always consciously recognised by the players, but they are used to bring focus to the types of enemies the player is against. The comment about desserts is heard when fighting the enemy known as 'Flan', also the name of a dessert, whereas the comment on time signifies the enemy's strength or resistance to the combat mechanics used by the player. As combat progresses the player fills a meter that, once full, allows the player to access a powerful special attack, called a technique, that is performed with a chosen companion; Noctis can also initiate a follow-up after these attacks if the player times a button press correctly.

### ***Voice Acting and Direction***

In *Transistor* we have witnessed the idea that other characters, whether playable or not, can direct the protagonist of the game. Typically for a JRPG, *FF XV* offers the player pre-set player-characters in the party, although they are not silent protagonists. The player has access to a single moveable protagonist, Noctis Lucis Caelum, and cannot physically control the three other characters Prompto

---

<sup>53</sup> 'Worst FF Combat? - Final Fantasy XV - Giant Bomb', *Giant Bomb*, 2017  
<<https://www.giantbomb.com/forums/final-fantasy-xv-84/worst-ff-combat-1804674/>> [accessed 27 May 2020].

Argentum, Ignis Scientia, Gladiolus Amicitia. The relationship between the four characters is stronger than the player initially sees because they have extensive backstories, which the player can witness through extra contextual Animé series and films.

Noctis is the crown prince of his kingdom, Lucis, and Prompto, Ignis, and Gladiolus are Noctis' royal guard for his coming-of-age journey. The player cannot directly control Prompto, Ignis, and Gladiolus, unlike other *FF* games, but they have control over the combined combat actions, their equipment, and skills. Immediately, the game puts emphasis on comradery within its combat system, allowing the three characters independent movement unless the player wants to cooperate in an attack on the enemy. There are only a few 'temporary' party members throughout the game because some characters leave on another 'mission' during the game; this is usually restricted to one chapter and the player does not have control over the skills and equipment of these extra characters. The regularity of the four characters in the game allows the voices of the game world to engage the player with Noctis and his team. As the player explores and engages with the game world, the four characters communicate with each other through voice lines consistently. *FF XV* puts an emphasis on this communication before the game begins, providing the game in multiple languages so that players worldwide can engage with the constant communication of the four characters' dialogue. The extra characters are also given enough dialogue and natural movement so that they feel like affective companions. Chris Carter examines, what Kristen Pullen might identify as, the naturalism acting style presented in the game as a key aspect of player to character relationship building:

[...] we're given the static viewpoint of Noctis and his boys. That would fall completely flat if Square didn't actually make that static cast interesting, but it did. I quickly picked up on certain cues and tropes and welcomed them, like Gladiolus' sense of overprotectiveness. Organic quests pop up not because they feel arbitrary, but because of, say, Prompto's love of photography.<sup>54</sup>

The extra dialogue, aimed towards the player-character and the player, can increase a sense of identification throughout the game as the player is constantly being communicated to; usually the characters ask about where to go next and where their next narrative goal is. These comments by the characters do not affect what the player wants to do with their time as they are only suggestions that may help the player if stuck or confused. The characters themselves also have their own personal interests, which are constantly communicated to the player. Voice lines personal to the

---

<sup>54</sup> Chris Carter, 'Review: Final Fantasy XV', *Destructoid*, 28 November 2016 <<https://www.destructoid.com/review-final-fantasy-xv-399859.phtml>> [accessed 29 March 2019].

characters build unique characterisations; Ignis' excitement whenever the player collects enough ingredients in the environment to create a new dish to cook, triggers the vocal line 'That's it! I've come up with a new recipe!'. This inclusion of consistent, and personalised, voice acting worldbuilds the game to include a sense that it is older, larger, and more complex than initially witnessed.

### ***Relationship Building: 'Stand by Me'***

The teamwork aspect of *FF XV* is unique because of the player's inability to physically control the other three characters, especially in combat. This mechanic was something tested in the previous single player instalment, *Final Fantasy XIII*,<sup>55</sup> as the player could only fully control the actions of one player-character in combat, but the game swapped players between different 'protagonists' throughout the game. This felt more like a way of storytelling multiple narratives rather than relationship building. The implication of only being able to fully control Noctis defines both the combat system and the identification process.

Although the combat system deviates, the music remains true to the *FF* series, boasting an orchestrated soundtrack that accompanies the game's cinematic cut-scenes, environments, and combat to worldbuild everything that surrounds the characters and their leitmotifs. The opening of *FF XV* emphasises the new systems of combat and storytelling for the game. This anticipatory scene launches the player into a stressful situation. The track 'Hellfire' is cued as the initial storytelling and worldbuilding begins with the message 'The Tale of the Chosen King, Saviour to the Star'. As low brass and a synthesised choir perform in a minor key, vocalisations come to the forefront of the soundscape. The heavy breathing of an unknown being continues as a male character, dressed in black, is presented physically and through text as Noctis Lucis Caelum; see Figure 15. The tense low brass exclaims Noctis, erupting into a major, fanfare chord as his name appears on-screen. The game flows between gameplay and cinematics as the characters Prompto Argentum, Ignis Scientia, and Gladiolus Amicitia are also named through on-screen text.

---

<sup>55</sup> Square Enix 1st Production Department, *Final Fantasy XIII*, Tokyo: Square Enix (PlayStation 3, Xbox 360, Microsoft Windows, iOS, Android, 2010).



*Figure 15. Final Fantasy XV: The ominous narrative foreshadowing of FF XV's opening, and the introduction of Noctis Lucis Caelum. Screenshot by Author.*

The cinematics transition from 'Hellfire', as the party is struggling against Ifrit, the fire summon used within all *FF* games, to a juxtaposed moment between a much younger Noctis and his father, the king of Lucis. The two say goodbye as the game implies Noctis has begun his coming-of-age journey. The game changes to a scene where Noctis, Prompto, Ignis, and Gladiolus are arguing over having to push their broken-down car in a desert environment. The comical scene of arguments and jests between the characters is again juxtaposed with the non-diegetic, angelic, singing voice of Florence + the Machine as she performs 'Stand by Me'.

It is not abnormal for a song with a female vocalist to be included in the *FF* series, however vocal performances are often written specifically for that *FF* game and is included several hours into gameplay, rather than in the opening. In *FF XV*, the cover of Ben E. King's 1961 song 'Stand by Me' is played during the title sequence and opening of the game. Before the song is triggered, however, the player is given mechanical control of Noctis and must push the car towards its goal. The player does this by pressing a trigger button which starts the car moving and triggers character dialogue. As shown in Figure 16, the player has to listen to the characters complain at each other, for the inconvenience of pushing a car, before the song begins.



Figure 16. *Final Fantasy XV*: The player's introduction to the characters' group dynamics. Screenshot by Author.

The game takes control away from the player again, when the car has been pushed a certain distance, so that the in-game camera can pan towards the sky, showing off the game's graphical capability; the title is presented across the sky as Florence sings the lyrics 'stand by me'. This early inclusion of the popular song thus solidifies the strong bond between the characters, and inevitably the player, foreshadowing that they will remain loyal and will stand by each other.

The cover of 'Stand by Me' retains the original lyrics of the song but uses female vocals instead of traditional male vocals in Ben E. King's original recording. The use of the female voice allows the player to hear both the song and the lower frequencies of the bickering dialogue of the male characters during the scene. The accompanying music for the vocals claims the song as part of the *FF* series. 'Stand by Me' replaces the introduction of the bass line, heard in the original Ben E. King recording, with the *FF* series' identifying 'Prelude' theme, a B $\flat$  major arpeggio on a harp; see Example 2.

## *Final Fantasy 'Prelude'*

The image displays a musical score for the 'Prelude' from Final Fantasy XV. It consists of two systems of music, each with a treble and bass clef staff. The tempo is marked as ♩ = 80. The key signature has two flats (B-flat and E-flat). The first system shows the initial eight bars of the piece. The second system begins with a measure number '3' above the treble clef, indicating the start of a new section or measure. The music features a prominent arpeggiated pattern in both hands, with accents (>) placed over many notes. The piece concludes with a double bar line and repeat dots.

Example 2. *Final Fantasy XV*: The original arpeggiated 'Prelude' theme. Transcription by Author.

Instead of the original 'Prelude' theme being played behind Florence's vocal performance, however, a variation of the 'Prelude' is heard. As shown in Example 3, the *FF XV* variation of 'Prelude' considers only a rising arpeggio in the key of A major, whereas the original arpeggiated 'Prelude' continuously rises and falls. As the vocals begin after these first eight bars, the theme on the harp continues as expected, with an ascending and immediately descending pattern. The vocals are highlighted by this reduction of the theme. The harp appears as an introduction to the vocals, highlighting the distinctiveness of the use of vocals in comparison to other *FF* introductions.



Example 3. *Final Fantasy XV: 'Prelude' theme in 'Stand by Me' with only its rising arpeggio to begin. Transcription by Josh Dibble.*

Throughout its rendition, the song draws parallels between the characters of the game and its lyrics:

*When the night has come  
And the land is dark  
And the moon is the only light we'll see*

One of the themes of the game is its use of Latin to identify its characters. The name of the player-character, Noctis, translates to night and thus relates to the first stanza of the song. Noctis' coming-of-age storyline resonates with the foreshadowing of the lyrics. The game consistently communicates with the player, through character dialogue, that it is dangerous at night-time due to daemons, and thus the party should make camp for the evening. If the player leaves the lit areas of camp or towns in the evening, high-level enemies will spawn and quickly defeat players at lower levels. The game's encroaching darkness during the game (night occurs earlier and earlier as the story progresses) is foreshadowed in the second stanza of 'Stand by Me'. The moon refers to Noctis' bride to be, Lunafreya Nox Fleuret (Luna translating to moon). Luna is the oracle of the game, a character who can speak to gods and summon god-like allies throughout the *FF* series. Childhood friend of Noctis, and political bride, Luna acts as a guide for the player-character as she continuously drives the actions of Noctis without being present. Luna is a crux of the player-character's identity, guiding Noctis to gods who can aid him on his coming-of-age quest. Luna's involvement in Noctis' story is important, highlighted by the lyrics, but is also problematic. Luna has no lack of agency, using



her voice and oracle powers to communicate with Noctis, gods, and the game world. Luna initially guides Noctis to his first summon, Titan, who resides in the centre of the world. When Noctis arrives at the area of Titan, he is spoken to in an ancient language which only the Titan understands, triggering a test of strength for the player. This cues the track ‘Apocalypsis Noctis’, intentionally written in Latin. As one of the first summons, ‘Apocalypsis Noctis’ identifies the gravitas of Titan and the importance of this test of strength.

Latin	English
Specie tua tantum carnem adme omnes habet cultum.	Through you alone there will be worship.
Ego sum qui mittit petras et gubernat omnes vitae. Caelum, Terra, aversione.	I am the one who sends the stone and rules all lives. Heaven, earth, aversion.
Magna, Parvus, reversusque. Fortis ductor, praecipio.	Great, small, return. Strong leadership and command.
Quod ligabis lucem vocis.	That binds to the voice of light.
Cum omnibus lucis plenus, Apocalypse finibus.	With all full light, the end of apocalypse.
Victum unus qui comedit hic vitae omnia: tenebris.	That which devoured all life, darkness.
Malum incarnatus surgit.	
Adducentes secum mortem.	Evil increases bringing death.
Para! Para! Para! Para!	For! For! For! For!
Fortis Rex, Bellator, in manus gladio.	The mighty king, warrior, with the sword in his hands.
Bellum in virtute, victum quod hostibus.	Battles in virtue, survives the enemy.
Magis mortem venit primum. Quod est quidam venire.	But death comes first. It is certain that it will come!
Para omnes! Para omnes! Para omnes!	To all! To all! To all!
Nunc est tempus, haec est locus ubi nos obstandum, omne tandem ex nobis unum! Omnes!	Now is the time, this is a place where we are established, finally, each of us! All!

Table 5. *Final Fantasy XV: Lyrics for ‘Apocalypsis Noctis’*. Translation partly by James Cook and partly gathered through an online fan translation.<sup>56</sup>

<sup>56</sup> U/omnialexmusic, ‘Translated Lyrics from APOCALYPSIS NOCTIS [Spoiler]’, *Reddit*, 2018 <[https://www.reddit.com/r/FFXV/comments/98wpi9/translated\\_lyrics\\_from\\_apocalypsis\\_noctis\\_spoiler/](https://www.reddit.com/r/FFXV/comments/98wpi9/translated_lyrics_from_apocalypsis_noctis_spoiler/)> [accessed 28 March 2019].

The lyrics of 'Apocalypsis Noctis', outlined in Table 5, are the main characteristics of this cue. The cue still uses the typical full orchestration that is known throughout *Final Fantasy* games, with strings, brass, woodwind, and percussion used. However, it is the male and female chanting choir that sings in Latin which highlights Noctis' identity. The lyrics tell of the might of Titan, who is predominantly constructed of earth: 'I am the one who sends the stone and rules all lives. Heaven, earth, aversion.' Although this fight seems almost impossible to the player as they begin the combat sequence, the lyrics, unknowing to any none-Latin speaking players, foretell Noctis' glory as a king and the possibility of him ending the darkness that has currently shrouded the world: 'With all full light, the end of apocalypse. That which devoured all life, darkness.' The lyrics complement the player-character's hidden identity as a powerful king in the future. It is not necessarily clear to the player that their character's identity is 'power', as Noctis is still coming-of-age. However, 'Apocalypsis Noctis' is an overbearing cue compared to the minimalist soundscape previously heard, identified in Chapter Four. Therefore, the cue is significant as it highlights Noctis' movement towards becoming more powerful during Titan's test. Although the cue is named after Noctis, and the lyrics foretell of his mastery of swords and his power as a king, Noctis still needs the assistance of his companions at this point. The use of a choir in this cue, especially when the sung voice is not used in other combat and exploration cues, portrays the significance of this particular act of combat. 'Apocalypsis Noctis' is used to signify the more important aspects of the narrative: Noctis fighting gods; gaining summons; and finding powers that will help Noctis to become king.

Further into the narrative, when Noctis has travelled to find Luna, Noctis fights another god, Leviathan, which uses the same cue as 'Apocalypsis Noctis' now named 'Apocalypsis Aquarius'. When Noctis attempts to bargain with Leviathan to aid him in his quest, he is defeated by the rampaging sea monster. Luna, who is already injured by an attack, gives her final life force to save Noctis, allowing the player-character to continue with their quest. Luna does not lose her agency because she chooses to sacrifice herself for Noctis. Instead, Noctis loses his agency because he is riddled by grief and anger, and the player fails one of their overarching goals to marry Luna. The light from the moon, as suggested in the lyrics, is removed, leaving both the player-character and player feeling a sense of loss. As Luna is lost, Noctis falls into depression and it is from this point onwards that, after some conflict because of Noctis' attitude, the team of four are the main relationship for the player.

Returning to the opening of the game. The chorus of 'Stand by Me' contains the lyrics 'stand by me', which highlight the relationship of the four characters to each other. Alongside this the use of the song parallels the 1986 coming-of-age film *Stand by Me*, in which the song is played over the end credits. The film follows four young boys on a journey to search for a rumoured human body. The

journey creates conflict amongst the boys, but the effects of the experiences on the journey result in friendship and comradeship by the end of the film. *FF XV* reflects the aspects of this film as *FF XV* breaks the tradition of the series by only providing the player with four characters. These four characters journey throughout the game and have internal conflicts but become strong comrades by the end.

Close to the final scenes of the game, Prompto speaks the line ‘we will stand by you’ to Noctis. The use of ‘Stand by Me’ in the opening of the game is a powerful tool when identifying the relationship between the player-character and functionally NPC companions. It is this message throughout the game that helps players to identify with this relationship. The significance of the comradeship theme, and the use of ‘Stand by Me’, can continue player identification with the four characters outside of completing the game; Table 6 shows some player’s reactions to listening to Florence + The Machine’s cover of ‘Stand by Me’ in isolation from the game.

Name	Comment
ohnosnakes	‘Me before playing: This looks fun, but I can’t see myself enjoying these boyband characters very much Me 1 minute into this song: If any of these precious boys is harmed in any way it will destroy me’
JAE 700	‘Anyone still EMOTIONALLY SCARRED by this song?’
Mr. Fluffles	‘This song really fits the game. When you hear it the first time when they are pushing the car, it’s like, ‘oh that’s cute they’re friends’ but it means so much more when it plays during the credits, and it means even more after playing through the Episode DLCs, because you get to watch them make the promise to protect Noctis no matter what, and to always stay by his side. The story of <i>FFXV</i> rivals even the best titles in the series, leaving players old and new astounded at the storytelling ability of this beautiful game, and the amazing people who made it.’
Tony Joestar	‘Game start: ‘Oh hey this is neat song, it goes with sticking by your friends’ Game end: UGLY CRYING’

Table 6. *Final Fantasy XV*: YouTube user comments from the official Florence + The Machine - ‘Stand by Me’ video.<sup>57</sup>

<sup>57</sup> Florencemachine, ‘Florence + The Machine - Stand By Me’, *YouTube*, 12 August 2016 <<https://www.youtube.com/watch?v=vv2DSmy3Tro>>.

The comments in Table 6 are only a few examples of the effect the comradeship theme had on players. The occurrence of the song both at the beginning and end of the game reinforces this theme, and players identify with the personal habits and voices of the characters.

The case studies of *Transistor*, *Divinity: Original Sin II*, and *Final Fantasy XV* consider differing uses of voices to build character characteristics. Although each case study portrays characterisation building, and a sense of identification for players, they do so through different handlings of male and female characters. In video games and, likely, multimedia as a whole the loss of the voice has mostly been appointed to the mute female character. Red and Lohse by no means lose combat agency in their worlds, but because of their associations of female characters, their gendered characteristics are handled through the narrative of their muteness. Regardless of the gendered use of the voice, however, each case study identifies voices that are used to signify player-characters, as opposed to the environment, people, or culture of the game world: The male voice of the sword in *Transistor* is used to define Red's personal story and herself as a performer; Lohse's regaining of her performative voice allows her to define herself with her song, 'Sing for Me'; and the use of female vocals in 'Stand by Me', alongside male dialogue, portrays the companionship of the four characters in *FF XV*.

The identification process between the player, their playable-character(s) or avatar, and other NPCs is contained within voices. Emotions such as empathy aid the player in being involved in the game world and allow the worldbuilding process to create identification through characterisations. The grain of the voice, accents used, gender and cultural stereotypes can all identify the agency of a character to the player, and their characteristics; the dialogue of characters can also build a sense of personality and agency in their goals. Returning to Tronstad's consideration of two types of identification, empathic and sameness, it is not to say that all players are engaged with the same characters and characteristics. Although a suspension of disbelief is often applied to games, the inclusion of certain races, cultures, and genders to a playable-character might entice some players to engage with certain characteristics over others. The player avatar is more likely a consistent method of allowing players to engage with the characteristics of their playable model in the game; players can make the avatar look how they would like and can reflect themselves in the process. However, the lack of voice, and therefore the character's own agency, with avatars changes the dynamic of the relationship; it may not be empathy that the player feels with their avatar, but a sense of insertion.

## Chapter Six: The Witcher 3: Wild Hunt

Chapters four and five have identified the ways in which environmental and characterisation worldbuilding can be enhanced with the inclusion of voices. Voices as worldbuilding tools can aid the construction of a coherent virtual world, providing audio signifiers of people and culture alongside characterisations which may lead to increased identification by the player. These previous chapters have provided shorter individual case studies that examine how the representative features of voices effectively enhance worldbuilding. The aim of this chapter is to show how these worldbuilding processes can be used in a single RPG, encompassing the environmental and characteristic components of diegesis worldbuilding together over a larger open-game world with connections to lore from pre-existing literature.

*The Witcher 3: Wild Hunt* (*The Witcher 3*) was given the status of one of the ‘greatest video games of all time’ by a variety of online video game-oriented newspapers, journalists, and critics.<sup>1</sup> The game’s status was given due to its complexity and coherence as an RPG, which provide the player with: intricate narratives and outcomes depending on player decisions; character developments that affect the narrative and can be controlled partly by the player; characters and NPCs that can be interacted with in a multitude of manners, each with personalities relating to their class, culture, and kingdom; meaningful quests and side-quests that affect narrative and character development; a varied combat system that allows the player the fight with sword combat, magic, or potions and bombs; an in-game card game, known as Gwent, which was adapted into its own game (*Gwent: The Witcher Card Game*)<sup>2</sup>; an open-world environment with a complex ecosystem full of objects, characters, and enemies to engage with (rather than an empty but expansive environment); and a soundscape that identifies each environment of the game effectively, reacts to player action, and identifies characters through voices and accent. *The Witcher 3* is part of a series of Polish medieval fantasy RPGs, based on short stories and novels written by Andrzej Sapkowski; Sapkowski’s literature was adapted by the Polish video game developer *CD Projekt Red* to create *The Witcher*,<sup>3</sup> *The Witcher*

---

<sup>1</sup> Examples include: PC Gamer, ‘The PC Gamer Top 100’, *PC Gamer*, 31 August 2018 <<https://www.pcgamer.com/the-pc-gamer-top-100-new/12/>> [accessed 4 July 2019]; Polygon, ‘The 500 Best Games of All Time: 100-1’, *Polygon*, 1 December 2017 <<https://www.polygon.com/features/2017/12/1/16707720/the-500-best-games-of-all-time-100-1>> [accessed 4 July 2019]; IGN, ‘TOP 100 VIDEO GAMES OF ALL TIME’, *IGN* <<https://uk.ign.com/lists/top-100-games/100>> [accessed 4 July 2019]; See John Tones, ‘Los 100 Mejores Videojuegos de La Historia’, *GQ*, 17 October 2018 <<https://www.revistagq.com/noticias/tecnologia/galerias/los-100-mejores-videojuegos-de-la-historia/8951?pagina=5>> [accessed 4 July 2019].

<sup>2</sup> CD Projekt Red, *Gwent: The Witcher Card Game*, Warsaw: CD Projekt (Microsoft Windows, PlayStation 4, Xbox One, iOS, Android, 2018).

<sup>3</sup> CD Projekt Red, *The Witcher*, New York: Atari (Microsoft Windows, OS X, 2007).

*II: Assassins of Kings*,<sup>4</sup> and *The Witcher 3: Wild Hunt*. The adaptation of Andrzej Sapkowski's abundance of *The Witcher* literature presents the case study as an excellent example of transmedia worldbuilding. CD Projekt Red had to establish the aesthetics of the books, alongside the extensive lore, people, and stories of the literature, whilst establishing their own narrative of Geralt and associated characters. Where *The Witcher 3* is critically acclaimed, the previous two iterations of the series received mostly positive receptions but were criticised because of the combat's difficulty. The combat system was refined for *The Witcher 3*, and the technological advancements that occurred as the series' development allowed for an extensive and detailed open-world environment to be built, enhancing the series' acclaimed RPG elements.

The series establishes a northern European medieval fantasy aesthetic, combined with Polish folklore, that grew in popularity in the early 2000s. The northernness of medieval fantasy texts is outlined by James Cook, in his paper 'Sonic Medievalism, World Building, and Cultural Identity in Fantasy Video Games', as a stereotypical cumulation of culture and people from mostly European Middle Ages history:

Northernness in this literature stretches far further than Scandinavian lands, however, encompassing Central Europe and especially Celtic traditions too. [...] Tolkien clearly mixed Celtic and Germanic/Scandinavian ideas under the umbrella of Northernness. [...] Celtic and Germanic/Scandinavian Northernness are, nonetheless, clearly differentiated with, according to Burns, the former associated with the fairy world, 'mysteries of the spirit,' and with charm, sensitivity, and imagination. Contrastingly, the latter is represented by ideals of heroism, courage, manhood, determination, and stamina.<sup>5</sup>

The northernness of medieval fantasy was prominent around *The Witcher's* release in 2007, between the completion of the *LoTR* film trilogy in 2003, and the first episode of *GoT* airing in 2011. Adaptations of medieval fantasy literature have remained popular since *The Lord of the Rings* trilogy popularised medieval fantasy with western audiences; Amazon announced the development of two medieval fantasy television series, *The Lord of the Rings* and *The Wheel of Time*, in 2018, whilst Netflix released a television series of *The Witcher* in 2019. In the video game industry, the developer and publisher Blizzard Entertainment released one of the most popular MMORPG *World of Warcraft* in 2004. *WoW's* aesthetics provide a similar medieval and fantasy style to *The Lord of the Rings*, providing players with immersive control of avatars in a fantasy environment. Bethesda Game

---

<sup>4</sup> CD Projekt Red, *The Witcher 2: Assassins of Kings*, Warsaw: CD Projekt (Microsoft Windows, Xbox 360, OS X, Linux, 2011).

<sup>5</sup> James Cook, 'Sonic Medievalism', in *Studies in Medievalism XXIX*, ed. by Karl Fugelso, Politics and Medievalism (Studies), NED-New (Boydell & Brewer, 2020), pp. 217–38 (p. 224) <<https://doi.org/10.2307/j.ctvxhrjqn.18>>.

Studios, having already released their first Elder Scrolls game *The Elder Scrolls: Arena*<sup>6</sup> in 1994, and second game *The Elder Scrolls II: Daggerfall*<sup>7</sup> in 1996, finally released *The Elder Scrolls III: Morrowind*<sup>8</sup> in 2002. *Morrowind* provided the aesthetic of medieval fantasy in an open-world RPG environment at the time of the genre's popularity in film.

*The Witcher 3* is different to the RPGs of *The Elder Scrolls*, *WoW*, and typical *D&D* style games, because Geralt is the player-character of the entire series, rather than a customisable avatar. As the player-character, and prevalent witcher of the game world, the player has little customisation of Geralt. Geralt is furnished with a personality, backstories, relationships, scars, and other non-customisable features that identify Geralt as a character with his own agency and goals; the player can only change the appearance of his hair and beard (by paying a barber) and his armour and swords. As a witcher, Geralt is a mutated human who belongs to a class of travelling monster hunters; witchers undergo trials and transformations as children to allow them to use magic and potions which enhance hunting capabilities. The witchers are a dwindling people by the time of *The Witcher 3*, feared by human civilisations due to their use of magic and bodily differences, such as mutated cat eyes. The unchangeable features of a witcher forces the identification process between Geralt and the player to rely on his pre-made characterisations.

The soundscape was composed by Marcin Przybyłowicz and performed by the Polish folk band Percival (Christina Bogdanowa, Katarzyna Bromirska, Joanna Lacher, and Mikołaj Rybacki), whose name represents the character Percival Shuttenbach from the original *The Witcher* novels.

Przybyłowicz discusses wanting to create an authentic soundscape for *The Witcher 3*, differing from other fantasy media of its type by producing music and sounds that you cannot create with modern instruments, as Przybyłowicz explains:

We want the whole audio to be as organic as possible, we want the player to feel the naturalness and not that we are creating some kind of illusion, creating a video game and implementing sound to the locations but we want to create an impression of being in that reality as Geralt himself.<sup>9</sup>

Alongside Percival, other talented instrumentalists were invited to perform and record medieval, Mediterranean, and Slavic instrumentation such as the lute, hurdy gurdy, renaissance fiddle, saz (Turkic string), kemenche (east Mediterranean string), yayli tanbur (Turkish string), gheychak (Asian

<sup>6</sup> Bethesda Softworks, *The Elder Scrolls: Arena*, Rockville: Bethesda Softworks (MS-DOS, 1994).

<sup>7</sup> Bethesda Softworks, *The Elder Scrolls II: Daggerfall*, Rockville: Bethesda Softworks (MS-DOS, 1996).

<sup>8</sup> Bethesda Game Studios, *The Elder Scrolls III: Morrowind*, Rockville: Bethesda Softworks (Microsoft Windows, Xbox, 2002).

<sup>9</sup> Gamemusic.pl, 'The Witcher 3: Wild Hunt - The World Through The Composer's Eyes', *YouTube*, 21 August 2014 <<https://www.youtube.com/watch?v=Q4DnfmhJiGM>> [accessed 1 August 2018]. - 01:53-02:18

string), bouzouki (Greek plucked string), mandolin, hammered dulcimer (percussion), and the bowed gusli (Russian plucked multi string). The main instruments and musical style heard in *The Witcher 3* are not typically used in Western cinema and television. The soundscape of the game builds an alien lore through coherent audio and visuals, symbolising a dangerous world filled with monsters and the unknown. The unique sounds, rhythms, and styles of the voice and instrumentation was produced through improvisation, rather than pre-composed pieces, as the instrumentalists were mostly self-taught and could not read the notation that Przybyłowicz originally created:

I found out that most of them are self-taught, and don't necessarily need notes. Only Kasia from Percival finished music school. The rest learned how to play instruments by themselves...And then I had an idea to just throw my initial approach out the window and try going about this another way. It looked something like, Mikołaj, take the saz and play something like this, you over there take that, you do this, and so on, and let's go from here and we'll see where it takes us...And that turned out to be the common language we were missing at first. From that moment things really started to come together quickly... A lot of the sound we did for *The Witcher 3* is the result of improvisation.<sup>10</sup>

My own experience of *The Witcher 3* began in 2017. My playthrough of the English version was initially casual, rather than a dedicated analytical playthrough. It was when I began to see the fascinating and careful placement of voices, to identify an array of narrative situations, the environment, diverse characters, and the player-character themselves, that I understood its significant contribution to using voices in video games. *The Witcher 3* includes voices as dialogue, song, and vocalisations within the diegetic, non-diegetic, and fantastical gap spaces of the soundscape, as both reactive and non-reactive sounds.

## Environments and Accents

Environmental worldbuilding, and spatiality, in *The Witcher 3* is created through reactive audio and visual representations of the various areas in the game world. The reactive audio of each environmental soundscape changes depending on where the player is and their current actions, whether this is exploring the world or in combat with enemies. Voices are used as a part of this reactivity to the player, existing amongst different languages and accents within the diegetic and non-diegetic space. As the player travels, and explores the game world, they are provided access to

---

<sup>10</sup> The Witcher, 'The Witcher 3: Wild Hunt || Creating the Sound', *YouTube*, 30 August 2016 <<https://www.youtube.com/watch?v=g3B2QHIT1Q4>> [accessed 4 September 2017].



various kingdoms, provinces, cities, and villages on the mainland. As the player explores, the accents of the game's NPCs change depending on the kingdom and province. These accents can alter radically between neighbouring areas and effectively distinguish the diversity of NPCs across, what is assumed to be, an expansive world.

The environments in *The Witcher 3* are divided into six initial areas, shown in Figure 17, and one DLC (downloadable content) area that was released with the *Blood and Wine* expansion in 2016: the province and village of White Orchard in the kingdom of Temeria; the Royal Palace in Vizima, a castle in the city of Vizima located in Temeria; the northern province of Velen located in Temeria; the northern city of Novigrad located in the kingdom of Redania; the witcher fortress, Kaer Morhen, in the kingdom of Kaedwen; the kingdom of The Skellige Isles; and (if players own the additional DLC) The Duchy of Toussaint in the Nilfgaardian Empire to the south. Each area has unique environmental geographies, such as climate and weather system, alongside distinctive human geographies such as societies, culture, and economy.

The game begins with a cinematic of Geralt and the character Vesemir tracking an unknown female character, later introduced as Yennefer. When the player is introduced to gameplay it is within a flashback dream sequence in the witcher fortress of Kaer Morhen. For players new to the series, they are introduced to key aesthetics and gameplay at Kaer Morhen, including Geralt and Lambert's unusual low American accent alongside the British accents of Yennefer, Vesemir, and Ciri. This tutorial sequence swiftly vanishes, as Geralt awakes from his dream after the player has shown they can control the player-character sufficiently, and an open-world village named White Orchard is presented. Although not a standardised tutorial area, White Orchard provides the player with a few quests of gathering ingredients, fighting enemies, and talking to NPCs to practice gameplay mechanics learnt in Kaer Morhen. The NPCs of the village incorporate southern British and East End London accents.



Figure 17. *The Witcher 3: Map of the Northern Territory, excluding The Duchy of Toussaint further south. Screenshot by Author.*

After the player has time to become accustomed to the game, the narrative begins as Yennefer and the antagonists, The Wild Hunt, are introduced as Geralt rides for the Royal Palace in Vizima. The palace in Vizima has been occupied by the Nilfgaardian empire, an empire at war with the northern realms and The Skellige Isles. The accents of Nilfgaardians are alien to the British and American accents of the game environments, creating a vocal otherness by contrast, to symbolise the invading empire of Nilfgaard.

After Geralt, begrudgingly, accepts a quest from the current emperor of Nilfgaard, Emhyr var Emreis, to find his daughter Ciri, the player is finally introduced to the game's open world environments, beginning with the marshes of Velen. The rural environment, pagan marshes, and farming villages of Velen incorporate west country accents to stereotypically portray a land of peasantry; Cook compares Velen to the 'Shire' in Peter Jackson's *The Lord of the Rings*, although the marsh terrain and stormy climate are not as pleasant.<sup>11</sup> Velen leads the player to the two main city hubs of the game, Oxenfurt and Novigrad, which the player can explore whenever they choose. The people and culture vary between the cities and the rural environments as Oxenfurt's residents use the queen's English to identify a wealthy, University city. The city of Novigrad adopts stereotypical 'Northern'

<sup>11</sup> James Cook, 'Playing with the Past in the Imagined Middle Ages: Music and Soundscape in Video Game.', *Sounding Out!*, 3 October 2016 <<https://soundstudiesblog.com>>.

British accents to represent itself as the northernmost point of the game's environments outside of Kaer Morhen; this stereotypical northern accent is a mixture of Yorkshire, Lancashire, Liverpool, and Newcastle accents which are commonly used in medieval fantasy genres, such as *GoT*, when representing a northern or relatable character.

The environments and people of White Orchard, the Royal Palace in Vizima, Velen, Oxenfurt, and Novigrad incorporate consistent British accents that alternate depending on the player's location. Britishness and northern European influences in medieval fantasy may be partially due to the construction of medievalisms as the 'historical'. Heavy stereotyping to reinforce real-world associations was discussed in Chapter four as a tool to aid an audience's recognition of environments, people, and culture. Therefore, leaning on British and European accents to portray a medieval fantasy environment may be due to the inaccurate associations of neomedievalisms. James Cook, Alexander Kolassa, and Adam Whittaker discuss both *The Witcher 3* and *GoT's* use of musical neomedievalisms to construct this believable, yet inaccurate medieval and fantasy world:

Rather than waste time with lengthy narrative expositions, such structures are instead sketched through widespread association with a distinctive vocabulary of medievalisms, be they pagan, Christian, urban, rural, etc. Strategies of this sort abound in fantasy. They serve to establish a complex pattern in which both the alterity and familiarity of the fantastical world is created and maintained. It is this contradiction that has much narrative power, making the fantasy world the perfect site to negotiate complex issues. *GoT* is no different in this respect, borrowing tokens of medievalism to construct a believable world whilst constantly making nuanced references to, from, and outside of an analogous Middle Ages, and/or present day.<sup>12</sup>

What Cook, Kolassa, and Whittaker suggest here is the inaccuracy of perceived medieval principles, visuals, and soundscapes. Present day moral codes and modernist ideologies are incorporated into medieval fantasy alongside general associations of chivalry and barbarisms associated with the era. This stereotypical view is also inherently western. Neomedievalisms focus on Britain's vast history but ignore the middle ages and medieval periods that would have occurred throughout the rest of the world.

This focus on British accents in *The Witcher 3* creates a sense of otherness and alienisms in juxtaposing accents. When the player eventually unlocks access to The Skellige Isles, they encounter Viking-esque NPCs. The micro-culture and people of Skellige are represented by a mixture of Celtic

---

<sup>12</sup> James Cook, Alexander Kolassa, and Adam Whittaker, 'Music in fantasy pasts: neomedievalism and Game of Thrones', in Cook, Kolassa, and Whittaker, *Recomposing the Past*, pp. 233–34.

and Nordic mythology, architecture, culture, and environment. The NPCs of the Isles have unique Irish accents, not heard on the mainland, and the environment changes from the marshes of Velen to pristine mountain ranges and an abundance of wildflowers. The emphasis on cultural, visual, and audio difference in the Isles considers Nordic and Celtic associations as otherness. Shiloh Carroll argues that this use of Celtic mythology is separate to Tolkienisms because of Tolkien's aversion to using Celtic mythology in *The Hobbit* and *The Lord of the Rings*.<sup>13</sup> As medieval fantasy media, and *The Witcher 3* often use Tolkienisms as influence, Celtic mythologies thus may appear unfamiliar to audiences; the Celtic is separated from the norm of the witcher world, removing Irish accents from the British by a sea.

It is not just differing accents that create a separation from the norm of the world, however. Vocalisations have a prominent position in the non-diegetic soundscape. Where the English language (known as common speech) narrates the humanisms of the world, the use of Polish and Slavic languages, during female vocal performances in the non-diegetic space, signify Geralt's otherness in the world.

### ***A Rare Glimpse of The Sun: Velen's Marshland***

Velen encompasses the greater environment of the witcher world as an expansive rural, marsh landscape and the two cities of Oxenfurt and Novigrad. The rural marshland of Velen consists of a vast, war-torn environment currently at war with the Nilfgaardian Empire. The climate reflects the ravished landscape with infrequent spells of sunshine interrupting the bleak cloudy, rainy, and stormy skies; the significance of Figure 18 is the lack of rain in the Velen marshland.

---

<sup>13</sup> Shiloh Carroll, *Medievalism in a Song of Ice and Fire and Game of Thrones* (Cambridge: D. S. Brewer, 2018), p. 113.



Figure 18. *The Witcher 3: A rare glimpse of the sun in Velen. Screenshot by Author.*

Away from the cities, and amongst the woodland and small farming villages of Velen, the soundscape incorporates naturalism in its diegetic sounds and discomfort in the non-diegetic music. Wind, rain, and bird sounds are heard in the diegetic soundscape of Velen, building the natural environment. However, with the abundance of monsters and bandits that live in the wilderness of Velen, the non-diegetic soundscape must portray an element of unease and danger; the abundance of enemies in the environment is partially due to the escalating war with Nilfgaard in the area. Although the aim of Velen's soundscape is to signify this war, the music still incorporates musical signifiers for woodlands and a 'pastoral' style, to identify the landscape of Velen before the war.<sup>14</sup>

<sup>14</sup> The musical cues of Velen incorporate drones and short repetitive melodies that are considered by Philip Tagg's analysis of ABBA's 'Fernando' as musical associations of the pastoral; Philip Tagg (no date) *Fernando, part 1*. Available at: [http://tagg.org/Clips/HTML5/Fernando00-04\\_VP8.webm](http://tagg.org/Clips/HTML5/Fernando00-04_VP8.webm) (Accessed: 14 August 2017).

The image shows a musical score for four instruments: Treble Lute in D, Violin I, Viola, and Violoncello. The music is in 4/4 time. The Treble Lute part has a short rhythmic pattern of eighth notes. The Violin I part is mostly silent. The Viola part has a drone on the first bar and a small movement in the second bar. The Violoncello part has a drone on the first bar and a small movement in the second bar.

Example 4. *The Witcher 3*: Opening musical cells of Velen's outskirts, including a short rhythmic pattern on the Lute, drones and small movements on the strings. Transcription by Author.

Example 4 identifies that the 'pastoral' elements of Velen are created through the cello's use of drones and small movements revolving around an A. This anchoring around an aeolian tonality halts the potential 'rising sun'-style motif in the second bar of the viola. This tension between pastoral-style melodic movement and the tension implied by stasis mirrors the grim climate and political situation of Velen, stopping its soundscape and musical leitmotifs from developing. The stunted growth of Velen's musical signifiers represents its failing economy and danger of its people falling to famine and the dangers of the wilderness.

Although Velen may appear to be trapped in its own grief, the soundscape includes multiple musical cues that adapt to micro shifts in location. As discussed in Chapter four, the effectiveness of a soundscape to worldbuild a video game environment is its ability to adapt not only to differing areas, but to the micro changes of visuals and gameplay within a single environment. An interview with Przybyłowicz identifies that the aim of *The Witcher 3*'s soundscape was to create nuance within a single environment: 'every nuance matters, every change in the weather affects how the world around you will sound. The wildlife, rustling of trees, a river's current – this will all sound different during a strong wind or a storm than it would otherwise'.<sup>15</sup> The everchanging soundscape, and landscape of Velen, is signified through the inclusion of multiple cues of music that adapt from the lute cell in Example 4

<sup>15</sup> The Witcher.

The image shows a musical score for five instruments: Marimba (Mar.), Treble Lute (Treb. Lute), Violin I (Vln. I), Viola (Vla.), and Violoncello (Vc.). The Marimba part is in the treble clef and features a rhythmic pattern of eighth and sixteenth notes. The Treble Lute part is also in the treble clef and is mostly silent. The Violin I part is in the treble clef and has long notes and a drone. The Viola part is in the bass clef and has long notes and a drone. The Violoncello part is in the bass clef and has long notes and a drone.

Example 5. *The Witcher 3: Velen's marimba leitmotif, an adapted rhythmic pattern from the lute. Transcription by Author.*

Example 5 shows the expansion of the lute's rhythmic pattern and its movement to the marimba, with drones continuing to centre around an A, but with the inclusion of a violin line. As the cue progresses, and the player explores, violins and other string instruments perform long notes and drones to create the basic sounds of Velen. Other cues of Velen signify the discomfort of the marshland and its dangers through uncommon instruments. Non-orchestral string instruments such as the yayli tanbur (a Turkish long-necked lute) are introduced in this area, with the performers applying pressure to the strings to create a feeling of tension in the sound; this tension is likely a sense that developers hope will be transmitted to the player. Percussive instrumentation often takes the forefront in Velen, symbolising the wooded environment. Cymbals are played using wooden sticks to make the sound of movement and creaking trees to put the player on edge. 'Monster' sounds are also used to portray the dangers and unease of Velen, with elephant-like sounds used infrequently in the cue accompanied by digeridoos. The combination of Australian, Turkish, and medieval instrumentation is significant in stopping the player feeling at home. Geralt is never truly accepted in Velen, and thus his associated 'home' remains in Kaer Morhen. The soundscape of Velen blends the non-diegetic and diegetic sounds so that the player might struggle to discern what sounds are a physical danger to Geralt. The wilderness of Velen builds an untamed reputation through a lack of voices in these cues because of voices' associations with safety, like in *Crypt of the NecroDancer*. Velen's voices are kept to NPC dialogue on the outskirts of the wilderness, as inhabited villages are often free from monsters.



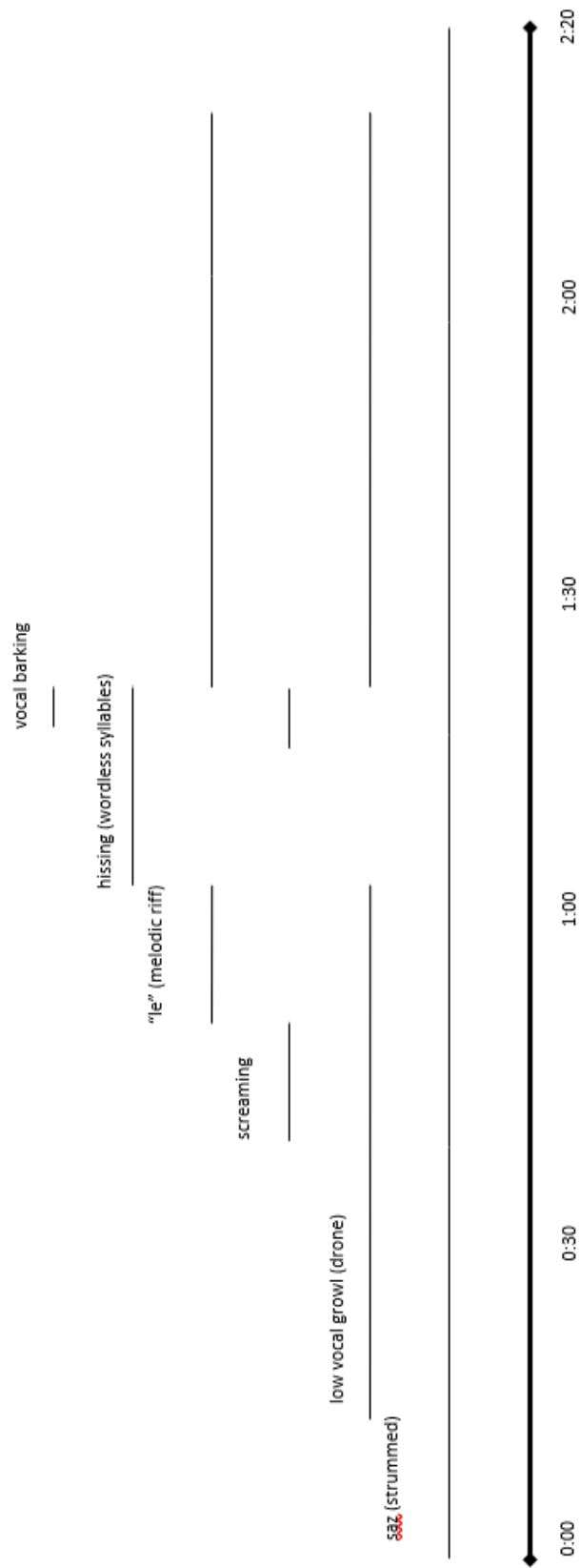
## ***Geralt's Signifiers***

As Velen's marshland builds musical signifiers of its dangers, there is little space in the soundscape to represent the movements of Geralt. Geralt's musical signifiers are thus attached to himself and his actions, in the format of combat music. If the player can be anywhere in the witcher world at one time, as they have access to fast-travel, it may be difficult to simultaneously signify an environment and the player-character in the non-diegetic space. Instead, Geralt's musical signifiers are located within the female voices of non-diegetic combat music; these vocal signifiers of Geralt portray him as a complex being which goes against the 'aggressive' monster hunter identifiers that the game-world's inhabitants believe him to be, the developers critiquing the views of Geralt as a 'combat machine'.

Combat, in *The Witcher 3*, is live-action and the player can attack enemies in a multitude of ways. As combat does not need to take time to load, players can move seamlessly between exploration, combat, and back to exploration or fleeing combat. Geralt also retains his agency as a character within combat as he automatically unsheathes a sword when he comes within a distinct distance of an enemy, without needing player interaction. The audio supports this fluidity by adapting with a fade-in of combat cues from the environmental music, likely incorporated through adaptive compositional methods such as vertical layering. The voices of the combat themes, recorded by the band Percival, are used to identify Geralt as a character, where Geralt is, and who or what Geralt is fighting. Where Geralt's agency is contained within the voices of the combat cues, the vocals adapt to enemy type and define either humans or monsters, using the steel or silver sword. The two most notable combat cues, 'Silver for Monsters...' and '...Steel for Humans', are heard in the provinces and kingdoms of the mainland and signify enemies through conflicting uses of the voice.

The cue 'Silver for Monsters...' triggers when the player fights monster enemies. Voice and language are not used in this cue as the female vocals of Percival use vocalisations to mimic unpleasant sounds of a monster attack. Nonsensical hisses, screams, and barks continue the boundary blurring of diegetic and non-diegetic sounds to unbalance the player and intensify the combat. The main vocalisation of the cue is the syllabic 'le', which acts as the cue's melody which is outlined in the graphic transcription in Example 6.





Example 6. *The Witcher 3: Graphic Transcription of instrumentation and vocal techniques in 'Silver for Monsters...'*.  
 Transcription by Author.

The nonsensicality of the vocalisations portrays the indiscriminate violence and ferocity of the monsters, and the lack of any possibility to reason or negotiate. The use of vocalisations in the sung voice also contains a tactic for creating unease with players. A study of the difference in visual information and perception of sung and spoken syllables was conducted by William Forde Thompson, Lena Quinto, Frank A. Russo, and Sandra E. Trehub, who discussed the differences in clarity between speech and singing, as when singing there is a constraint on pitch movement.<sup>16</sup> Pitch movement, and the rhythmic and intonation properties of prosody are considered by Annabel J. Cohen as emotional meaning makers for listeners, however listeners tend to focus only consciously on this meaning rather than the prosodic information given.<sup>17</sup> The creation of unease and anxiety in the player's emotional state when the cue of 'Silver for Monsters...' occurs creates this focus on combat and the grimness of fighting monsters. The possibility of a lack of clarity in sung syllables, combined with Cohen's notion of 'inattentive deafness' on prosodic information, may cause discomfort in listeners not receiving audio information they are used to in these situations.

These emotionally unclear vocalisations dominate the musical cue. The cue, which may loop or cut short depending on the player's length in combat, contains mostly vocalisations, with the exception of a strummed saz, as seen in Example 6. Although these female voices do not signify emotions, they do engage with the enemy type and Geralt himself. The significance of these Polish vocals being within the non-diegetic soundscape during combat is their consistency of appearance when Geralt is in combat.

The twinned cue '...Steel for Humans' uses the voice, instead of mostly vocalisations, to signify combat with humans. As human contact occurs mostly in human settlements, or at bandit camps, language is then involved in the non-diegetic space. The voice includes Bulgarian lyrics that portray the sentience of humans and their ability to discriminate attack Geralt. '...Steel for Humans' was initially composed by Percival, as an adaptation of folk song named 'Lazare (Bulgaria)', for the album *Slava! Pieśni Słowian Południowych* (translated to *Slava! Songs of South Slavs*) in 2012. Although the song uses voice to signify humanistic behaviours, the lyrics portray both Geralt's backstory and the pagan and pastoral environment of Velen. The lyrics, as seen in Table 7, portray Geralt's relationship with women and marriages, or joinings. A narrative in the literature and Netflix adaptation, as opposed to the game, shows Geralt's relationship with Ciri's mother and father and his involvement with their union. Geralt himself is also fated to be joined forever with the character Yennefer, due to

---

<sup>16</sup> William Forde Thompson, Lena Quinto, Frank A. Russo, and Sandra E. Trehub, 'A Comparison of the McGurk Effect for Spoken and Sung Syllables', *Attention, Perception & Psychophysics*, 72.6 (2010), 1450–54 (p. 1451).

<sup>17</sup> Annabel J. Cohen, 'Music as a source of emotion in film', in Juslin and Sloboda, p. 884.

a wish he made through a Djinn. Geralt's fate and relationships with women is something teased through female voices throughout the game, and eventually acts as his normaliser in the world, which I shall discuss further later in the chapter.

Traditional Lyrics	English Translation
Raduvay se, raduvay	Be happy, be happy
Oy Lazare, Lazare. Tuka ni sa kazali	Oh, Lazarus, Lazarus. Here we're told
Kolko liste po gorach tolko zdrave na taz kysycha.	That as there are leaves in the forest That much health be given to this home
Tervo tuka doydome Moma momche naydeme	We have come here To find the boy a girl
Ya momata godete Ya momcheto zhenete.	Wed the girl Wed the boy

*Table 7. The Witcher 3: Lyrical Translation of '...Steel for Humans' from Latinised Bulgarian to English.*

The simplicity of the lyrics, speaking of wedding boy and girl amongst leaves, forest, and good health, refer to the pastoral elements of Velen. Although the climate has not changed in Velen since the outbreak of the war with Nilfgaard, the ravished environment and frequent human deaths are a breeding ground for enemies which are created from, or feed on, corpses, such as Drowners and Rotfiends. Before the war, however, the landscape of Velen would have been relatively rural and pastoral, with now abandoned farms and villages inhabited. The mention of the natural in the song somewhat relates back to how the player could imagine Velen's environment before the destruction by the Nilfgaardian invasion. The lack of translation to English, however, maintains discomfort for non-Bulgarian speaking players even with the inclusion of the voice. The uncommon instrumentation which accompanies the voice, such as the saz, bowed gusli, hurdy-gurdys and hammered dulcimers, alongside the use of dissonance and heavily accented singing does not mirror the pleasant lyrics presented.

Both cues, 'Silver for Monsters...' and '...Steel for Humans', address the unease and tension of Velen, and the act of combat, through the use of uncommon instrumentation. Humans and monsters are thus divided by the distinctions between voice, language, and vocalisations, rather than creating any player comfort. The musical cue for combat adapts itself from the environmental cue, taking the instrumental timbres into consideration and using the leitmotif. Combat cues do adapt to the kingdom Geralt is in, for example using male voices as opposed to solely female, but they expand on the game's already established leitmotifs as opposed to creating completely new cues.

### ***Celticisms and the Other: The Skellige Isles***

Velen's wilderness environment only uses sung non-diegetic voices in combat to identify Geralt and enemy types, where environmental worldbuilding outside of combat remains voiceless. As the mainland keeps this rule of voices relatively consistent, when voices are treated differently in the world it often identifies an otherness to an environment or dissociation between the environment and Geralt.

The Skellige Isles are an environment locked off to the player until they have completed a certain amount of the story. The Celtic, Nordic, and Viking-influenced environment is also separated from the game's mainland by a sea, cultural differences, and variations in visual and sonic representations. Geralt arrives on the largest isle, Ard Skellig, after being washed up on a beach; Geralt's ship bound to Ard Skellig was ironically attacked by the local inhabitants. This drastic entrance to the islands, after the player witnesses a cutscene of Geralt falling overboard, disconnects Skellige to the mainland as the player is literally abandoned on an alien beach; thankfully the player's trusty steed Roach is waiting patiently for Geralt when he wakes up from being shipwrecked. The components of the island are decisively different to the aesthetics of the mainland that the player is used to. The geographic landscape of Ard Skellig boasts woodlands and mountain ranges, whilst the climate is colder, yet sunnier than Velen; see Figure 19 compared to Figure 18.



Figure 19. *The Witcher 3: Ard Skellig's environment. Screenshot by Author.*

The inhabitants of the island reflect this difference in landscape and climate. The NPCs adopt associations with stereotypical Viking culture as they live in clans and reside in great halls and spend their time pillaging ships and the mainland for food and wealth. The Viking associations also stems from their use of long boats (seen on The Skellige Isle flag in Figure 17) and the worship of Nordic gods, such as Freya. The NPCs' aesthetical design shows an association not unique to Nordic culture, however. The tartan and kilt fashion, alongside Irish accents, places the people of Skellig in a mixture of Northern European culture which symbolises the 'Celtic', here including Scottish and Irish traditions.

This divide in aesthetics and culture from the mainland is enhanced through the soundscape of the archipelago as Przybylowicz explains, in an interview with ICO-Radio, that the Isles reference Celtic, Scottish Gaelic, and Norse cultures and mythologies in the music, using bagpipes, flutes, and Scandinavian folk instrumentation.<sup>18</sup> The deviation from Slavic mythology places Geralt and the player in an unknown world. The soundscape also showcases this shift in culture and mythology by using different instrumentation and vocal styles alongside new environmental and combat cues from the mainland. One of the most notable aspects that isolates the islands is the lack of female voices in the combat themes, and almost lack of voices altogether except for male vocalisations in the Skellige

<sup>18</sup> ICO-Radio, 'The Witcher 3 Komponist Im Interview', 2015 <<http://www.ico-radio.de/2015/06/show-79-the-witcher-3-komponist-im-interview/>> [accessed 22 August 2017].

combat cue 'The Hunt Begins'. Unlike Geralt's female signifiers on the mainland, however, the male voices are not dominant in the cue, and focus on vocalising phonetics such as 'ee', 'yah', 'keh', 'huh', 'hah'; the vocalisations create the sense that the male voices are reciting an unrecognisable, possibly ancient language. The inclusion of these male voices enhances the exoticisms of the archipelago. The instrumentation also highlights the Celticisms of the islands, as the cue includes new instruments, unheard on the mainland, such as the fiddle, bagpipes, flutes, and the Jew's harp. Simon Nugent categorises this 'authentic' use of Celtic components in historical films of modern cinema:

A constructed idea of Celtic music initially gained traction during the latter half of the twentieth century, differing from the actual 'traditional' musics of the various Celtic lands which are understood to share a common ancestry. Numerous instruments, dance forms, and musical styles are analogous throughout musics of the Celtic peripheries, including the pipes, harp, tin whistle (penny whistle), fiddle, and kindred drums. Melodic ornamentation is one of the defining features of music of the Celtic regions, with slides, mordents, and grace notes being archetypal.<sup>19</sup>

The appropriation of Celtic inspired melodic riffs can be heard through the wider tonal intervals and ornamentation that Nugent highlights. The construction of a stereotypical Celtic landscape is also created through the composer's use of these instrumentations that are analogous to the Celtic.

### ***'The Fields of Ard Skellig'***

The soundscape focuses on creating a sense of otherness on the Skellige Isles, through an appropriation of the Celtic. When the player is outside of human civilisations, exploring the wilderness, the non-diegetic cue 'The Fields of Ard Skellig' is triggered. The instrumentation of the cue does not differ from the mainland as much as the combat cues, but its inclusion of a soprano voice provides its otherness. Although there is a female vocalist, her voice is not there to represent Geralt but to emphasise the beauty of the islands. This distinction between female voices is created through several factors: integration of the voice into the exploration theme; the inclusion of a solo voice, rather than multiple voices; the lack of vocalisations; the use of Gaelic language; and the use of fewer instruments so that the vocal performance can be performed with softer dynamics. Example 7 illustrates the fewer instruments that are used to accompany the voice compared to the

---

<sup>19</sup> Simon Nugent, 'Celtic music and Hollywood cinema: Representation, stereotype, and affect', in Cook, Kolassa, and Whittaker, *Recomposing the Past*, pp. 109–10.

mainland, with rhythmic cues on the viola and lute identifying the leitmotifs heard throughout the isles.

**'The Fields of Ard Skellig'**

*J* = 92

The musical score is arranged in three systems, each with three staves: Soprano (S.), Viola (Vla.), and Lute (Lt.).

- System 1 (Measures 1-9):** The Soprano part is silent. The Viola part begins with a melody marked *mf* and includes a fermata over the final measure. The Lute part provides a rhythmic accompaniment marked *mp*.
- System 2 (Measures 10-17):** The Soprano part remains silent. The Viola part continues the melody from measure 10, marked *mf*. The Lute part continues its accompaniment.
- System 3 (Measures 18-25):** The Soprano part begins with the lyrics: "Fhir a' Bhà-ta 'Stric mi se-all-tainn on". The Viola part continues with a steady accompaniment. The Lute part continues with its accompaniment, marked *mp*.
- System 4 (Measures 26-33):** The Soprano part continues with the lyrics: "chnoc as deir - de gach àit' an téid thu". The Viola part remains silent. The Lute part continues with its accompaniment.

Example 7. *The Witcher 3: The opening and first verse of 'The Fields of Ard Skellig'*. The score includes the melody and rhythm on the viola and lute, and the solo female voice. Transcription by Author.

The female voice is significant in its use outside of combat cues. The voice's use of language identifies an environment that is less dangerous than Velen. The Skellige Isles are still full of enemies, which are a higher-level and difficulty because the environment is initially locked off to the player, but the enemy types are mythological creatures as opposed to war-environment monsters. This is shown through the vocal line's quieter dynamics, slower tempo, and longer notation with fewer accents. As opposed to the mixed use of English and Slavic languages in Velen, Novigrad, and Oxenfurt, 'The Fields of Ard Skellig' incorporates the Scottish Gaelic language of the traditional, real-world song 'Fear a' Bhàta' ('The Boatman').<sup>20</sup> The lyrics, as shown in Table 8, express the love of the singer for a young fisherman, waiting patiently on a hill overlooking the sea and awaiting his return.

Scottish Gaelic Translation	English Translation
Fhir a' bhàta	O Boatman
'S tric mi sealltainn on chnoc as deirde	I often look from the highest hill
gach àit' an téid thu	Wherever you go
Dh'fheuch am faic mi fear a' bhàta	To try and see the boatman
gach àit' an téid thu	Wherever you go
Fhir a' bhàta	O Boatman
'S tric mi sealltainn on chnoc as àirde	I often look from the highest hill

Table 8. *The Witcher 3: Lyrical Translation of 'Fear a' Bhàta'. The Gaelic lyrics are used in 'The Fields of Ard Skellig'.*

The cue resemblances tropes of Hollywood Celticism, such as Enya's 'May it be' in the ending credits of *The Lord of the Rings: The Fellowship of the Ring*. The mordent ornamentations accompany the main rhythm of the cue in the vocals and viola, attempting to portray its Celticism.<sup>21</sup> Example 7 shows the mordents and dotted rhythms used in the vocal line that are typical of these implied Celticisms. The rising line of 'Fhir a', moving from an A to D with a quick stop on C, is used in the traditional song alongside the grace notes which highlight the improvisatory nature of traditional folk music. The movement in the notes, the use of ornamentations, and long lyrical verses remove 'The Fields of Ard Skellig' from being stagnant, unlike the static behaviour of the musical cues in

<sup>20</sup> 'Fear a' Bhata' was written at the end of the 18<sup>th</sup> century by Sine NicFhionnlaigh.

<sup>21</sup> The combination of cultures can be seen in the song's similarity to Scandinavian popular song, which is identified in the high, yet soft, soprano singing style. Nightwish, a Finnish metal band, uses these soprano female voices alongside mordents and instrumentation, such as bagpipes.



Velen; the freedom of the music removes Ard Skellig from the war that is occurring on the mainland. There are some differences between 'The Fields of Ard Skellig, and 'Fear a' Bhàta', however. The mordents within Przybyłowicz's adaptation are not present in the original song, as the traditional piece does not use these stereotypical ornamentations to identify with the Celtic. 'The Fields of Ard Skellig' uses these wrongly 'authentic' signifiers to symbolise the differences between the clans and the mainland, but alongside this it also shows variations within itself. The other five islands of the archipelago are smaller than Ard Skellig, however they differ slightly in climate, clan priorities, and present more quests for the player; the islands often feel completely isolated from Ard Skellig at this part of the game due to the lack of unity between clans. To worldbuild this separation of the islands' people, the non-diegetic cues differ slightly for each island.

The Faroe Isle, the smallest isle of the archipelago (most likely named after the Faroe Islands between Scotland and Iceland), consists of two villages, one belonging to Clan Dimun, and the other taken over by bandits which the player can destroy when visiting the island. In the soundscape of Faroe, the melodic lines on the viola and harp in 'The Fields of Ard Skellig' can be heard. This cue, however, lacks the solo voice of the Ard Skellig but contains an altered version of its leitmotifs. Faroe's environmental cue, 'The Tree Where We Sat Once', adapts the melodic and rhythmic ideas of the main island to a faster speed with more ornamentation and a slight deviance from the melodic line; seen in Example 8. The instrumentation remains like Ard Skellig but the roles of each instrument changes in the cue, and more instruments are introduced; a harp accompanies the main melody on the viola, which is doubled with a flute.



Example 8. *The Witcher 3: 'The Tree Where We Sat Once'*. Transcription by Author.

'The Tree Where We Sat Once' both aligns Faroe with the archipelago through its adaptation of the Ard Skellig's melodic and rhythmic ideas, but the instrumentation used situations it in Celtic culture more so than Ard Skellig. Synthesised bagpipes are used alongside the flute to worldbuild the wilder, colder environment that has a slightly different culture amongst its clan residents. Instead of the solo soprano voice, a choir is used sporadically within the cue. The choir only vocalises the same rhythmic line as the string, the lack of language showing the greater dangers and wilderness of the isle. The cue is shorter than 'The Fields of Ard Skellig' due to the shorter exploration times of the

small island, but the high female voices in the choir accentuate the beauty of the island, as snow falls almost constantly when in the environment.

### **Characterisations: Geralt and Himself**

As I have discovered, comradery in video games is important to player identity as it gives a way for the player to interact with other characters. Interestingly, due to the nature of witchers, Geralt works alone and the player does not have access to a party of other characters. It is made a point of humour in the Netflix adaptation of *The Witcher* that Geralt likes to speak to his horse, and subsequently suffers when followed by the bard Jaskier who sings of his glory. Geralt's vocalised signifiers are thus contained to the non-diegetic space. When the player needs to be aware that Geralt does not belong in a place, his female voices are lost as we saw on The Skellige Isles – instead, the female voices of the environment itself take over, highlighting Geralt's lack of fit.

### ***Lack of Geralt's Voices: 'Ladies of The Wood'***

One of the first decision making quests the player must engage with introduces the dense swamp environment within the heart of Velen, Crookback Bog, and the characters known as the ladies of the wood. The area presents itself as a twisted fairy-tale, an environment where the local children were sent by the parents to follow a candy trail to the bog. The children are managed by a woman named 'Gran', who keeps them fed for the ladies of the wood, the witches of Crookback bog, a storyline which takes similarities to the tale of Hansel and Gretel. The witches offer to aid Geralt in his search for Ciri but only if he does a favour for them first. This part of the game narrative gives the player two choices to resolve the witch's quest that both have large impacts on the greater narrative. The quest is for Geralt to kill a devil creature that resides in a tree, a being that has been transformed by the witches. The player either does as the witches ask or grants the being's wish to be set free. Neither choice is received well, as at least one innocent character will die whether Geralt does as the witches ask or not. However, if the player-character goes against the witches then the innocent's death has a greater impact on the narrative.

'Gran' takes Geralt to see the three witches, called 'crones' by Geralt, in a locked silo building. The imagery that is presented to the player is striking, the game's lighting drastically changes and appears dark as the player-character enters the candle lit room. The building exhibits pagan imagery, with hanging herbs, meats, bones and other curiosities that symbolise stereotypical pagan worship;

see Figure 20. The soundscape accompanies the discomforting visuals as the diegetic sounds of the outside environment disappear as the player enters the building. Replacing the natural sounds, such as wind and birdsong, is a looped non-diegetic musical cue performed on a bowed gusli as Geralt speaks to the witches, who are residing within a tapestry, giving no respite from the excerpt whilst the player is in the room.



Figure 20. The Witcher 3: The ladies of the wood's silo residence. The images show the hanging oddities around the portrait and the room, and the tapestry of the witches which is made from human hair. Screenshots by Author

The 'Ladies of the Woods' cue consists of a repeating eight bars, two sections of four bars, by the bowed gusli which presents the theme, accompanied by a drone performed by synthesised male

bass voices; the use of synthesised male voices removes any male agency from the ladies' presence, that they can control. The gusli creates a drone affect in a high frequency, alongside a quiet drum beat and lute accompaniment. The cue is a chaotic performance to represent the witches, as high squeaks, or instrumental 'mistakes', are heard randomly across the cue to keep the player in the uncomfortable shoes of Geralt. The rhythm is static and unnerving as the crotchet movements remain within a similar pattern and a single octave. The jump from the first note to its fifth (G to D), shown in Example 9, is not resolved by any stepwise motion in the opposite direction or returning to the third, putting an emphasis on the minim and empowering the uncomfortable cue. The centring of the music, around the note D, follows a similar pattern to the lack of growth in the melody and sounds of Velen and creating a similar unease. The static and unnatural movements of the cue signify the ancientness of the witches, and their disturbing nature, which James Cook considers an experience that is 'unsettling and claustrophobic.'<sup>22</sup>

Example 9. *The Witcher 3: Vocal bass, bowed gusli, and contrabass drones and melody of 'Ladies of the Wood'.*  
Transcription by Author.

The cue lacks the identifiers of Geralt, the consistent female vocals of the rural areas of Velen, isolating the player as the voices do not reach Crookback Bog. As the 'Ladies of the Wood' cue begins, and loops during conversation, the player hears the death of Geralt's signifiers repeatedly. At the beginning of the cue muted vocalised screams of female voices, rising in pitch, are heard but abruptly cut off by the unnatural synthesised drone of male voices. This suffocation of Geralt's signifiers removes his agency from the room. The witches clearly control the environment and attempt to control the feminine of this environment by smothering Geralt's female voices, speaking through 'Gran', and making derogatory comments towards the women of the local communities. Alongside dominating any feminine agency in the environment, the witches' control over men is shown through an optional commentary about the tapestry's creation out of the weaving of boys' hair, supposedly the first hair they cut when becoming men. This domination is forced on Geralt in

<sup>22</sup> Cook, 'Playing with the Past in the Imagined Middle Ages', p. 232.

conversation as the ladies' taunt Geralt about his sexuality repeatedly and remove the player from their usual dominant and agent role.

After the player has completed interacting with the witches during this quest line, the tapestry can no longer be accessed (along with the 'Ladies of the Wood' cue) because the witches have manifested into their physical characters. The swamp environment, inside and outside the silo, now uses aspects of the original cue to keep the anxiety of the player locked within this environment; the player will always be reminded of the decisions they made on this quest line whenever they return to this environment. The witches remain in control of this environment, aided by their unique, and effectively orientalist, Welsh accents in a geographical area otherwise dominated by West-country voices. The ladies remain full of agency as unique characters and enemies not found elsewhere. Cook highlights the use of microtonal inflections in the soundscape to portray the disturbing otherness of the ladies.<sup>23</sup> Although Cook analyses the soundscape of the ladies as their non-western and non-northern behaviours through this microtonal disturbance, he does not discuss their Welsh accents in a game world where this accent is rare amongst the other Scottish, Irish, and British inflections.

The lack of dominance of Welsh accents in the fantasy environment is due to its relationship with ancient languages. Elvish speech, known as Elder speech, often uses words closely linked to the Welsh language. Within the world of *The Witcher 3*, Elves are persecuted by the northern realms, in a very similar manner to witchers. Elves and other non-human races, such as Dwarves, use a language called Elder Speech (Hen Llinge) as opposed to common speech (English). Elder Speech is a language created from Welsh and Northern European languages.<sup>24</sup> The ancient language, mostly unused by humans, contains references to Welsh when defining the witcher fortress in Kaer Morhen. Kaer Morhen, and its kingdom location Kaedwen, refers to the Welsh words caer (fortress), môr (sea), hen (old), and wen (white). As Kaer Morhen may translate to a 'fortress of the old sea' in Welsh, in elder speech it translates to 'Keep of the Elder Sea'. When the player visits Kaer Morhen for the first time, its joint Welsh and Elder speech translation is prevalent in its aesthetic as a, mostly derelict, old white fortress where witchers were trained; see Figure 21.

---

<sup>23</sup> Cook, 'Sonic Medievalism', p. 232.

<sup>24</sup> For the Netflix television series, Elder Speech was latinised and translated and can be found on The Witcher wikia: Witcher Wiki, 'Elder Speech', *Fandom* <[https://witcher.fandom.com/wiki/Elder\\_Speech](https://witcher.fandom.com/wiki/Elder_Speech)> [accessed 1 June 2020].





Figure 21. *The Witcher 3: Kaer Morhen, the white fortress. Screenshot by Author.*

Elder speech takes influence not only from Welsh but also Irish and Scottish Gaelic. Where *môr* is considered the word for sea, the similar Scottish Gaelic and Irish word *muir* also refers to sea, and the Irish word *mór*, translating to large, can be attributed to the *mor* in *Morhen*. The accumulation of languages to symbolise the aesthetics of Kaer Morhen, an ancient, large white fortress which used to be adjacent to a sea, shrouds the language in a sense of the historic. As Nugent identified, the constructed otherness in film can often be identified by the analogous instrumentation used. The exoticism of Kaer Morhen is highlighted in its environmental soundscape as it incorporates the tin whistle to perform its melodic line.

The limited use of Elder Speech, because of the persecution of non-human races, mimics the loss of the Welsh language after the Middle Ages. The anglicisation of Wales, after an era of prosperity in the Middle Ages, eroded the use of the Welsh language by the landed gentry as England attempted to unify Wales, as Janet Davies discusses:

The process took at least 250 years and was virtually complete by the late eighteenth century. It had profound consequences. Welsh culture, which had been essentially aristocratic, came into the guardianship of the peasantry and the 'middling sort of people' – craftsmen, artisans and the lower clergy. As the inhabitants of the gentry houses ceased to speak Welsh, the system of patronage that had maintained the Welsh poets over the

centuries collapsed, and the standardized Welsh they had jealously defended was in peril of deteriorating into an assortment of mutually unintelligible dialects.<sup>25</sup>

This collapse of Welsh culture and language, as highlighted by Davies, is mirrored in the breakdown of witcher culture. As humans began to unify the northern realms, building over ruined Elven cities, and persecuting non-human races and culture, so too did witchers begin to lose their status. This persecution saw the reduction of non-human races in the northern realms, children sent to Kaer Morhen to be trained as witchers, and the use of Elder Speech; the human language of common tongue became common place.

Returning to the ‘Ladies of the Wood’, their southern Welsh accents situate them as ancient beings, the accent possibly a remnant of inflection from what would have been their original language. However, unlike the Elder Speech as it would have been used by Elves, the Welsh accent by the ladies highlights their dangerousness. The Welsh language’s reference in Elder Speech identifies the use of magic, and the ladies’ abilities to use magic it dangerous for the player. Where sorcery in the world is conducted by witches, and is not inherently used for evil, witchcraft itself is attributed to uncomfortable and grotesque rituals.

### ***Lack of Geralt’s Voices: The Human Amongst the Inhuman***

It is not only monsters, witches, ancient accents, language, and microtonal inflections that suffocate Geralt’s signifiers. His vocal manifestations are consistent in the wilderness, outside of city environments and Crookback Bog. As Geralt enters the city environments of Novigrad and Oxenfurt, it becomes clear that witchers are not necessarily well received in these environments, nor does Geralt appear to fit into the cities.

---

<sup>25</sup> Janet Davies, *The Welsh Language: A History* (Cardiff: University of Wales Press, 2014), p. 36.



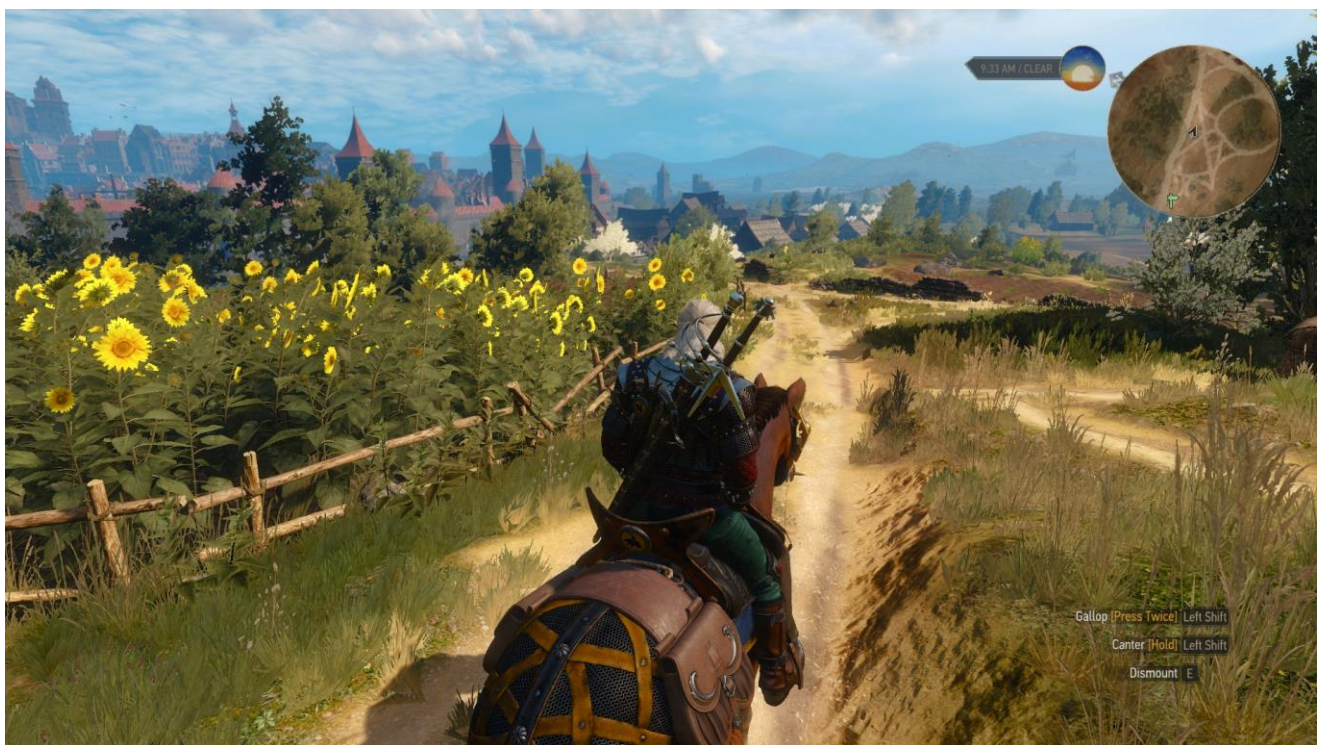


Figure 22. The Witcher 3: The inside and outskirts of Novigrad. Screenshots by Author.

Novigrad is the largest city in the game and resides to the north of Velen. The environment changes as the player starts to move towards Novigrad, as they move away from Velen's marshes and the war-torn environment. The worldbuilding of Novigrad is transitional as the player moves closer to



the city. The outskirts of Novigrad consists of a belt of agricultural land that separates the marshes and pagan imagery of Velen to the medieval city; see Figure 22. Farmland inhabits this area, rather than untamed swamp land, and the soundscape begins to show this transition by reducing the instrumentation used and the anxiety riddled female vocals are diminished. Rhythms are performed by the strings rather than drones, losing the tension from the bow, in order to identify the environment's lack of monsters. The instrumentation itself moves to harps and flutes in the non-diegetic space, performing tonal harmonies. As this happens the imagery become lighter and filled with the brighter yellows of crops and white blossom of spring trees. Rain is less frequent in the outskirts of Novigrad, but the land is still damp and the characters use Velen accents, reminding the player they are still in the same province.

The city almost completely removes itself from the paganistic and pastoral environments of Velen's countryside. The city's architecture, fashion, and soundscape build a truly medieval identity and introduces diegetic performers in the streets. Baroque flutes, lutes, renaissance fiddles, and hammered dulcimers are used to portray the period's style; the soundscape aims to remove Novigrad from the 'exotic' Slavic and folk instruments used in the deep swamps of Velen. The characterisations of Novigrad's inhabitants assumes a wealth divide due to the colourful fashion of the city. This draws parallels from other depictions of medieval cities in media, such as the film and television series of *Robin Hood: Prince of Thieves*,<sup>26</sup> *GoT*, *Merlin*,<sup>27</sup> and *Galavant*,<sup>28</sup> where neo-medieval cities and countryside are divided dramatically. In the drama *Vikings*,<sup>29</sup> the city of Paris distinguishes itself from Scandinavia and England by its almost impenetrable architectural environment and advanced war machines. The visuals inside Paris, compared to the French countryside environments and Scandinavia, consists of vivid colours, and medieval architecture and clothing. The accompanying soundscape changes inside Paris to stereotypical choral music consisting of diegetic male voices alongside catholic worship, as opposed to the non-diegetic pagan female voices used to identify Scandinavia and the Vikings. The audio and visual change between the countryside and city of Paris assumes those that reside outside the city are poor, non-Christian individuals. The change between Velen and the cities of Novigrad and Oxenfurt exhibits these divides of tropes and wealth. Characterisations differ between the pagan west country accents to a northern accent when the player enters Novigrad, and changes to the Queen's English as the player roams the academic city of Oxenfurt. The inclusion of male voices to signify the organised religious

---

<sup>26</sup> Kevin Reynolds, *Robin Hood: Prince of Thieves*, Warner Bros (Michael Kamen, 1991).

<sup>27</sup> Julian Jones, *Merlin*, FremantleMedia (Rob Lane & Rohan Stevenson, 2008-2012).

<sup>28</sup> Dan Fogelman, *Galavant*, Disney-ABC Domestic Television (Alan Menken & Christopher Lennertz & Glenn Slater, 2015-2016).

<sup>29</sup> Michael Hirst, *Vikings* (MGM Television, 2013).

environments, as discussed in Chapter four, is included in *The Witcher*. As the player explores the district of 'Temple Isle' in Novigrad they are introduced to the 'Temple of the Eternal Fire'. This environment introduces a similar worldbuilding and characterisation situation to *Horizon: Zero Dawn* where the visuals and audio suddenly switch to a neo-Christian like setting; the soundscape changes to diegetic male voices and Gregorian chant in this area. The 'Temple of the Eternal Fire' highlights the problematic differential use between paganistic non-diegetic female voices and introducing the male dominated chorus. The male voice is privileged in both the diegetic and non-diegetic space here, as the voice is assumed to embody the powerful Eternal Fire priests.

The change in environment encourages the divide of its people. Although Geralt is not warmly welcomed by the rural villages of Velen, the NPC comments made at Geralt without player interaction are continuously derogatory and aggressive. The guards of Oxenfurt and Novigrad frequently comment towards Geralt with terms such as 'mutant', 'misborn clod', and he is continuously told to 'sod off' with the threat of violence. The relationship between Geralt and the people of the cities results in their characterisations. As with *Vikings*, the environmental soundscape encourages the representation of individuals and cultures, whether poor, wealthy, educated, and more. This dichotomy is indicated in the segregated areas of Novigrad, with various districts showcasing diverse people and cultures: poorer peasantry live in The Bits and Silverton districts; the sorcerers reside in Putrid Grove; harlots live in Lacehalls; merchants exist in the Fish Market; and the social elite and nobles are found in Gildorf and Hierarch Square. The areas of higher class are marked in the use of diegetic music in Gildorf and Hierarch Square, with diegetic street performances occurring within these cultured districts of the city. In Oxenfurt, Geralt's female voices even betray him, and their Slavish identifiers, by singing the cue 'Whispers of Oxenfurt' in a Celtic inspired, Irish/Scottish English:

*We're a' laid idle  
 Wi' the keepin' o' the bairn  
 The lad will nay work and the lass will nay learn  
 The lad will nay work and the lass will nay learn  
 Thou lies ower lang in thy bed  
 Bonny at morn  
 Canny at e'en  
 Bonny at morn  
 Canny at e'en  
 Bonny at morn  
 Thou lies ower lang in thy bed*

*Bonny at morn*

*Thou hinders thy mother at every turn*

The female voice sings through the lyrics without the screams and loud dynamics heard in Velen, emphasising the safety of Oxenfurt; the solo soprano voice is used as a replacement for the guttural cries of the variety of female voices in the swamps. The identification process is thus difficult for Geralt as his voice, demeanor, and social skills are disconnected from the people of the world, monotone in the environment. It is the sounds and NPCs around Geralt that engage the player with his characteristics which can be found in the non-diegetic space.

As part of the main story quest, the player witnesses a live lute performance in the game which is also heard in the district of Hierarch Square, in the Kingfisher Inn where the character Priscilla performs the song 'Wolven Storm'. Geralt encounters Priscilla as he searches for his friend from previous Witcher games, Dandelion. Geralt is led to The Kingfisher Inn by his Dwarven ally Zoltan to meet Priscilla conveniently during her show. The performance occurs in an environment with stylistic lighting and royal red and purple colours taking place on a visually rich stage. The tavern is filled with various characters clothed in expensive and colourful cloth, unlike the plain, tattered, and sometimes non-existent clothing of the peasants of the city. The visuals represent a medieval, almost Shakespearean imagery that indicates a cultural and literate group of NPCs. The player is put in an environment far from Geralt's musical identifiers but, surprisingly, Priscilla's song acts as identifier of Geralt instead.

***The Female Entertainer: Geralt's Redemption***

The player is forced into a quest where they must meet the performer Priscilla in order to find Dandelion to help complete their quest to find Ciri. Voices fill the diegetic space of the tavern, such as conversation, coughing, drinking, and eating sounds, and if Geralt arrives at the Inn before the designated time of the quest, they hear set songs. If Geralt enters the Kingfisher Inn before sunset they can listen to non-vocal performances also heard when playing the in-game card game Gwent, such as 'Another Round for Everyone' and 'The Nightingale'. Diegetic dialogue from tavern patrons remains consistent with the city as patrons do not let Geralt or the player feel at home in this non-rural environment; one patron in particular comments 'You shouldn't have come to the city, you would have been better off staying in the trees.' It is not until Priscilla takes the stage that the tavern

becomes a place of significance for Geralt, whether the patrons of the tavern are aware of the song's source material, being Geralt.

The intrigue of Priscilla's song is that Geralt and Priscilla have never met until this point. It is thus a surprise to Geralt, if not the player, that Priscilla sings of Geralt and identifies his likeness and love interests. 'Wolven Storm' refers to Geralt's signifiers as he is known as the 'white wolf' across the game's world. The tavern turns into a performance space as Priscilla takes the stage, begins to tune her lute, and the tavern sounds previously heard are diminished to make space for her song; Priscilla's dominance both off-stage and on-stage is shown in Figure 23. James Cook highlights the performance as an important diegetic moment:

Perhaps the most significant diegetic moment, however, is the song by the Trobaritz Priscilla. The audio and visuals are surprisingly well matched, and the tuning of the lute adds emphasis to the fact of live performance. As a video cut-scene, this is one of few moments in the game where the player has no ability to affect their environs and must simply watch and listen.<sup>30</sup>

As Cook proposes, the player loses gameplay functions. The player is locked into the space (although they can press a button to skip stanzas) and cannot control movement of Geralt or the camera. The jests at Geralt cease in this moment, and Geralt is removed from his isolation in the tavern, becoming a member of the audience, normalising Geralt in the space. The player too becomes an audience member in this moment, although they are privy to different views during the song, including different audience members and close ups on Priscilla and her lute. This reduction of player agency by the female performer bares similarities to the soundscape of the theatre in *NieR: Automata*, outlined in Chapter four. As Simone apprehends the soundscape from the player's narrative to identifying the theatre and her story, Priscilla takes both the soundscape and the player's ability to move away from them. The uniqueness of the performance space, also seen in *Transistor* and *Divinity: Original Sin II*, tends to provide the female performer with agency during this scene. Even if the player triggers this event, such as Red's hum function in *Transistor*, the soundscape highlights the story of the performer.

---

<sup>30</sup> Cook, 'Playing with the Past in the Imagined Middle Ages'.

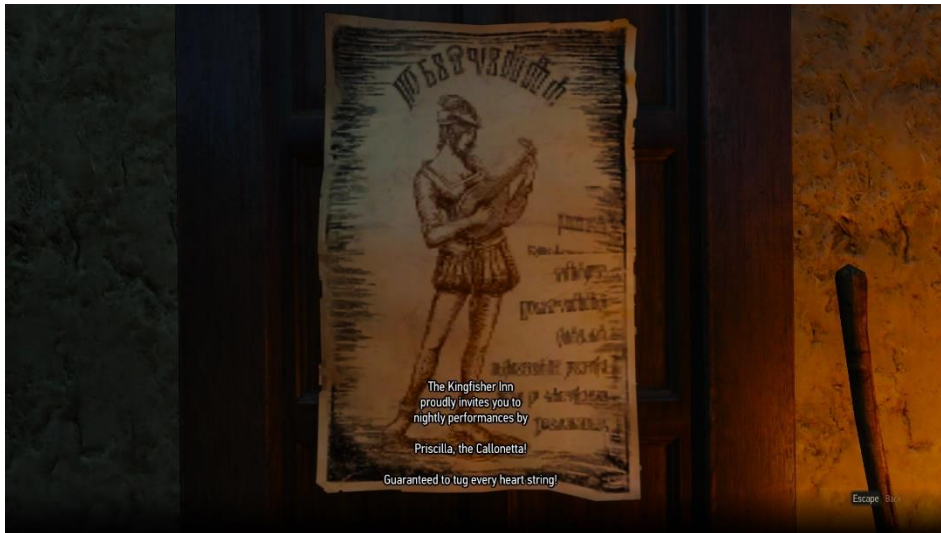


Figure 23. *The Witcher 3: Priscilla's poster as entertainer, Priscilla and her lute during performance, and Geralt looking reminiscent during Yennefer's descriptors. Screenshots by the Author.*

Whilst the player is listening to the performance, the characters in the tavern are highlighted by the camera, showing couples kissing and both women and men being brought to tears. There is still an emphasis on the cultural elite in this environment, with most people in the game being aware of witchers, but it is the elite and cultured people of the game that are aware of Geralt and his past triumphs and failings. The performance of 'Wolven Storm' is a space for the player to relax and delve into the lore of the witcher. Priscilla's song, although highlights her agency as a performer, changes from identifying her story to that of Geralt's past; this difference to the usual highlight of the performers story, as seen in previous case studies, is significant. 'Wolven Storm' speaks of the past narratives of *The Witcher* and *The Witcher 2*, and the continued love story between Geralt and Yennefer, which the player has the option to ignite or destroy later in the game narrative; part of the lyrics sung are as follows:

*You flee my dream come the morning,  
Your scent – berries tart, lilac sweet,  
To dream of raven locks entwisted, stormy,  
Of violet eyes, glistening as you weep.*

These lyrics identify certain traits of Yennefer (her looks and smell) and the convoluted relationship between Geralt and Yennefer as they never remained together. To find Yennefer is the initial goal set by the game and Geralt only manages to contact her once before this performance; Yennefer is also a primary character in Andrzej Sapkowski's books and the Netflix *The Witcher* series. Although Yennefer is a primary character in *The Witcher 3*, and her combined story with Geralt is sung by Priscilla, Yennefer is only mentioned in the previous two games. Yennefer is bound to folk lore, sometimes seen in flashbacks and dreams in the previous games, until the third game where she is given a body. The player can therefore be conflicted by Priscilla's song. It becomes common knowledge for the player (especially if the player has played all three games) that Yennefer and Geralt are bound to each other by a spell. Yennefer is Geralt's first love, and many players believe that she should remain the love interest for the third game, to remain true to the legend set by the game and the books. However, Geralt's love interest across the series is Triss Merigold, a character who can also be a love interest in the third game. Before the first game in the series, Geralt loses his memories, including those of his true love Yennefer, and Triss becomes the love interest. Geralt recovers his memories at the end of the second game, and thus the player has a choice to make

between Triss and Yennefer, Geralt's partner over the three games or the lore built true love of Yennefer.

Yennefer is removed from her place in lore in the third game, only to be put back by Priscilla's song. Priscilla's song specifically characterises Yennefer, likely causing uncertain feelings in players that may have already chosen Triss as their love interest. Priscilla's lyrics regarding Yennefer's scent refers to her signature scent of lilac and gooseberries which Geralt was tracking at the beginning of *The Witcher 3*. The performance has been translated into seven different languages, highlighting its focal point as an identifier of Geralt and a space for players to reminisce or think about their choice between Triss and Yennefer.

The troubadours, and trobairitz, of the witcher diversify the medieval fantasy world and act as Geralt's identifiers where he does not act as a storyteller himself. This inclusion of performers draws clear parallels to the medievalist and fantasy tropes of bards and travelling performers. John Haines describes, 'the singing minstrel is a commonplace in medieval film. More often than not, the minstrel's song takes up a brief but arresting moment in the narrative'.<sup>31</sup> This is true for the cue here, portraying past narratives and potentially the future relationship of Geralt and Yennefer. Ben Winters discusses the emotional connection of characters to audience members, less through non-diegetic themes and more so possibly through emotional diegesis:

[...] there are other issues that may be of only tangential interest to a narratological approach to this scene: namely, how music makes the characters feel, and how an audience reacts to that emotional expressivity. For me, and I suspect for most viewers, the emotional content of the scene's score is far more significant and worthy of exploration than any narrating omniscience or distanced authority apparent in the music's thematic structures.<sup>32</sup>

Unlike the gendered muteness in the entertainer characters seen in *Divinity: Original Sin II* and *Transistor*, Priscilla is not mute when introduced to Geralt and takes control of her vocal agency. However, it is the player that can disrupt Priscilla's performative life, taking her vocal agency. If the player continues on a seemingly innocent quest for Dandelion to renovate his business, Priscilla is attacked, with 'formaldehyde forced in her throat, and an incision made into her voice box'.<sup>33</sup> Priscilla is not mute forever as Dandelion makes a comment later in the game that she is performing again, but at a lower alto frequency and thus losing her own vocal style.

---

<sup>31</sup> Haines, p. 89.

<sup>32</sup> Winters, 'Musical Wallpaper?', pp. 51–52.

<sup>33</sup> 'Priscilla', *Wiki, Witcher*, 2015 <<https://witcher.fandom.com/wiki/Priscilla>> [accessed 30 June 2019].

Priscilla's brutal silencing removes her from the female character and companion space inhabited by Yennefer, Triss, and Ciri. The game's focus on these three female characters provides a distinct space for their stories, and little room for any other female character. Where female agencies are diminished in areas such as the ladies of the wood, the player is provided with gameplay sequences where they control Geralt's ward, Ciri. For a notable amount of the narrative, Geralt and the player are searching for Ciri. During some of Ciri's story, the player switches control to Ciri and guides her through, mostly, linear gameplay and narratives. As the player has Ciri as their player-character they are presented with a unique and agent female space. Ciri and Geralt are joined by destiny and the soundscape represents this through Ciri's combat theme, 'Hunt or Be Hunted'. The combat theme uses the female signifiers of Geralt as they vocalise over the yayli tanbur and hammered dulcimer, which play fast rhythmic patterns to portray the intensity of combat. Although Ciri has her own agency during these unique gameplay sequences, the soundscape is still signifying Geralt because of their shared destiny. With the exception of the ladies of the wood, almost all female characters reflect Geralt and his story, from Priscilla and her 'Wolven Storm' to Ciri's female voices in her combat cue.

*The Witcher 3: Wild Hunt* is known as an RPG video game that, at the time, set the standards for open-world video games. The worldbuilding of the game occurs through both its environmental identifiers and the characterisations of its characters throughout the game world. Geralt is the player-character, but it is the diegetic and non-diegetic uses of voices that can help create a sense of self between the himself and the player. The female voices in the non-diegetic soundscape act as muses for Geralt and, as the player spends more time in the world, it emphasises environments not meant for Geralt when these voices are not present. *The Witcher 3* follows some tropes of cinema and television regarding reflections of identity, as Priscilla the entertainer character becomes muted when attacked off-screen. Priscilla's song is never heard again by the player, Priscilla's agency is in the hands of the player and whether they choose to take on certain quests; the game gives the player choices without them knowing, and thus the player always has consequences to their actions which creates an involvement with each character.

Voice and vocalisations are both used in the game world, vocalisations often referring to Geralt and the monsters he faces, whilst voice is reserved for character characterisations and to highlight human enemies. Alongside this, the varying environmental non-diegetic cues and changing cues for combat highlights where the player is in the game world. When the player fast-travels in *The Witcher 3*, the clear change of visual and aural environmental signifiers helps to identify exactly where the



player is. The changes in vocal signifiers between the mainland and the Skellige isles is a key example of this, as the visuals and aural signifiers clearly show that the land belongs to the clans of Skellige and are separate from the mainland and Geralt. The everchanging accents through different kingdoms, cities, and villages worldbuilds the varying cultures of a 'realistic' and organic world. Language and accents highlight the varying classes and race of the world's people, alongside the player-character. The diverse use of British accents throughout the world effectively identify both the environment and characterisations of NPCs, and enemies such as the ladies of the woods.

## Conclusions and Further Research

Throughout the case studies discussed, I have highlighted the varied and complex uses of recorded voices in RPGs, considering voices as dialogue, song, and vocalisations in the diegetic, non-diegetic, and Stilwell's 'fantastical gap' space of the soundscape. I have shown that voices of an RPG soundscape, whether western or eastern, can effectively aid the worldbuilding process of both environmental features and character identifiers. My research shows that the application of voices as tools to engage the player with the game world requires thoughtful composition and implementation. The inclusion of voices in an RPG, or possibly video games in general, without proper attention may have a negative effect on worldbuilding and player immersion.

As voices are complex, there have been definitions put in place to highlight the varying natural functionalities of voices. I separate voices into two components, voice and vocalisations, which both reflected differing roles as worldbuilding tools in the case studies. Voice acts as a mediator of language-based meaning and, when implemented in the case study worlds, acts as an effective worldbuilding tool for the human. The incorporation of dialogue and language in *NieR: Automata*, and its adoption by machines, is used to make the player question their actions when attacking sentient beings. In *Final Fantasy XV* language is used to identify comradery between characters and engage the player within this comradery as the game progresses. In *The Witcher 3: Wild Hunt*, language is used to differentiate human and non-human enemies, Geralt's signifiers, and environments that hold people and cultures which do not react positively to Geralt.

I define vocalisations as non-language-based meaning that affect feedback and emotional understanding. Vocalisations often encompass vocal manipulations that feedback to the player without language, such as to signify the character's health, providing aural information so that the player does not have to always be engaged with visual UIs. Vocalisations as performance can also be used to provide feedback, as seen within *Transistor* and *Divinity: Original Sin II*, but they are mostly implied as emotional signifiers for the player throughout case studies. The non-diegetic vocalisations in environmental cues, and the combat cue 'Silver for Monsters...', in *The Witcher 3* signify the dangerous environment and enemies within the war-torn environment of Velen. Both simple and complex emotions, such as unease and fear, can be created from this lack of language and identification of the other. What I have found through my analysis of *The Witcher 3* is the conclusion that where voice and its language-based-meaning can portray the human, vocalisations can portray the inhuman. However, as seen through the cue 'The Fields of Ard Skellig', both can imply a sense of otherness.

It is not just the inclusion of these voices that is important however, the implementation of voices through vertical layering and creating reactive voices has been shown to be imperative in RPGs such as *Transistor* and *NieR: Automata*. Where language signifies human qualities, and vocalisations can identify danger and emotion, the loss of voices and performance is shown to inhibit characterisations, such as the loss of Lohse and Red's ability to perform in *Divinity: Original Sin II* and *Transistor*. The transmedia nature of my research also engages with gender studies in multimedia. As I outline in Chapter five, the display of mutism amongst the performative female characters is a trope of female mutism that was developed in opera and film. However, I identified in my research that through the relationship of the voice actor, character, and audience, in Jarman-Iven's 'third-space', video games can begin to remove the gender stereotypes of voices. I discuss in Chapter five that players do not need to identify with all characteristics of a character and can select certain characterisations to identify with. As the voice is used to worldbuild all character characteristics, in video games, rather than focusing on gender, games can surpass the stereotypes and binary genders associated with voices. I conclude that voices are used in games, like *Final Fantasy XV*, to focus on the player's relationship with the character, regardless of their gender; this can provide a space for transgender, and non-binary players to reflect their sense of self.

My model from Chapter two has worked throughout the thesis as an effective visual that highlights the systematic creation of a soundscape. The layout of the model identifies the order in which elements are needed to create an effective, and diverse soundscape for the player. I initially identify the needs of a working video game, starting with the video game diegesis and interfaces, which creates a space for diegesis and outsourced worldbuilding that finally leads to the intricate worldbuilding styles of environmental, characterisations and identification. My model provides a space for voices to be analysed amidst video game worldbuilding components. I found throughout the process of this PhD, and through my own casual play of games, that the most fundamental and effective worldbuilding components are environments and characterisations, whether within the diegesis or outsourced. This distinction provides a clear space for voices to be analysed, and without this initial, visual, display of the soundscape's creation, it would have been difficult to focus my research entirely on analysing the voices of case studies. Without this distinction, I would have spent considerable time in the case study chapters discussing where voices appeared in the soundscape, rather than focusing their effect on the player's engagement with the game world.

My focus on RPGs, and video games which I have both casually and analytically played, has scoped the conclusions of voices as worldbuilding tools. Music and voices in games outside of the RPG genre do vary in their use and effectiveness. There are video games that are voiceless, yet player enjoyment and immersion is not affected, and the soundscape is not lacking. As seen in *Transistor*, a

lack of voices can be an active part of the soundscape itself, especially within minimalist soundscapes, as seen in the building simulation game *Kingdom: Two Crowns*.<sup>1</sup> Even within the RPG genre, the 2018 rogue-light action-RPG *Moonlighter*<sup>2</sup> only uses silent text-boxes to portray the dialogue of the game's NPCs. However, the lack of recorded or synthesised voices is not overly damaging to the worldbuilding of the game because of its retro aesthetic design and indie status. As discussed in Chapter four, the minimalist non-diegetic music in *Final Fantasy XV* and *The Legend of Zelda: Breath of the Wild* allow the player to play the game at their own pace without driving them to any narrative conclusions. *Moonlighter* and *Kingdom: Two Crowns* are games which do not have many narrative or gameplay restrictions, allowing the player to move through the game at any pace they feel. The minimalist soundscapes act to engage the open behaviours of the visuals and gameplay, and even if the player 'completes' the games they do not officially 'end', and the player can continue play after credits.

I focused on analysing RPGs that incorporate dynamic voices that worldbuild a game by providing vocal information and feedback for the player, through accents that reflect environments, adapted voices as direction, comradery, and performance. I have shown that the inclusion of recorded voices, that follow a player's actions and decisions, can engage players in the environments and characters of a world. This importance of recorded voices can be seen in remakes and remasters of RPGs that could not implement recorded voices in their original games because of technological constraints. The remaster, and definitive edition, of the JRPG *Xenoblade Chronicles*<sup>3</sup> has incorporated sung non-diegetic female vocalisations into its soundscape that did not exist in the original game. *Xenoblade Chronicles* was not originally voiceless, however, and used a cast of British voice actors, in its English localisation, as dialogue to create a similar comradery affect between characters, as outlined in my discussion of *Final Fantasy XV*. The decision to introduce vocalisations in the non-diegetic soundscape has provided *Xenoblade Chronicles* with extra environmental and emotional worldbuilders in a comparable manner to *NieR: Automata*. My thesis reflects the increased need for voices in video games as worldbuilding tools in the gaming industry, where possible, and the remastering of soundscapes alongside gameplay and visuals is proof.

---

<sup>1</sup> Noio and Coatsink, *Kingdom: Two Crowns, Stockholm: Raw Fury* (Microsoft Windows, macOS, Linux, PlayStation 4, Xbox One, Nintendo Switch, 2018).

<sup>2</sup> Digital Sun, *Moonlighter, Warsaw: 11 Bit Studios* (Microsoft Windows, macOS, Linux, PlayStation 4, Xbox One, Nintendo Switch, 2018).

<sup>3</sup> Monolith Soft, *Xenoblade Chronicles: Definitive Edition, Kyoto: Nintendo* (Nintendo Switch, 2020).

## Future Directions for my Research

In Chapter two, I discounted FMVs and interactive movies from my research due to the lack of gameplay, physical action, and movement which defines a video game. However, the recent increase in popularity of interactive movies, since *Bandersnatch* and *Erica*, is notable; this is especially owing to audience reactions to the branching paths of *Bandersnatch* and their willingness to try and find every possible ending. I provide a steppingstone to the study of these interactive movies and FMV games through my interdisciplinary study and close link with film literature within Chapters two and three. Interactive movies are by no means linear in narrative, but often their music only reacts to the images on-screen and not player movement.

Where I adapted my definitions of voice and vocalisations to apply to video games, and the study of worldbuilding and RPGs, any further theoretical and empirical study I undertake can test and expand upon my hypotheses and model in this thesis. As I created and used my model to engage with my subjective view of RPGs, I designed it also to have a legacy as a theoretical guide for empirical study which may consult focus groups and gamer responses to voices in video games. The breadth of analysis I undertook across the case studies and literature reviewed can act as the groundwork for any future studies in player and gamer responses to voices, worldbuilding, and identification. The model also considers the positives of video games and its community through its relation to outsourced worldbuilding. The ability for players to engage further with the game world, outside of play and with other players or even the developers, could mean that players can be further immersed in the game when playing. A community effect can be created through outsourced worldbuilding and is something that is relevant in modern times. The COVID-19 outbreak in 2020 coincided with the release of the life and casual simulation game *Animal Crossing: New Horizons* on March 20<sup>th</sup>, 2020. The, already, highly anticipated game was released as countries were entering a lockdown status. Where friends and families could not unite in the flesh, those that had bought the game made a concerted effort to play with their friends and family, meeting up in-game through the embodiment of a personal-avatar; with creative players, this led to mock wedding proposals and ceremonies in the game world, where they could not in reality, as seen in Figure 24. There are no voices in *Animal Crossing: New Horizons*, and players cannot insert their voice into the diegesis. However, players can speak through text-chat bubbles which emerge above their avatar's heads and through gestural reactions, such as clapping.



Figure 24. *Animal Crossing: New Horizons*: Reddit user ashmush posted an image of their in-game mock wedding ceremony.<sup>4</sup>

Social media, such as twitter and Instagram, saw an increase in players, famous personalities, and streamers sharing their play of *Animal Crossing: New Horizons*. Players can become friends in the game and visit each other's islands, functionally socialising whilst being isolated at home during the pandemic. This engagement with a diegesis positively affected the wellbeing and mental health of players who could not leave the house.

As outlined by the model, interfaces can also affect the implementation of voices in video games. Interface worldbuilding is likely a research area that needs considering through empirical studies. The noises of physical interfaces can change depending on a gamers set-up of technology, whether they have a console, PC, headphones, speakers, etc. These studies of interface noises may not particularly include the discussion of voices, but the noise created from interfaces may well effect diegesis worldbuilding and player immersion within the game. This would open the conversation of whether interfaces can negatively affect the engagement with worldbuilding, and the voices outlined in this study. The model also provides a space for player voices to be analysed as part of the diegesis. Player voices, or player related text messages, in the diegesis were not discussed. As shown through Cheng's active research in the game worlds of *Team Fortress 2* and *LTRO*, player voices in

<sup>4</sup> Ashmush, 'My Fiance and I Had to Cancel Our Upcoming Wedding Due to Covid-19, so Our Best Friends Gave Us a Surprise Animal Crossing Wedding Instead', *Reddit*, 2020  
<[https://www.reddit.com/r/AnimalCrossing/comments/fnbgzd/my\\_fiance\\_and\\_i\\_had\\_to\\_cancel\\_our\\_upcoming/](https://www.reddit.com/r/AnimalCrossing/comments/fnbgzd/my_fiance_and_i_had_to_cancel_our_upcoming/)> [accessed 7 June 2020].

the diegesis is a separate, lengthy, and diverse study. Co-operative study can focus on the affect character voices have in competitive online play and esports; do character voices provide useful information to players during a competition, or do they get turned off? The online multiplayer battle arena video game *Dota 2*<sup>5</sup> includes in-game voice lines that can be used to congratulate and encourage a team or discourage and taunt the opposing team; these voice lines can be used in casual play and within the game's esports world championships. Although these voices were not considered here, the psychology literature that I engage with in Chapter three can be used to observe player voices in esports.

Though it may seem that a focus group or interview of gamers may have greatly informed research that includes notions of player identification, I focused on creating the groundwork to understanding voices as worldbuilding tools. Throughout this thesis, I engaged with the study of ludomusicology on more specified theoretical levels and brought together, and expanded on, the disparate and limited literature that focused on voices, worldbuilding, and RPGs. By bringing these components together in this theoretical text, I have created an effective and comprehensive study of voices in video games.

---

<sup>5</sup> Valve, *Dota 2, Bellevue: Valve Corporation* (Microsoft Windows, Linux, OS X, 2013).

# Bibliography

## List of Text Sources

- Anderson, Ethan, 'Animal Crossing Player Puts on a Persona 5 Concert in New Horizons', *Twinfinite*, 13 April 2020 <https://twinfinite.net/2020/04/star-citizen-videos/>
- Ashmush, 'My Fiance and I Had to Cancel Our Upcoming Wedding Due to Covid-19, so Our Best Friends Gave Us a Surprise Animal Crossing Wedding Instead', *Reddit*, 2020 [https://www.reddit.com/r/AnimalCrossing/comments/fnbgzd/my\\_fiance\\_and\\_i\\_had\\_to\\_cancel\\_our\\_upcoming/](https://www.reddit.com/r/AnimalCrossing/comments/fnbgzd/my_fiance_and_i_had_to_cancel_our_upcoming/) [accessed 7 June 2020]
- Barthes, Roland, 'The Grain of the Voice', *Image Music Text*, 1977, 179–89
- Bartle, Richard, 'Virtual Worlds: Why People Play', *Massively Multiplayer Game Development*, 2.2 (2005), 3–18 [https://www.researchgate.net/publication/308073596\\_Virtual\\_worlds\\_Why\\_people\\_play](https://www.researchgate.net/publication/308073596_Virtual_worlds_Why_people_play) [accessed 24 October 2018]
- Calleja, Gordon, *In-Game: From Immersion to Incorporation* (London: The MIT Press, 2011)
- Carroll, Shiloh, *Medievalism in a Song of Ice and Fire and Game of Thrones* (Cambridge: D. S. Brewer, 2018)
- Carter, Chris, 'Review: Final Fantasy XV', *Destructoid*, 28 November 2016 <https://www.destructoid.com/review-final-fantasy-xv-399859.phtml> [accessed 29 March 2019]
- Chattopadhyay, Budhaditya, 'The Auditory Spectacle: Designing Sound for the "Dubbing Era" of Indian Cinema', *The New Soundtrack*, 5.1 (2015), 55–68
- Cheng, William, *Sound Play: Video Games and the Musical Imagination* (New York: Oxford University Press, 2014)
- Chion, Michel, *Audio-Vision: Sound on Screen*, ed. by C Gorbman, trans. by Claudia Gorbman (New York: Columbia University Press, 1994)
- , *The Voice in Cinema*, trans. by Claudia Gorbman (Chichester: Columbia University Press, 1999)
- Collins, Karen, *From Pac-Man to Pop Music: Interactive Audio in Games and New Media* (Aldershot: Ashgate, 2008)
- , *Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design* (Cambridge, Massachusetts: MIT Press, 2008)
- , 'Making Gamers Cry: Mirror Neurons and Embodied Interaction with Game Sound', *Proceedings of AudioMostly 2011*, March, 2011, 39–46 <https://doi.org/10.1145/2095667.2095673>
- Connell, John, and Chris Gibson, *Sound Tracks: Popular Music, Identity and Place* (London: Routledge)
- Cook, James, 'Playing with the Past in the Imagined Middle Ages: Music and Soundscape in Video Game.', *Sounding Out!*, 3 October 2016 <https://soundstudiesblog.com>



- , 'Sonic Medievalism', in *Studies in Medievalism XXIX*, ed. by Karl Fugelso, Politics and Medievalism (Studies), NED-New (Boydell & Brewer, 2020), pp. 217–38  
<https://doi.org/10.2307/j.ctvxhrjqn.18>
- Cook, James, Alexander Kolassa, and Adam Whittaker, eds., *Recomposing the Past: Representations of Early Music on Stage and Screen* (London: Routledge, 2018)
- Corneliussen, Hilde G., and Jill Walker Rettberg, eds., *Digital Culture, Play, and Identity: A World of Warcraft Reader* (London: MIT Press, 2008)
- Coulthard, Lisa, 'The Violence of Silence: Vocal Provocation in the Cinema of Michael Haneke', *Studies in European Cinema*, 9.2–3 (2012), 87–97 [https://doi.org/10.1386/seci.9.2-3.87\\_1](https://doi.org/10.1386/seci.9.2-3.87_1)
- Coyle, Rebecca, ed., *Drawn to Sound: Animation Film Music and Sonicity* (London: Equinox Publishing Ltd, 2010)
- Cross, Ian, Susan Hallam, and Michael Thaut, *The Oxford Handbook of Music Psychology, The Oxford Handbook of Music Psychology*, 2008  
<https://doi.org/10.1093/oxfordhb/9780199298457.001.0001>
- Csikszentmihalyi, Mihaly, *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* (London: Springer, 2014)
- Davies, Janet, *The Welsh Language: A History* (Cardiff: University of Wales Press, 2014)
- Davison, Annette, 'Reviews of Books: Beyond the Soundtrack: Representing Music in Cinema', *Music & Letters*, 90.1 (2008), 113–16
- Day9TV, 'Day[9]'s Day Off - Shadow of the Colossus Remastered P5', *YouTube*, 25 February 2018  
[https://youtu.be/6H\\_mFS6\\_6il?t=790](https://youtu.be/6H_mFS6_6il?t=790)
- Deterding, Sebastian, and José P. Zagal, eds., *Role-Playing Game Studies: A Transmedia Approach* (Oxford: Routledge, 2018)
- Deutsch, Diana, ed., *The Psychology of Music*, 3rd edn (California: Academic Press, 2012)
- Dolar, Mladen, *A Voice and Nothing More* (MIT Press, 2006)
- Donnelly, K.J., ed., *Film Music: Critical Approaches* (New York: Continuum, 2001)
- Donnelly, K.J., William Gibbons, and Neil Lerner, eds., *Music in Video Games: Studying Play* (London: Routledge, 2014)
- Dunn, Leslie, and Nancy Jones, eds., *Embodied Voices: Representing Female Vocality in Western Culture* (Cambridge, New York and Melbourne: Press Syndicate, 1994)
- @emirevans, 'Emi Evans', *Twitter*, 2 April 2018  
<https://twitter.com/emirevans/status/980820204274855936?s=20>
- Escoffier, Nicolas, Jidan Zhong, Annett Schirmer, and Anqi Qiu, 'Emotional Expressions in Voice and Music: Same Code, Same Effect?', *Human Brain Mapping*, 34.8 (2013), 1796–1810  
<https://doi.org/10.1002/hbm.22029>
- Fisher, Linda, 'Feminist Phenomenological Voices', *Continental Philosophy Review*, 43.1 (2010), 83–95
- Florencemachine, 'Florence + The Machine - Stand By Me', *YouTube*, 12 August 2016  
<https://www.youtube.com/watch?v=vv2DSmy3Tro>
- Fritsch, Melanie, 'Game-Music-Performance: Ludomusicology Meets Performance Studies', in

*Ludomusicology Conference, 2017*

- Frühholz, Sascha, Wiebke Trost, and Didier Grandjean, 'The Role of the Medial Temporal Limbic System in Processing Emotions in Voice and Music', *Progress in Neurobiology*, 123 (2014), 1–17 <https://doi.org/10.1016/j.pneurobio.2014.09.003>
- Gamemusic.pl, 'The Witcher 3: Wild Hunt - The World Through The Composer's Eyes', *YouTube*, 21 August 2014 <https://www.youtube.com/watch?v=Q4DnfmhJiGM> [accessed 1 August 2018]
- Gamer, PC, 'The PC Gamer Top 100', *PC Gamer*, 31 August 2018 <https://www.pcgamer.com/the-pc-gamer-top-100-new/12/> [accessed 4 July 2019]
- Gehlawat, Ajay, *Reframing Bollywood: Theories of Popular Hindi Cinema, Reframing Bollywood: Theories of Popular Hindi Cinema* (London: SAGE Publications Ltd, 2010)
- Gibbons, William, and Steven Reale, eds., *Music in the Role-Playing Game: Heroes & Harmonies* (New York, London: Routledge, 2020)
- Goldmark, Daniel, Lawrence Kramer, and Richard D. Leppert, *Beyond the Soundtrack: Representing Music in Cinema* (London: University of California Press, 2007)
- Goosey, Peter, 'Creating Immersion' (unpublished master's thesis, University of Huddersfield, 2017)
- Gorbman, Claudia, 'The Master's Voice', *Film Quarterly*, 68.2 (2014), 8–21  
 ———, *Unheard Melodies: Narrative Film Music* (Bloomington: Indiana University Press, 1987)
- Greene, Liz, 'Speaking, Singing, Screaming: Controlling the Female Voice in American Cinema', *The Soundtrack*, 2.1 (2009), 63–76 <https://doi.org/10.1386/st.2.1.63/1>
- Grover-Friedlander, Michal, *Vocal Apparitions: The Attraction of Cinema to Opera* (Princeton: Princeton University Press, 2005)
- Haines, John, *Music in Films on the Middle Ages: Authenticity vs Fantasy* (Oxford: Routledge, 2014)
- Halfyard, Janet K., ed., *The Music of Fantasy Cinema* (Sheffield: Equinox Publishing Ltd, 2014)
- Hall, Stuart, and Paul du Gay, eds., *Questions of Cultural Identity* (Los Angeles: SAGE Publications, 1996)
- Hanus, Michael D, and Ted M Dickinson, 'The (Faulty) Assumption That Male Players Prefer Male Characters: How Character Desirability and Likability Influence Video Game Purchase Intentions and Enjoyment', *Psychology of Popular Media Culture*, 8.4 (2019), 395–401 <https://doi.org/10.1037/ppm0000191>
- Hart, Casey, 'Getting Into the Game: An Examination of Player Personality Projection in Videogame Avatars', *Game Studies*, 17.2 (2017) <http://gamestudies.org/1702/articles/hart>
- Haworth, Catherine, 'Introduction: Gender, Sexuality, and the Soundtrack', *Music, Sound, and the Moving Image*, 6.2 (2012), 113–35  
 ———, "'Something beneath the Flesh": Music, Gender, and Medical Discourse in the 1940s Female Gothic Film', *Journal of the Society for American Music*, 8.3 (2014), 338–70 <https://doi.org/10.1017/S1752196314000236>
- Horton, Justin, 'The Unheard Voice in the Sound Film', *Cinema Journal*, 52.4 (2013), 3–24
- Hunt, Beth, 'Identity & the Audience: Concerts of Moving Image Soundtracks', in *Constructing the Moving Image: Identity and the Soundtrack Study Day*, 2019

- ICO-Radio, 'The Witcher 3 Komponist Im Interview', 2015 <http://www.ico-radio.de/2015/06/show-79-the-witcher-3-komponist-im-interview/> [accessed 22 August 2017]
- Igami, Laure, 'Relaxing Video Game Music (with Rain and Thunder Sounds)', *YouTube*, 7 May 2019 <https://www.youtube.com/watch?v=Dl3Bd1rdleE>
- IGN, 'TOP 100 VIDEO GAMES OF ALL TIME', *IGN* <https://uk.ign.com/lists/top-100-games/100> [accessed 4 July 2019]
- ImagineBaggins, '1 Hour of Beautiful Video Game Music (Playlist 1)', *YouTube*, 10 January 2017 <https://www.youtube.com/watch?v=EETFIPEnBh4>
- Ivănescu, Andra, *Popular Music in the Nostalgia Video Game: The Way It Never Sounded*, Palgrave Studies in Audio-Visual Culture (Cham: Palgrave Macmillan, 2019)
- Jarman-Ivens, Freya, *Queer Voices: Technologies, Vocalities, and the Musical Flaw* (Basingstoke: Palgrave Macmillan, 2011)
- Jenkins, Henry, *Textual Poachers: Television Fans and Participatory Culture*. 2nd edn. (London: Taylor & Francis, 2013)
- Joe, Jeongwon, *Opera as Soundtrack* (Burlington, Vt: Ashgate, 2013)
- Jørgensen, Kristine, "What Are Those Grunts and Growls Over There?" (Copenhagen University, 2007)
- , *Gameworld Interfaces* (London: The MIT Press, 2013)
- Juslin, Patrik N., and John A. Sloboda, eds., *Handbook of Music and Emotion: Theory, Research, Applications* (Oxford: Oxford University Press, 2010)
- Kamp, Michel, Tim Summers, and Mark Sweeney, eds., *Ludomusicology* (Sheffield: Equinox Publishing Ltd, 2016)
- Kapralos, Bill, Karen Collins, and Holly Tessler, eds., *The Oxford Handbook of Interactive Audio*, Oxford Handbooks (New York: Oxford University Press, 2014)
- Kassabian, Anahid, *Hearing Film* (London: Routledge, 2001)
- Kinne, Jesse, 'The Solo Female Voice as Destiny Topos in Fantasy Media', in *Music and the Moving Image XII*, 2019
- Kozloff, Sarah, *Invisible Storytellers: Voice-over Narration in American Fiction Film* (London: University of California Press, Ltd., 1988)
- , *Overhearing Film Dialogue* (London: University of California Press, Ltd., 2000)
- Lackey, 'Harry Potter - Hedwig's Theme || Symphonic Metal Cover', *YouTube*, 31 October 2017 [https://www.youtube.com/watch?v=z7fbcNW\\_2pg](https://www.youtube.com/watch?v=z7fbcNW_2pg) [accessed 9 June 2020]
- Laing, Heather, *The Gendered Score: Music in 1940s Melodrama and the Woman's Film* (Aldershot: Ashgate, 2007)
- Lash, Dominic, 'Film as Film in Twenty-First Century: The Function and Evaluation of Diegesis and Disruption', *Screen*, 58.2 (2017), 163–79
- Leimbacher, Irina, 'Hearing Voice(s): Experiments with Documentary Listening', *Documentary Audibilities*, 39.3 (2017), 292–318
- Loni, Deepali Y, and Shaila Subbaraman, 'Robust Singer Identification of Indian Playback Singers',

- Audio, Speech, and Music Processing*, 2019.10 (2019), 1–14 <https://doi.org/10.1186/s13636-019-0153-0>
- Macpherson, Ben, 'A Voice and so Much More (or When Bodies Say Things That Words Cannot)', *Studies in Musical Theatre*, 6.1 (2012), 43–57
- Marcus, Kenneth H., 'Mexican Folk Music and Theater in Early Twentieth-Century Southern California: The Ramona Pageant and the Mexican Players', *Journal of the Society for American Music*, 9.1 (2015), 26–60 <https://doi.org/https://doi-org.libaccess.hud.ac.uk/10.1017/S1752196314000534>
- Marshall, Bill, and Robynn Stilwell, eds., *Musicals: Hollywood and Beyond* (Exeter: Intellect Ltd, 2000)
- Martin, Crystle, 'Video Games, Identity, and the Constellation of Information', *Bulletin of Science, Technology & Society*, 32.5 (2012), 384–92
- McAlpine, Kenneth B., *Bits and Pieces: A History of Chiptunes* (New York: Oxford University Press, 2019)
- Meizel, Katherine, 'A Powerful Voice: Investigating Vocality and Identity', *Voice and Speech Review*, 7.1 (2011), 267–74 <https://doi.org/10.1080/23268263.2011.10739551>
- Mera, Miguel, Ronald Sadoff, and Ben Winters, eds., *The Routledge Companion to Screen Music and Sound* (Oxon: Routledge, 2017) <https://doi.org/10.4324/9781315681047>
- Mercer, Matthew, *Critical Role, Twitch, YouTube* (Geek & Sundry, Critical Role Productions LLC, 2015- )
- Mithen, Steven, *The Singing Neanderthals: The Origins of Music, Language, Mind and Body* (London: Weidenfeld & Nicolson, 2006)
- Moayeri, Lily, 'Woody Jackson and Vox Studios - Mixonline', *MIX*, 11 January 2019 <https://www.mixonline.com/recording/woody-jackson-and-vox-studios-red-dead-redemption-2> [accessed 12 April 2019]
- Mollaghan, Aimee, 'An Audio-Visual Gallivant: Psychogeographical Soundscapes in the Films of Andrew Kötting', *The Soundtrack*, 3.2 (2010), 125–30 [https://doi.org/10.1386/st.3.2.125\\_1](https://doi.org/10.1386/st.3.2.125_1)
- Montgomery, Colleen, 'Pixarticulation', *Music, Sound, and the Moving Image*, 10.1 (2016), 1–23 <https://doi.org/10.3828/msmi.2016.1>
- Muriel, Daniel, and Garry Crawford, *Video Games As Culture: Considering the Role and Importance of Video Games in Contemporary Society* (London: Routledge, 2018)
- Nettl, Bruno, *Folk Music in the United States: An Introduction*, 3rd edn (Detroit: Wayne State University Press, 1976)
- Newell, Gabe, 'Steam' (Bellevue: Valve Corporation, 2003)
- Newman, James, *Videogames*, 2nd edn (London: Routledge, 2013)
- Novak, Jelena, 'Voices beyond Corporeality: Towards the Prosthetic Body in Opera', *Studies in Musical Theatre*, 6.1 (2012), 73–88 [https://doi.org/libaccess.hud.ac.uk/10.1386/smt.6.1.73\\_1](https://doi.org/libaccess.hud.ac.uk/10.1386/smt.6.1.73_1)
- Okabe, Keiichi, 'One Year Later, Composer Keiichi Okabe Looks Back at Nier:Automata', *Playstation.Blog*, 7 March 2018 <https://blog.us.playstation.com/2018/03/07/one-year-later-composer-keiichi-okabe-looks-back-at-nierautomata/> [accessed 4 July 2018]
- Ombler, Mat, 'Doom Eternal's Bungled Soundtrack Highlights the Mistreatment of Video Game

- Music', VG247, 22 April 2020 [https://www.vg247.com/2020/04/22/doom-eternal-soundtrack-bad-mixing/?utm\\_medium=social&utm\\_source=twitter](https://www.vg247.com/2020/04/22/doom-eternal-soundtrack-bad-mixing/?utm_medium=social&utm_source=twitter)
- Peterson, Jon, *Playing at the World: A History of Simulating Wars, People and Fantastic Adventures, from Chess to Role-Playing Games* (San Diego: Unreason Press, 2012)
- Pfister, Linda, and Sabiha Ghellal, 'Exploring the Influence of Non-Diegetic and Dieget', in *Proceedings of the 30th Australian Conference on Computer-Human Interaction* (New York, 2018), pp. 490–94
- Phillips, Winifred, *A Composer's Guide to Game Music* (London: The MIT Press, 2014)
- Polygon, 'The 500 Best Games of All Time: 100-1', *Polygon*, 1 December 2017 <https://www.polygon.com/features/2017/12/1/16707720/the-500-best-games-of-all-time-100-1> [accessed 4 July 2019]
- 'Priscilla', *Wiki, Witcher*, 2015 <https://witcher.fandom.com/wiki/Priscilla> [accessed 30 June 2019]
- Pullen, Kristen, *Like a Natural Woman: Spectacular Female Performance in Classical Hollywood* (New Brunswick, New Jersey, and London: Rutgers University Press, 2014)
- Purcell, John, *Dialogue Editing for Motion Pictures : A Guide to the Invisible Art* (Focal Press, 2014)
- Quillfeldt, Thomas, 'Interview: Talos Principle Composer on Humanity in the Digital Realm', *Laced Records*, 11 October 2017 <https://www.lacedrecords.co/blogs/news/interview-talos-principle-composer-on-humanity-in-the-digital-realm> [accessed 9 July 2018]
- Rain, Marcelus Castle, 'Keiichi Okabe & Takahisa Taura Interview (NieR Automata - Pax East 2017)', *YouTube*, 16 April 2017 [https://youtu.be/\\_CmYrYhtvGs?t=1m53s](https://youtu.be/_CmYrYhtvGs?t=1m53s) [accessed 4 July 2018]
- Reeder, Chandler, 'How Music Shapes Houses and the Domestic Space as Vessels Towards a Romanticized Reality in Lady Bird (2017)', in *Music and the Moving Image XII*, 2019
- Reggy, 'Glixel: Interview with Composer Meguro and Artist Soejima About the Development of Persona 5', *Persona Central*, 8 May 2017 <https://personacentral.com/glixel-interview-composer-meguro-artist-soejima-persona-5/> [accessed 4 July 2019]
- Reid, George, 'Chiptune: The Ludomusical Shaping of Identity', *The Computer Games Journal*, 2018, 279–90 <https://doi.org/https://doi.org/10.1007/s40869-018-0070-y>
- Roberts, Maddy Shaw, 'Harry Potter Soundtrack: "Hedwig's Theme" and Everything to Know about the Film... - Classic FM', *Classic FM*, 12 March 2020 <https://www.classicfm.com/discover-music/periods-genres/film-tv/harry-potter-soundtrack-hedwig-theme-john-williams/> [accessed 9 June 2020]
- Samuels, David, Louise Meintjes, Ana Maria Ochoa, and Thomas Porcello, 'Soundscapes: Toward a Sounded Anthropology', *Annual Review of Anthropology*, 39 (2010), 329–45 <https://doi.org/https://doi.org/10.1146/annurev-anthro-022510-132230>
- Schlichter, Annette, 'Do Voices Matter? Vocality, Materiality, Gender Performativity', *Body & Society*, 17.1 (2011), 31–52 <http://journals.sagepub.com/doi/10.1177/1357034X10394669>
- Schreier, Jason, 'The No Man's Sky Hype Dilemma', *Kotaku*, 18 August 2016 <https://kotaku.com/the-no-mans-sky-hype-dilemma-1785416931> [accessed 7 November 2018]
- Silverman, Kaja, *The Acoustic Mirror: The Female Voice in Psychoanalysis and Cinema* (Bloomington: Indiana University Press, 1988)
- Simpson, Campbell, 'We Talked To The People That Made Horizon Zero Dawn Sound So Good |

- Kotaku Australia', 20 April 2017
- Sjogren, Britta, *Into the Vortex: Female Voice and Paradox in Film* (Urbana and Chicago: University of Illinois Press, 2006)
- Smith, Jennifer, 'A Beautiful Song: The Adaptive Music of NieR: Automata', *Medium*, 4 April 2019 <https://medium.com/the-sound-of-ai/a-beautiful-song-the-adaptive-music-of-nier-automata-4b29e5d7f2f3>
- , 'The Musical Parameters of Immersion and Flow' (unpublished master's thesis, University of Huddersfield, 2016)
- Sterling, Jim, 'Demolishing The Excuses Made For "AAA" Publishers' Exploitative Greed (The Jimquisition)', *YouTube*, 15 October 2018 <https://www.youtube.com/watch?v=ycz3U1HdfJo> [accessed 7 November 2018]
- Summers, Tim, 'Opera Scenes in Video Games: Hitmen, Divas and Wagner's Werewolves', *Cambridge Opera Journal*, 29.3 (2018), 253–86
- , *Understanding Video Game Music* (Cambridge: Cambridge University Press, 2016)
- Team, Wizards RPG, *Player's Handbook* (Uxbridge: Wizards of the Coast, 2014)
- The Witcher, 'The Witcher 3: Wild Hunt || Creating the Sound', *YouTube*, 30 August 2016 <https://www.youtube.com/watch?v=g3B2QHIT1Q4> [accessed 4 September 2017]
- Thompson, William Forde, Lena Quinto, Frank A. Russo, and Sandra E. Trehub, 'A Comparison of the McGurk Effect for Spoken and Sung Syllables', *Attention, Perception & Psychophysics*, 72.6 (2010), 1450–54
- Till, Nicholas, ed., *The Cambridge Companion to Opera Studies*, Cambridge Companions to Music (Cambridge: Cambridge University Press, 2012)
- Tones, John, 'Los 100 Mejores Videojuegos de La Historia', *GQ*, 17 October 2018 <https://www.revistagq.com/noticias/tecnologia/galerias/los-100-mejores-videojuegos-de-la-historia/8951?pagina=5> [accessed 4 July 2019]
- U/omnialexmusic, 'Translated Lyrics from APOCALYPSIS NOCTIS [Spoiler]', *Reddit*, 2018 [https://www.reddit.com/r/FFXV/comments/98wpi9/translated\\_lyrics\\_from\\_apocalypsis\\_noctis\\_spoiler/](https://www.reddit.com/r/FFXV/comments/98wpi9/translated_lyrics_from_apocalypsis_noctis_spoiler/) [accessed 28 March 2019]
- Waggoner, Zach, *My Avatar, My Self: Identity in Video Role-Playing Games* (Jefferson: McFarland & Company, Incorporated Publishers, 2009)
- Weiner, Robert G., B. Lynn Whitfield, and Jack Becker, eds., *James Bond in World and Popular Culture: The Films Are Not Enough*, 2nd edn (Newcastle upon Tyne: Cambridge Scholars Publishing, 2011)
- Wiki, Witcher, 'Elder Speech', *Fandom* <[https://witcher.fandom.com/wiki/Elder\\_Speech](https://witcher.fandom.com/wiki/Elder_Speech)> [accessed 1 June 2020]
- Wilde, Tyler, 'The Anatomy of Hype: How No Man's Sky Became the Best and Worst Game Ever', *PC Gamer*, 15 August 2016 <https://www.pcgamer.com/uk/the-anatomy-of-hype-how-no-mans-sky-became-the-best-and-worst-game-ever/> [accessed 7 November 2018]
- Wildt, Lars de, "'Things Greater than Thou": Post-Apocalyptic Religion in Games', *Religions*, 9.6 (2018), 1–20 <https://doi.org/doi:10.3390/rel9060169>
- Wilkins, Heidi, *Talkies, Road Movies and Chick Flicks: Gender, Genre and Film Sound in American*



*Cinema* (Edinburgh: Edinburgh University Press, 2016)

Winters, Ben, 'Musical Wallpaper? Towards an Appreciation of Non-Narrating Music in Film', *Music, Sound, and the Moving Image*, 6.1 (2012), 39–54 <<https://doi.org/doi:10.3828/msmi.2012.5>>

———, 'The Non-Diegetic Fallacy: Film, Music, and Narrative Space', *Music & Letters*, 91.2 (2010), 224–44

Wolf, Tim, Nicholas D. Bowman, John A. Velez, and Johannes Breuer, 'Once Upon a Game: Exploring Video Game Nostalgia and Its Impact on Well-Being', *Psychology of Popular Media Culture*, 2018, 1–18 <https://doi.org/10.1037/ppm0000208>

'Worst FF Combat? - Final Fantasy XV - Giant Bomb', *Giant Bomb*, 2017  
<https://www.giantbomb.com/forums/final-fantasy-xv-84/worst-ff-combat-1804674/> [accessed 27 May 2020]

Yates, David, *Fantastic Beasts and Where to Find Them, Warner Bros. Pictures Heyday Films* (Warner Bros. Pictures, 2016)

Yin-Poole, Wesley, 'No Man's Sky Changed the Video Game Hype Train Forever', *Eurogamer*, 27 December 2016 <https://www.eurogamer.net/articles/2016-12-20-no-mans-sky-changed-the-video-game-hype-train-forever> [accessed 7 November 2018]

## List of Video Game Sources

2K Marin, *BioShock 2, Novato: 2K Games* (Microsoft Windows, PlayStation 3, Xbox 360, Mac OS X, Nintendo Switch, 2010)

Bethesda Game Studios, *Fallout 3, Rockville: Bethesda Softworks* (Microsoft Windows, PlayStation 3, Xbox 360, 2008)

———, *Fallout 4, Rockville: Bethesda Softworks* (Microsoft Windows, PlayStation 4, Xbox One, 2015)

———, *The Elder Scrolls III: Morrowind, Rockville: Bethesda Softworks* (Microsoft Windows, Xbox, 2002)

———, *The Elder Scrolls V: Skyrim, Rockville: Bethesda Softworks* (Microsoft Windows, PlayStation 3, Xbox 360, PlayStation 4, Xbox One, Nintendo Switch, 2011)

Bethesda Softworks, *The Elder Scrolls: Arena, Rockville: Bethesda Softworks* (MS-DOS, 1994)

Bethesda Softworks, *The Elder Scrolls II: Daggerfall, Rockville: Bethesda Softworks* (MS-DOS, 1996)

Blizzard Entertainment, *Warcraft III: Reign of Chaos, Irvine: Blizzard Entertainment* (Microsoft Windows, Classic Mac OS, Mac OS X, 2002)

———, *World of Warcraft, Irvine: Blizzard Entertainment* (Microsoft Windows, macOS, 2004)

Bluepoint Games, *Shadow of the Colossus, San Mateo: Sony Interactive Entertainment* (PlayStation 4, 2018)

CD Projekt Red, *Gwent: The Witcher Card Game, Warsaw: CD Projekt* (Microsoft Windows, PlayStation 4, Xbox One, iOS, Android, 2018)

———, *The Witcher, New York: Atari* (Microsoft Windows, OS X, 2007)

———, *The Witcher 2: Assassins of Kings, Warsaw: CD Projekt* (Microsoft Windows, Xbox 360, OS X,

- Linux, 2011)
- , *The Witcher 3: Wild Hunt*, Warsaw: CD Projekt (Microsoft Windows, PlayStation 4, Xbox One, Nintendo Switch, 2015)
- Croteam, *The Talos Principle*, Austin: Devolver Digital (Microsoft Windows, Linux, OS X, Android, PlayStation 4, iOS, XBOX One, 2014)
- Digital Sun, *Moonlighter*, Warsaw: 11 Bit Studios (Microsoft Windows, macOS, Linux, PlayStation 4, Xbox One, Nintendo Switch, 2018)
- Dontnod Entertainment, *Life Is Strange*, Tokyo: Square Enix (Microsoft Windows, PlayStation 3, PlayStation 4, Xbox 360, Xbox One, iOS, Android, 2015)
- EA Redwood Shores, *James Bond 007: From Russia with Love*, Redwood City: Electronic Arts (GameCube, PlayStation 2, Xbox, PlayStation Portable, 2005)
- Expansive Worlds, *TheHunter: Call of the Wild*, Stockholm: Avalanche Studios Group (Microsoft Windows, PlayStation 4, Xbox One, 2017)
- Firaxis Games, *Sid Meier's Civilization IV*, Novato: 2K Games, Austin: Aspyr (Microsoft Windows, Mac OS X, 2005)
- , *Sid Meier's Civilization VI*, Novato: 2K Games (Microsoft Windows, macOS, Linux, iOS, Nintendo Switch, PlayStation 4, Xbox One, 2016)
- Flavourworks, *Erica*, San Mateo: Sony Interactive Entertainment (PlayStation 4, 2019)
- FromSoftware, *Dark Souls*, Tokyo: Namco Bandai Games (Playstation 3, Xbox 360, Microsoft Windows, PlayStation 4, Xbox One, Nintendo Switch, 2011)
- Game Arts, and Japan Art Media, *Lunar: Silver Star Story Complete*, Tokyo: Kadokawa Shoten (Sega Saturn, PlayStation, Microsoft Windows, iOS, 1996)
- Game Freak, *Pokémon: Let's Go, Pikachu! And Let's Go, Eevee!*, Tokyo: The Pokémon Company (Nintendo Switch, 2018)
- , *Pokémon Yellow*, Kyoto: Nintendo (Game Boy, Game Boy Color, 2000)
- Giant Squid Studios, *ABZÛ*, Milan: 505 Games (Microsoft Windows, PlayStation 4, Xbox One, Nintendo Switch, 2016)
- Griptonite games, KnowWonder, Argonaut Games, Westlake Entertainment, and Warthog Games Limited, *Harry Potter and the Philosopher's Stone*, Redwood City: Electronic Arts (Game Boy Color, Game Boy Advance, Microsoft Windows, PlayStation, Mac OS X, GameCube, PlayStation 2. Xbox, 2001)
- Hangar 13, *Mafia III*, Novato: 2K Games (Microsoft Windows, PlayStation 4, Xbox One, macOS, 2016)
- Hello Games, *No Man's Sky*, Guildford: Hello Games (PlayStation 4, Microsoft Windows, Xbox One, 2016)
- Human Head Studios, *The Quiet Man*, Tokyo: Square Enix (Microsoft Windows, PlayStation 4, 2018)
- Id Software, *Doom Eternal*, Rockville: Bethesda Softworks (Microsoft Windows, PlayStation 4, Stadia, Xbox One, Nintendo Switch, 2020)
- Intelligent Systems, *Paper Mario: The Thousand-Year Door*, Kyoto: Nintendo (GameCube, 2004)
- IO Interactive, *Hitman: Blood Money*, London: Eidos Interactive (Microsoft Windows, PlayStation 2,



Xbox, Xbox 360, 2006)

Iridium Studios, *There Came an Echo*, Los Angeles: Iridium Studios (Microsoft Windows, PlayStation 4, 2015)

Konami Computer Entertainment Japan, *Metal Gear Solid 3: Snake Eater*, Tokyo: Konami (PlayStation 2, 2004)

Konami Computer Entertainment Tokyo, *Silent Hill*, Tokyo: Konami (PlayStation, 1999)

Larian Studios, *Divinity: Original Sin*, Tokyo: Bandai Namco Entertainment (Microsoft Windows, PlayStation 4, Xbox One, Linux, OS X, 2014)

———, *Divinity: Original Sin II*, Tokyo: Bandai Namco Entertainment (Microsoft Windows, PlayStation 4, Xbox One, macOS, Nintendo Switch, 2017)

Level-5, *Ni No Kuni: Wrath of the White Witch*, Tokyo: Namco Bandai Games (PlayStation 3, Microsoft Windows, Nintendo Switch, PlayStation 4, 2013)

———, *Ni No Kuni II: Revenant Kingdom*, Tokyo: Bandai Namco Entertainment (Microsoft Windows, PlayStation 4, 2018)

Level-5, and Square Enix, *Dragon Quest VIII*, Tokyo: Square Enix (PlayStation 2, Android, iOS, Nintendo 3DS, 2004)

Lionhead Studios, *Fable II*, Redmond: Microsoft Game Studios (Xbox 360, 2008)

Maxis, *The Sims 4*, Redwood City: Electronic Arts (Microsoft Windows, macOS, PlayStation 4, Xbox One, 2014)

Media Molecule, *LittleBigPlanet 2*, San Mateo: Sony Computer Entertainment (PlayStation 3, 2011)

Monolith Soft, *Xenoblade Chronicles: Definitive Edition*, Kyoto: Nintendo (Nintendo Switch, 2020)

———, *Xenoblade Chronicles 2*, Kyoto: Nintendo (Nintendo Switch, 2017)

———, *Xenoblade Chronicles X*, Kyoto: Nintendo (Wii U, 2015)

Naughty Dog, *The Last of Us*, San Mateo: Sony Computer Entertainment (PlayStation 3, PlayStation 4, 2013)

———, *The Last of Us Part II*, San Mateo: Sony Interactive Entertainment (PlayStation 4, 2020)

Nintendo EAD, *Super Mario Maker*, Kyoto: Nintendo (Wii U, Nintendo 3DS, 2015)

———, *The Legend of Zelda: Majora's Mask*, Kyoto: Nintendo (Nintendo 64, GameCube, 2000)

———, *The Legend of Zelda: Ocarina of Time*, Kyoto: Nintendo (Nintendo 64, GameCube, iQue Player, 1998)

Nintendo EPD, *Animal Crossing: New Horizons*, Kyoto: Nintendo (Nintendo Switch, 2020)

———, *The Legend of Zelda: Breath of the Wild*, Kyoto: Nintendo (Nintendo Switch, Wii U, 2017)

Noio, and Coatsink, *Kingdom: Two Crowns*, Stockholm: Raw Fury (Microsoft Windows, macOS, Linux, PlayStation 4, Xbox One, Nintendo Switch, 2018)

Obsidian Entertainment, *Fallout: New Vegas*, Rockville: Bethesda Softworks (Microsoft Windows, PlayStation 3, Xbox 360, 2010)

- , *Pillars of Eternity*, Stockholm: Paradox Interactive & Austin: Versus Evil (Microsoft Windows, Linux, OS X, PlayStation 4, Xbox One, Nintendo Switch, 2015)
- On-Line, Sierra, *The Beast Within: A Gabriel Knight Mystery*, Los Angeles: Sierra On-Line (MS-DOS, Macintosh, Microsoft Windows, 1995)
- P-Studio, *Persona 5*, Höfen: Deep Silver (PlayStation 3, PlayStation 4, 2017)
- Platinum Games, *Nier: Automata*, Tokyo: Square Enix (PlayStation 4, Microsoft Windows, Xbox One, 2017)
- Rare, *GoldenEye 007*, Kyoto: Nintendo (Nintendo 64, 1997)
- Rockstar North, *Grand Theft Auto: San Andreas*, New York: Rockstar Games (PlayStation 2, Microsoft Windows, Xbox, Mac OS X, iOS, Android, Windows Phone, Fire OS, Xbox 360, PlayStation 3, 2004)
- Rockstar Studios, *Red Dead Redemption 2*, New York: Rockstar Games (PlayStation 4, Xbox One, Microsoft Windows, Stadia, 2018)
- Ryu Ga Gotoku Studio, *Yakuza 0*, Tokyo: Sega (PlayStation 4, Microsoft Windows, Xbox One, 2017)
- Santa Monica Studio, *God of War, San Mateo: Sony Interactive Entertainment* (PlayStation 4, 2018)
- SCE Japan Studio, and Team Ico, *Shadow of the Colossus, San Mateo: Sony Computer Entertainment* (PlayStation 2, 2005)
- SCE Santa Monica Studio, *God of War, San Mateo: Sony Computer Entertainment* (PlayStation 2, PlayStation 3, PlayStation Vita, 2005)
- Square, *Final Fantasy IX*, Tokyo: Square (PlayStation, iOS, Android, Microsoft Windows, PlayStation 4, Nintendo Switch, Xbox One, 2000)
- , *Final Fantasy VI*, Tokyo: Square (Super NES, PlayStation, Game Boy Advance, Nintendo Wii, Android, iOS, Microsoft Windows, Super NES Classic Edition, 1994)
- , *Final Fantasy VII*, San Mateo: Sony Computer Entertainment (PlayStation, Microsoft Windows, iOS, PlayStation 4, Android, Nintendo Switch, Xbox One, 1997)
- , *Parasite Eve*, Tokyo: Square Electronic Arts (PlayStation, 1998)
- Square Enix, and Acquire, *Octopath Traveler*, Tokyo: Square Enix (Nintendo Switch, Microsoft Windows, 2018)
- Square Enix, *Dragon Quest XI: Echoes of an Elusive Age*, Tokyo: Square Enix (PlayStation 4, Microsoft Windows, Nintendo Switch, 2018)
- Square Enix, *Final Fantasy XII: The Zodiac Age*, Tokyo: Square Enix (PlayStation 4, Microsoft Windows, Xbox One, Nintendo Switch, 2017)
- Square Enix 1st Production Department, *Final Fantasy XIII*, Tokyo: Square Enix (PlayStation 3, Xbox 360, Microsoft Windows, iOS, Android, 2010)
- Square Enix Business Division 2, *Final Fantasy XV*, Tokyo: Square Enix (PlayStation 4, Xbox One, Microsoft Windows, Google Stadia, 2016)
- Square Product Development Division 1, *Final Fantasy X*, San Mateo: Sony Computer Entertainment (PlayStation 2, 2002)

- Supergiant Games, *Bastion*, Burbank: Warner Bros. Interactive Entertainment (Xbox 360, Microsoft Windows, Linux, OS X, iOS, PlayStation 4, PlayStation Vita, Xbox One, Nintendo Switch, 2011)
- , *Transistor*, San Francisco: Supergiant Games (Microsoft Windows, Linux, OS X, PlayStation 4, iOS, Nintendo Switch, 2014)
- Supermassive Games, *Until Dawn*, San Mateo: Sony Interactive Entertainment (PlayStation 4, 2015)
- Thatgamecompany, *Journey*, San Mateo: Sony Computer Entertainment (PlayStation 3, PlayStation 4, Microsoft Windows, iOS, 2012)
- Team Ninja, *Nioh*, San Mateo: Sony Interactive Entertainment (PlayStation 4, Microsoft Windows, 2017)
- The Chinese Room, and SCE Santa Monica Studio, *Everybody's Gone to the Rapture*, San Mateo: Sony Computer Entertainment (PlayStation 4, Microsoft Windows, 2015)
- The Quinnsspiracy, *Depression Quest*, Independent: The Quinnsspiracy (Browser, Microsoft Windows, OS X, Linux, 2013)
- Treyarch, Vicarious Visions, Eurocom, and Beenox, *007: Quantum of Solace*, Santa Monica: Activision (Microsoft Windows, Nintendo DS, PlayStation 2, PlayStation 3, Wii, Xbox 360, 2008)
- Toby Fox, *Undertale*, Independent: Toby Fox (Microsoft Windows, Linux, OS X, PlayStation 4, PlayStation Vita, Nintendo Switch, 2015)
- Toys for Bob, *Spyro Reignited Trilogy*, Santa Monica: Activision (Microsoft Windows, PlayStation 4, Xbox One, Nintendo Switch, 2018)
- Turbine, *The Lord of the Rings Online*, Chicago: Midway Games (Microsoft Windows, OS X, 2007)
- Ubisoft, *Assassin's Creed III: Liberation*, Montreuil: Ubisoft (PlayStation Vita, 2012)
- Ubisoft Montpellier, *Rayman Legends: Definitive Edition*, Montreuil: Ubisoft (Nintendo Switch, 2017)
- Ubisoft Montreal, *Assassin's Creed II*, Montreuil: Ubisoft (PlayStation 3, Xbox 360, Microsoft Windows, Mac OS X, PlayStation 4, Xbox One, 2009)
- Valve, *Dota 2*, Bellevue: Valve Corporation (Microsoft Windows, Linux, OS X, 2013)
- , *Team Fortress 2*, Bellevue: Valve Corporation (Microsoft Windows, Xbox 360, PlayStation 3, Mac OS X, Linux, 2007)
- ZA/UM, *Disco Elysium*, Lutsck: ZA/UM (Microsoft Windows, PlayStation 4, Xbox One, 2019)

## List of Film and Television Sources

- Charles Walters, *Summer Stock*, MGM (Loew's, Inc., 1950)
- Chris Columbus, *Harry Potter and the Philosopher's Stone*, Burbank: Warner Bros. Pictures (Warner Bros. Pictures, 2001)
- Dan Fogelman, *Galavant*, Disney–ABC Domestic Television (Alan Menken & Christopher Lennertz & Glenn Slater, 2015-2016)
- David Benioff, and D.B. Weiss, *Game of Thrones*, HBO (Warner Bros. Television Distribution, 2011)

David Slade, *Black Mirror: Bandersnatch* (Netflix, 2018)

James Gunn, *Guardians of the Galaxy*, Burbank: *Marvel Studios* (Walt Disney Studios Motion Pictures, 2014)

Jerry Rees, *The Brave Little Toaster*, *Hyperion Pictures* (David Newman, 1987)

John Lasseter, *Toy Story*, *Walt Disney Pictures Pixar Animation Studios* (Buena Vista Pictures Distribution, 1995)

Julian Jones, *Merlin*, *FremantleMedia* (Rob Lane & Rohan Stevenson, 2008-2012)

Kevin Reynolds, *Robin Hood: Prince of Thieves*, *Warner Bros* (Michael Kamen, 1991)

Lauren Schmidt Hissrich, *The Witcher (TV Series)*, *Los Gatos: Netflix* (Netflix, 2019)

Peter Jackson, *The Hobbit (Film Series)*, *New Line Cinema & MGM* (Warner Bros. Pictures, 2012)

———, *The Lord of the Rings: The Fellowship of the Ring*, *WingNut Films* (New Line Cinema, 2001)

Michael Hirst, *Vikings* (MGM Television, 2013)

Nicholas Ray, *A Woman's Secret* (RKO Radio Pictures, 1949)

Rian Johnson, *Star Wars: The Last Jedi*, *Lucasfilm Ltd.* (Walt Disney Studios Motion Pictures, 2017)

Robert Siodmak, *The Spiral Staircase*, *Vanguard Films* (RKO Radio Pictures, 1946)

Ron Clements, and John Musker, *The Little Mermaid*, *Buena Vista Pictures* (Alan Menken, 1989)

Thomas Anderson, Paul, *The Master*, *Annapurna Pictures* (The Weinstein Company, 2012)

Tom Clegg 'Sharpe' (ITV, 1993-1997)

Sofia Coppola, *Lost in Translation*, *American Zoetrope*, *Elemental Films* (Focus Features, Tohokushinsha Film, 2003)