



**How do the soundtracks enhance immersion in computer games?—
A case study of 《NieR : Automata》**

Master Thesis

Center for Applied Music Research
Department for Arts and Cultural Sciences

Music for Applied Media

Supervisor

Miguel Kertsman

Supervisor

Thomas Wernbacher

By:

Hua-Chun Fan

2020.10. 22



Abstract

Thema/Theme :

- What is the role of soundtracks for immersion in computer games?
- How do adaptive music and audio design build up the believability of game world?

Name Autor / Name Author : Hua-Chun Fan

Lehrgang/Jahrgang / Course/Year : MA, Music for Applied Media / 2016-2020

Seitenanzahl / Pages : 159

Inhalt / Content :

The purpose of this study was to investigate how soundtracks and audio design can contribute to the establishment of immersion in computer games. Another aim was to find out how the game audio works with the visual can intensify the emotional bonds between the players and the fictional world of games. 《NieR: Automata》, the Japanese action role-playing game, was the subject of this study. The score analysis of selected soundtracks and the related gameplay videos of five main endings (ABCDE) were scrutinised in this study.

The Audio enhancement framework in SCI-model was proposed by this study as the research method, which can divide into three domains : Sensory immersion, Challenged-based immersion, and Imaginary immersion. The concept of Flow was distinguished from immersion in the study. Besides, theories from Michel Chion and Ernst Toch were utilised on inspecting the functionality of audio in multimedia and the music analysis. Finally, the embodied cognition approach was applied to examine imaginary immersion, whose findings suggested that the auditory realm of games is crucial to turns the characters as the player's 'extension of self' to sense and react to the game world.

The results revealed that the immersion in computer games benefits from the 'emotion-inducing' function of soundtracks to make players identify with the scenario and characters in games. Next, it was discovered that spatial audio design combined with adaptive music can effectively elicit the player's empathy with characters, which supports stronger emotional bonds and the believability of game world by providing high sensorimotor contingency.

The principal conclusion was that to evaluate the immersive level of a game is to assess how much it renders the fidelity to aural memory. Lastly, it was found that players would immerse themselves further if the thematic depth is offered via the recursive structure of games.

Betreuer / Supervisor : Thomas Wernbacher



Statutory Declaration

I,**Hua-Chun Fan**.....

born the:**1987, January 6th**..... in:**Taipei City, TAIWAN (R.O.C.)**.....

hereby declare,

1. that I have written my Master Thesis myself, have not used other sources other than the ones stated, and moreover have not used any illegal tools or unfair means,
2. that I have not published my Master Thesis in my domestic or any foreign country in any form to this date, and/or have not used it as an exam paper.
3. that, in case my Thesis concerns my employer or any other external cooperation partner, I have fully informed them about the title, form and content of the Master Thesis, and have his/her permission to include the data and information in my written work.

..**Vienna, 2020, October 22**.....

Place, Date

Signature

Acknowledgement

I'd like to thank all the people whose assistance was a milestone in the completion of this study :

Academic program director, *Miguel Kertsman*, for his inspiring education through the past 4 years. Without his resourceful guidance and encouragements, I wouldn't have been able to scrutinise this study so comprehensively from the audio design to music aesthetics.

My supervisor, *Thomas Wernbacher*, whose insights into immersion not only intrigued me to the in-depth analysis of game studies, enriching the scale of this research, but his support which enabled me to overcome several vital moments during the past year.

Head of Center for Applied Music Research, *Eva Maria Stöcker*, whose generous help on my visa extension made it possible for me to successfully reach my goals in Austria. I'd also like to show my gratitude to Ms. *Andrea Kaufmann*, not only for her continuous assistance on my registration issue, but also for her kindness in storing my keyboard safely in the office from time to time.

I wish to pay my special regards to *Dan Forden*, the audio director at Warner Bros. Games–Chicago. Attending his workshop opened my eyes to the evolving games industry, Wwise community, without losing sight of profound indie games. Additionally, I would like to express my sincere thanks and appreciation to composer *Garry Schyman*, whose recommendation on Toch's book truly broadened my horizons and turned this study into a success.

Further, I would like to gratefully acknowledge *Nancy Wu*, the editor-in-chief at United Digital Publications, and Pro. *Steve Fan* at Greatman Consulting Group, for providing job opportunities during the tough time of COVID-19. I'm also deeply indebted to *Aunt Lee*, my boss *Mr. Huang*, and all the supportive colleagues in *Formosa Vienna*. Without their thoughtfulness and bounty, I wouldn't have survived the summer of 2020 or realised this study in time.

I owe my accomplishments to my genius classmates–*Manual Brandtner*, *Richard Pleil*, and *Ingo Budweg*. It was a privilege to start my music adventure with three open-minded gentlemen, who always provided productive feedbacks on my awkward assignments :) Finally, I wish to pay my special thanks to all the talented musicians and the coordination team of *Tuning-In at mdw*, for being the spark behind my music writings and the new vision influencing the world through music.

Content

Introduction	p.01
1. Research question	p.03
2. Subject : Music and Gameplay	p.03
3. Method : Audio enhancement in SCI-model	p.07
4. Literature Review	p.11
4.1 Terminology of Immersion and Presence	p.11
4.2 Music in multimedia : Film and Game	p.13
5. Sensory immersion	p.18
5.1 <u>Feeling of PRESENCE</u>	p.19
I. Details in world design	p.20
II. Spatial audio	p.27
II-1. Simple 3D and Subjective EQ	p.31
II-2. Interactive reverb : Raycast, K-verb, Occlusion	p.33
II-3. Doppler effect : Flight unit, Emil	p.36
III. Audio-only asset	p.38
5.2 <u>Sensory gratification</u>	p.42
IV. Dynamics	p.42
① Vertical Remixing : Multi-layers, Multi-functions	p.43
② Silence : Seamless, Decisive moments	p.54
V. Appealing audio – 8-bit music, Tone Filter	p.59
6. Challenge-based immersion	p.65
6.1 <u>TEMPO of gameplay</u>	p.68
I. Motor challenges	p.70
II. Cognitive challenges	p.78
6.2 <u>STRUCTURE of gameplay</u>	p.89

7. Imaginary immersion	p.96
7.1 <u>Characters and Events</u>	p.96
I. Emotional responses to sound effects	p.99
II. Practicing empathy via Sounds	p.105
① Music as the extension of Personality : Pascal	p.105
② Silence with Subtlety : Campaigns of Companion characters	p.107
③ Witness as Experience : Mirror neurons & Body schema	p.115
7.2 <u>Setting</u>	p.122
I. World setting	p.123
I-1. Contradictory Aesthetics : family, war, violence	p.123
I-2. Comparison with previous 《Nier》	p.127
I-3. What is 'Nier-like atmosphere'?	p.132
II. Magic circle	p.139
7.3 <u>Story</u>	p.141
I. Overcoming the barriers of immersion by sounds	p.141
II. Supporting Primary & Secondary emotions–Ending A & B	p.143
III. Creating empathy with the avatar–Ending C	p.144
IV. Music as evaluation–Ending D & E	p.145
8. Result and Conclusion	p.148
9. Bibliography	p.151
Appendix	p.157

List of Tables & Figures

Table 1. Nier:Automata Full Soundtrack List in 3 categories	p.06
Table 2. Comparison between Flow and Immersion	p.67
Table 3. Tempo of combat music	p.71
Table 4. Composition structure of <34. Alien Manifestation>	p.83
Table 5. Scenario progression in Gameplay 7	p.91
Table 6. Campaign topics of Companion characters and Enemies	p.107
Figure 1. Overview of Audio Enhancement framework in SCI-model	p.11
Figure 2. Score analysis of < 17. Pascal >	p.22
Figure 3. Score analysis of < 35. Tower >	p.24
Figure 4. Camera angles : side, top, hacking, shooting	p.30
Figure 5. Simple3D DSP diagram	p.32
Figure 6. Subjective EQ	p.32
Figure 7. Raycast system	p.33
Figure 8. Raycast in game	p.34
Figure 9. Absorption coefficient & Location size in game world	p.34
Figure 10. K-verb DSP diagram	p.35
Figure 11. Emil's shop in City Ruins with 2B and 9S	p.38
Figure 12. Score analysis of < 12. Emil's Shop >	p.39
Figure 13. The Yerkes-Dodson Law	p.41
Figure 14. Horizontal re-sequencing–Crossfading scores	p.44
Figure 15. Vertical remixing	p.44
Figure 16. Score analysis of < 2. City Ruins >	p.46
Figure 17. Frequency analysis of < 2. City Ruins >	p.50
Figure 18. Frequency analysis of < 12. Emil's Shop >	p.51

Figure 19. Larger form of music and artworks	p.53
Figure 20-22. Discussion on **** of gameplay 14	p.57
Figure 23. Screen transition of Hacking : normal, process, hacking in	p.59
Figure 24. Tone Filter DSP diagram	p.60
Figure 25. Transition by Crossfading only (without Tone Filter)	p.61
Figure 26. Transition with Tone Filter	p.61
Figure 27. Using Tone Filter for no corresponding 8-bit music	p.62
Figure 28. Score analysis of < 10. Grandma–Destruction >	p.72
Figure 29. Score analysis of < 34. Alien Manifestation >	p.80
Figure 30. Cognitive challenge–transport the package	p.86
Figure 31. Score analysis of < 3. Peaceful Sleep >	p.87
Figure 32. Score analysis of < 20. Copied City >	p.92
Figure 33. Lo-Fi DSP diagram	p.99
Figure 34. Noise & Screen flicker present 2B’s vitality	p.100
Figure 35. Score analysis of < 37. The Sound of the End >	p.102
Figure 36. Pods’ campaign ③ in gameplay 19	p.113
Figure 37. Ending E in gameplay 21	p.114
Figure 38. Sound & Music turn characters into the "extension of self" for players	p.119
Figure 39. Beauvoir’s campaign in gameplay 11	p.121
Figure 40. Score analysis of < 26. Rebirth & Hope >	p.144
Figure 41. 9S’ will at Ending D	p.146
Figure 42. Chorus comes in when supporting messages appear on screen	p.148

Introduction

In recent decades, advancements in technology have made 'immersion' a major research subject, especially in the entertainment industry. More and more game companies emphasise the immersive experiences they provide to players via Virtual Reality headset, such as : *Torn VR* (Aspyr Media, 2018), *Fallout 4 VR* (Bethesda, 2017). Non-VR games, such as : *Metro Exodus* (4A Games, 2019), have a high demand to implement 'at least head tracking support' to gain 'deep immersion'.¹

However, too little attention has been paid to the connection between the audio and immersion. Most recent discussions have focused on the real-time processing technology on visual data, while previous research, for example in psychology or media studies has primarily concentrated on the epistemological distinction between 'immersion' and 'presence'. As highlighted by Sander Huiberts, 'it is often stated that sound is also important for immersion, but a more direct or explicit connection is difficult to find'.²

Although some developers have noticed the potential of binaural technology and integrated it into sound design as a HRTF (Head Related Transfer Function) plugin³, the effectiveness of localisation is disputable and mainly used on shooting games, like the Korean online multiple battle royale game *PlayerUnknown's Battlegrounds* (PUBG).⁴ Besides, the scientific survey on the notion of music, referring to "the language of emotion", has been scarcely reported. 'There seem to be consensus around the crucial need for further research in the area', as Eladhari, Nieuwdorp and Fridenfalk proposed in

¹ Jagneaux, D. (2019, February 14). 5 Reasons Why Metro Exodus VR Would Be Absolutely Amazing. [Web log post] UploadVR. Retrieved from <https://uploadvr.com/5-reasons-metro-exodus-vr/>

² Huiberts, S. (2010). Captivating Sound: the Role of Audio for Immersion in Games (Doctoral dissertation). Available from ResearchGate https://www.researchgate.net/publication/255968332_Captivating_Sound_the_Role_of_Audio_for_Immersion_in_Games

³ HRTF plugin can output the simulated 3-dimension audio effect through stereo headphones.

⁴ z83420123 (VoLTsRiNe). (2018, May 23). [情報] HRTF聲音系統實裝. [Web log post]批踢踢實業坊 > BattleRoyale > PUBG. Retrieved from <https://www.ptt.cc/bbs/BattleRoyale/M.1527073356.A.74F.html>

2004. The investigation on “how different factors in musical structure affect the perceived emotional expressions” is pivotal yet insufficient.⁵

Therefore, this thesis will take a comprehensive approach—not only will it examine how the audio technologies enhance player’s sensory engagement, it will also scrutinise how the music and sound design work together with the visual to intensify the emotional bonds between players, characters and the game world. That is, the immersive state where ‘the player experiences a certain sense of “reality” occurs while playing the game’.⁶

As the winner of “Best Score/Soundtrack” in The Game Award 2017, the “Excellence in Music Score” and the “Excellence in Technical Achievement” of SXSW Gaming Award⁷ 2018, the Japanese role-playing game 《NieR: Automata》 is a good example of how the music, sound design and its narrative intertwine with each other to build up an immersive world of game. Not only did 《NieR: Automata》 reach over 4.5 million sales worldwide⁸, but its original soundtrack album also ranked as Golden Record in May in 2019 by The Recording Industry Association of Japan (RIAJ).⁹ Among other J-POP music, it is the only one album of game soundtrack that has sold over 100,000 records since 2017. Furthermore, the oversea tour of NieR Automata Music Concert in Taiwan all sold out in just 1 minute.¹⁰

Given the huge commercial success and great acclaim for its soundtracks, with this study I aim to evaluate the audio strategies of 《NieR: Automata》 within the framework of immersion model through score and gameplay analysis. Since “computer game” is an

⁵ Eladhari, M., Nieuwdorp, R., Fridenfolk, M. (2006). The Soundtrack of Your Mind : Mind Music—Adaptive Audio for Game Characters. Paper presented at ACE’06: Proceedings of the 2006 ACM SIGCHI international conference on Advance in computer entertainment technology, June14-16, 2006, Hollywood, CA,USA. <https://dl.acm.org/doi/10.1145/1178823.1178887>

⁶ Huiberts, *op. cit.*, p.39

⁷ NieR: Automata (2017 Video Game). Awards. *IMDb*. Retrieved from <https://www.imdb.com/title/tt5238626/awards>

⁸ History. (2020, March 29). *NieR:Automata*. Retrieved from PlatinumGames website <https://www.platinumgames.com/games/nier-automata?age-verified=25e5d8d17e>

⁹ 統計情報. (2019, May). ゴールドディスク認定. Retrieved from RIAJ 一般社団法人日本レコード協会 <https://www.riaj.or.jp/f/data/cert/gd.html#02>

¹⁰ nakedjehuty. (2017, August 28). 「尼爾」不只是致鬱遊戲? 《尼爾:自動人形》製作人與聲優群劇透演出心得. [Web log post] gamebase. Retrieved from <https://www.gamebase.com.tw/news/topic/98869270/>

interactive medium, it is essential to see how the narrative and gaming mechanism interweave, as an immersive system that motivates players to engage in and stays close to their hearts for good.

1. Research question

As such, this paper seeks to address the following questions:

- (1) What is the role of soundtracks for immersion in computer games?
- (2) How do adaptive music and audio design build up the believability of game world?

2. Subject : Music and Gameplay

The subject of this study is the first edition of 《NieR: Automata》, which was initially released in February 2017. The research materials consist of two parts—music scores and gameplay videos. The former are 20 soundtracks selected from 46 tracks of 《NieR: Automata Original Soundtrack》¹¹ and 《NieR : Automata Official Score Book》¹², mainly based on the high frequency of usage and the show-up timing through the gameplay.

The latter are 21 gameplay videos by the professional gamer Shirrako¹³ who has more than 1.12 million subscribers on his Youtube channel, and the first gameplay video –“Prologue” of Nier Automata which has reached over 2.8 millions view since 2017 February.¹⁴ The greatest advantage of Shirrako’s videos is that he doesn’t talk during his

¹¹ Square Enix official website (2017, March). *NieR: Automata Original Soundtrack-Tracklist*. Retrieved from Square Enix. Music <https://www.jp.square-enix.com/music/sem/page/nier/automata/>

¹² Japanese Books. (2017). *NIER: AUTOMATA OFFICIAL SCORE BOOK (PIANO MUSIC COLLECTION)*. Retrieved from Play-Asia <https://www.play-asia.com/nier-automata-official-score-book-piano-music-collection/13/70bfhb>

¹³ Shirrako is a UK-based gamer, whose channel is one of the fastest growing Gaming Channel on YouTube, showcasing game trailers, walkthrough videos, soundtracks in 4K 60FPS and 1080P HD. Retrieved from <https://www.youtube.com/user/Shirrako/about>

¹⁴ Shirrako. (2017, February 22). Nier Automata - Gameplay Walkthrough Part 1 - Prologue (Full Game) PS4 PRO. [Youtube]. *Nier Automata PS4 Gameplay Walkthrough Ps4 PRO 1080p 60fps Full Game Guide (playlist)*. Retrieved from <https://www.youtube.com/watch?v=4MU0yMgu3bQ&list=PLCLLeSTzz6trYrOxS1anPprphUxjuOXBAA&index=2&t=0s>

play.¹⁵ Thus, those high-quality gameplay videos become the best material to see how the audio goes with the visual without any external noise/voice interfering. Due to time limits, this thesis would focus on the English gameplay, including 5 main endings (A, B, C, D, E), while the 21 remaining endings lettered through F to Z and DLC (downloadable content) endings won't be discussed in this thesis.¹⁶ Nevertheless, a few Japanese gameplay clips will also supplement the discourse to verify the correlation between immersion and the whole structure design.

The plot of 《NieR: Automata》 is set in the distant future, 5042A.D. An extra-terrestrial force has unleashed a ruthless army known as “machine lifeform” on Earth, driving mankind into exile on the moon. The Council of Humanity deploys a new unit of android infantry—YoRHa, as the resistance force to reclaim the planet.¹⁷ Players play the role of android 2B, 9S, A2 to fight against enemies in real-time in various environment settings within the game, ranging from shoot 'em ups (bullet-hell)¹⁸ to text adventure¹⁹ segments. Besides the slash action-based combat (with/without weapons)²⁰, a hacking function is also automatically activated when the player plays the character as 9S. Each character is assigned one Pod, a flying robot assistant which can launch customisable gunfire attack as well.

¹⁵ Regarding 'no talk' during the gameplay, there is one message written by Дима Марамыгин, saying 'some games are made to be played in silence. thank you for your silence' below of the 'Prologue' streaming video. This is the most popular message which gained 3045 likes among others. (Retrieved on 2020, October).

¹⁶ These additional endings are triggered by certain actions as the game-over events, which don't influence the narrative progress. Therefore, they are outside the scope of this research. see Wu,Chao Min. (2017, June 2). NieR: Automata Guide: How to get all 26 Endings, including the Best Ending. [Web log post] RPG Site. Retrieved from <https://www.rpgsite.net/feature/5392-nier-automata-guide-how-to-get-all-26-endings-including-the-best-ending>

¹⁷ Summarised from the official plot. Story. (2017, February). *NieR:Automata*. Retrieved from PlatinumGames website <https://www.platinumgames.com/games/nier-automata?age-verified=03d0cebbce>

¹⁸ The “Shoot 'em up” refers to the sub-genre of shooter games where players need to dodge seemingly endless waves of coloured bullets. The term is derived from its Japanese name “Danmaku” (弹幕), meaning “Bullet Hell”. Retrieved from GIANT BOMB <https://www.giantbomb.com/bullet-hell/3015-321/> (Retrieved on 2020, May)

¹⁹ “Text adventure” is a type of adventure game where players use text commands to control characters and proceed the narrative. It's also considered as one type of “interactive fiction” or “interactive narration”. see Schindel, D. (2018, November 12). More than Zork: Four of the Modern Text Adventures. [Web log post] GAMECRATE. Retrieved from <https://www.gamecrate.com/more-zork-four-best-modern-text-adventures/21377>

²⁰ There are 4 types of combat styles, including 40 weapons across all categories. see Weapons. (2020, January 31). *Equipment & Weapons - NieR Automata*. Retrieved from NieR:Automata wiki <https://nierautomata.wiki.fextralife.com/Weapons>

《NieR: Automata》 is an open-world Action Role-Playing Game (ARPG), developed by PlatinumGames and published by Square Enix in 2017. It is a sequel to the video game 《Nier》 (Cavia, 2010) which is the spin-off of the 《Drakengard》 game series. Although the developers have changed, the director Yoko Taro and the music production team—MONACA, Inc., led by composer Keiichi Okabe, remained the same in both Nier games. Therefore, some music pieces of 《NieR: Automata》 were indeed inherited from 《Nier》 but with different arrangements and were recorded in different styles. I will treat them equally as other original pieces to see how the music plays the role of unfolding a new narrative and connecting to the previous one.

Table 1 shows the 46 soundtracks of 《NieR: Automata Original Soundtrack》, presented in the order of official track list with minor adjustment.²¹ I grouped a few varying pieces that share the same composition together and marked them with the same number. For instance, the ending song 《*Weight of the World*》 has four versions—each of them is sung by different vocalists with different instrumentations that correspond to alternative endings of the game. Therefore, I assembled them as one group, marked as 15 for original song (English version), 15-1 for Japanese, 15-2 for Nouveau-FR, and 15-3 for “the End of YoRHa” version. Other similar groups include 《2. *City Ruins*》 《9. *Voice of no Return*》 《11. *Faltering Prayer*》 《14. *Vague Hope*》 and 《21. *Wretched Weaponry*》. There are, however, several pieces which couldn’t be found in the 21 gameplay videos by Shirrako : 《6. *The Colour of Depression*》 and 《11. *Faltering Prayer*》, might show up in other gamers’ playing routes ; these couldn’t be analysed in this paper due to the research limit. Hence, the track list has been reduced to 37 pieces.

Then, I divided them into three categories : location music, combat music, ending and others, which are marked in light blue, green, and noted in purple respectively. Firstly, **Location music** comprises 12 tracks, illustrating ten primary places in game : city ruins, resistance camp, desert zone, amusement park, machine village, forest zone, copied city, abandoned factory, bunker, and tower. Secondly, **Combat music** consists of 11 tracks,

²¹ Square Enix official website (2017, March). *NieR: Automata Original Soundtrack-Tracklist*. Retrieved from Square Enix. Music <https://www.jp.square-enix.com/music/sem/page/nier/automata/>

corresponding to various battle scenarios of different levels : 5, 10, 19, 21, 25, 27, 28, 31, 34, 36, 37. Only one track <21. *Wretched Weaponry*> isn't marked in green, because it also refers to the location—abandoned factory. This multi-functional usage of music is an important feature of 《NieR: Automata》 that would be discussed later on in chapter 5. Thirdly, the music that only appears close to the ending (including key characters' final scenes) or beginning sections is labelled as the category of **Ending and Others**, which has 13 soundtracks.

Table 1. *Nier:Automata Full Soundtrack List in 3 categories: Location, Combat, Ending & Others*

⊕Disc I⊕	⊕Disc II⊕	⊕Disc III⊕
1. Significance–Nothing • Initial start & After ending A	1-1. Significance • Start after ending B & C	26. Rebirth & Hope • Ending A & B
2. City Ruins–Rays of Light • ▪ city ruins	2-1. City Ruins–Shade	27. War & War •
3. Peaceful Sleep ▪ resistance camp	16. End of the Unknown • ▪ Adam final	28. Crumbling Lies–Front
4. Memory of Dust • ▪ desert zone	9-1. Voice of no Return–Normal	29. Widespread Illness
5. Birth of a Wish • ▪ Adam & Eve, Father Servo	17. Pascal ▪ machine village	30. Fortress of Lies ▪ bunker
6. The Colour of Depression	18. Forest Kingdom • ▪ forest zone	14-1. Vague Hope–Spring Rain
7. Amusement Park • ▪ amusement park	19. Dark Colossus–Kaiju (怪獣) •	31. Song of the Ancients–Atonement • • variation of 'NIEr' (former game)
8. A Beautiful Song • ▪ Opera boss (Beauvoir)	20. Copied City ▪ copied city [6] ▪ hack space [20]	32. Blissful Death ▪ hack space [20, A2 memory]
9. Voice of no Return–Guitar ▪ Pascal's philosophy	21. Wretched Weaponry : Medium / Dynamic • ▪ abandoned factory	33. Emil–Despair • variation of 'NIEr' (former game)
10. Grandma–Destruction • variation of 'NIEr' (former game)	22. Possessed by Disease •	11-1. Faltering Prayer–Starry Sky • variation of 'NIEr' (former game)
11. Faltering Prayer–Dawn breeze • variation of 'NIEr' (former game)	23. Broken Heart • Preview after ending B	34. Alien Manifestation

12. Emil's shop ▪ Emil	21-1. Wretched Weaponry : Quiet ▪ abandoned factory [2,7] ▪ hack space of factory [14]	35. The Tower • ▪ hack space [14, 9S memory] ▪ Red Girl boss [20]
13. Treasured Time	24. Mourning ▪ hack space [20] • Ending C (before the monologue of A2)	36. Bipolar nightmare
14. Vague Hope–Cold Rain • Ending A. E (melody only)	25. Dependent Weakling • ▪ Eve final [8,14,15]	37. The Sound of the End •
15. Weight of the World (English version) • • Ending A	15-1. Weight of the World (Kowaretasekainouta 壊レタ世界ノ歌) • Ending B	15-2. Weight of the World Nouveau–FR version • Ending C. D
green = combat music, number-1 = variations, [gameplay] ▪= location ▪=characters, fights •= start & end •= 8-bit version (source: this study)		15-3. Weight of the World–the End of YoRHa • Ending E

Finally, based on the official score book and the frequency of music usage, I selected the most representative soundtracks–20 pieces in total–from these three categories as the main materials for score analysis. 11 tracks, marked in bold, out of the chosen pieces were examined carefully with scores. The selected music is shown in track number as below:

Location music (6 tracks) – **2. 3. 7. 17. 20. 35**

Combat music (7 tracks) – 5. **10. 21. 25. 31. 34. 37**

Ending & others (7 tracks) – 9. **12. 14. 15. 16. 24. 26**

3. Method : Audio enhancement in SCI-model

The SCI-model was proposed by Ermi & Mäyrä as the gameplay experience model with a particular focus on immersion. It comprises three fundamental aspects : sensory immersion, challenge-based immersion and imaginative immersion, abbreviated as SCI-model.²² However, a study by Huiberts found that ‘Ermi and Mäyrä do not provide

²² Ermi, L. and Mäyrä, F. (2005). Fundamental Components of the Fame Experience : Analysing Immersion. [Conference paper] *Changing Views: Worlds in Play. Selected Papers of the 2005 Digital Games Research Association's Second International Conference*, pp. 15–27. Eds. Suzanne de Castell and Jennifer Jenson.

examples on how these dimensions of immersion could be realised in the design of a game, nor in the audio used in it'.²³ So, this paper will use the structure map for game audio design, advised by Huiberts, as the analysing framework which is evolved from SCI-model and other well-summarised relevant researches of immersion in games.

According to the investigation by Huiberts, the SCI-model has the smallest overlap between three categories amongst other definitions of immersion, and its scope can cover the largest aspects of immersion, which is 'transportation into the game world, absorption in the activity and identification with the situation'.²⁴ Furthermore, after a comprehensive review of game audio within the SCI-model, Huiberts confirmed that as an essential component in games, the audio 'can stimulate all three dimensions of immersion by enhancing the **sensory connection**, the feeling of **flow** and the feeling of **empathy** of the player'.²⁵

Nonetheless, as Huiberts intended to present a conceptual structure as a resource for game audio design, he didn't touch upon the 'actual-often very technical and specific-design parameters in relation to immersion', such as : sound frequency, musical chord progressions.²⁶ Therefore, this paper will probe into the audio sphere of «NieR: Automata» by utilising Huiberts' conceptual structure to examine how those audio factors complement each other to make 'the player's sense of actually being in the game world', as Rolling and Morris described.²⁷

The SCI-model was derived from the observation by Ermi & Mäyrä²⁸, presenting that the "audiovisual quality and style", "level of challenge", "imaginary world and fantasy" are the most central attributes of good digital games. Based on this discovery, Huiberts

²³ Huiberts, *op. cit.*, p.50

²⁴ Before he decided to choose SCI-model, Huiberts had inspected many scholars' definitions and classifications on immersion, such as: Dansky & Kane (2006), Brown & Cairns (2004), Adams (2004b), etc. Huiberts, *op. cit.*, p.37-49

²⁵ Huiberts, *op. cit.*, p.118 * The emphasis is added by this study.

²⁶ Huiberts, *op. cit.*, p.119

²⁷ Rollings, A. & Morris, D. (2000). *Game Architecture and Design*. Coriolis Group, LLC. Arizona.

²⁸ Ermi and Mäyrä, *op. cit.*, p.6

delineates each category of SCI-model more clearly with subdivisions. Firstly, he indicates two aspects of sensory immersion : feeling of presence, and the sensory gratification. The former is defined as 'rich and detailed auditory world which absorbs the player', that can be stimulated by the details in world design, spatial audio and the audio-only assets.²⁹ The latter refers to 'high quality auditory output which makes the sensory experience more intense' ; in other words, the audio is regarded as beautiful or impressive. And "sensory gratification" could also be established by sound dynamics, suggesting depth and appealing audio. Since the "depth" is related to spatial design, I will merge the discussion of depth into the "spatial audio" section.

Secondly, Huiberts divides the challenged-based immersion into two facets : the tempo of gameplay and the structure of gameplay. As highlighted by Ermi & Mäyrä, players feel the immersion at its most powerful level 'when one is able to achieve a satisfying balance of challenges and abilities'.³⁰ Therefore, the tempo of gameplay should be studied differently on different kinds of challenges. Two main characteristics of these challenges are found to be **motor skills** (such as reacting rapidly to specific events) and **cognitive challenges** (e.g., strategic planning or logical problem solving) which request the player to reflect upon his/her decisions.³¹ In Huiberts' classification, he included "audio-driven gameplay" and "audio-based gameplay" in his discussion. However, these two types of gameplay won't be studied in this paper because neither of them is relevant to 《NieR: Automata》 . We will only investigate the general structure of gameplay ; that is, how the audio supports the changes and the progress in game, and how 'the music correspond with the "peaks and valleys" of gameplay'.³²

Thirdly, there are three perspectives under the imaginary immersion : character and events, setting, and story. Although Ermi & Mäyrä didn't consider game characters as a

²⁹ Huiberts, *op. cit.*, p.57

³⁰ Ermi and Mäyrä, *op. cit.*, p.8

³¹ Huiberts, *op. cit.*, p.69 *The bold keywords are emphasised by Huiberts in his original paper.

³² Huiberts, *op. cit.*, p.75

fundamental factor contributing to immersion³³, a number of studies have shown that the use of characters is as the driving narrative force³⁴ which can 'create deeper involvement through characters and emotions'.³⁵ Besides, Brown and Cairns pointed out that empathy with the game character is very important during the deepest stage of immersion, Total Immersion.³⁶ As 《NieR: Automata》 is an ARPG which has three main characters, I will adopt Huiberts' framework to inspect yet with minor adjustments. Hence, this paper will primarily focus on **characters'** emotional response to sound effects, while the part of voice acting will be excluded due to the research scope.

Further, in order to clarify my discourse, the analysis of "affects in game" will merge with the discussion of "world setting" since they both address mood induction by music or ambient sounds. Therefore, in the **setting** section, the world setting and the magic circle with audio will be investigated, which Huiberts considered the cornerstone for the building "believability" of a game.

Finally, in the **story** section, not only two perspectives– primary & secondary emotions, music as evaluation–suggested by Huiberts will be reviewed, but two more points would be integrated into discussions as well: overcoming barrier of immersion between different routes and creating empathy with avatars through the progression of alternative story lines. Five main endings (A,B,C,D,E) and its associated soundtracks are the main focus of this part.

Originated from the conceptual map by Huiberts³⁷, the analysis framework of this thesis is visually presented in Figure 1. Generally speaking, the upper part of this chart (i.e, the sensory immersion and the challenge-based immersion) focuses more on the

³³ Because it is obviously that not all games provide the possibility for identification with a game character, the sense of identification with characters is optional, as Huiberts explained. see Huiberts, *op. cit.*, p.82

³⁴ Sheldon, L. (2004) *Character Development and Storytelling for Games*. Boston, MA: Thomson Course Technology.

³⁵ As cited in Jørgensen, K. (2010). Game Characters as Narrative Devices. A Comparative Analysis of Dragon Age: Origins and Mass Effect 2. *Eludamos. Journal for Computer Game Culture*, 2010:4 (2), p. 315-331. Retrieved from <https://www.eludamos.org/index.php/eludamos/article/view/vol4no2-13/198>

³⁶ As cited in Huiberts, *op. cit.*, p.82. see Brown, E. & Cairns, P. (2004). A Grounded Investigation of Game Immersion. [Conference paper]. *Paper presented at Human Factors in Computing Systems*. Vienna, Austria. p.2

³⁷ Huiberts, *op. cit.*, p.102

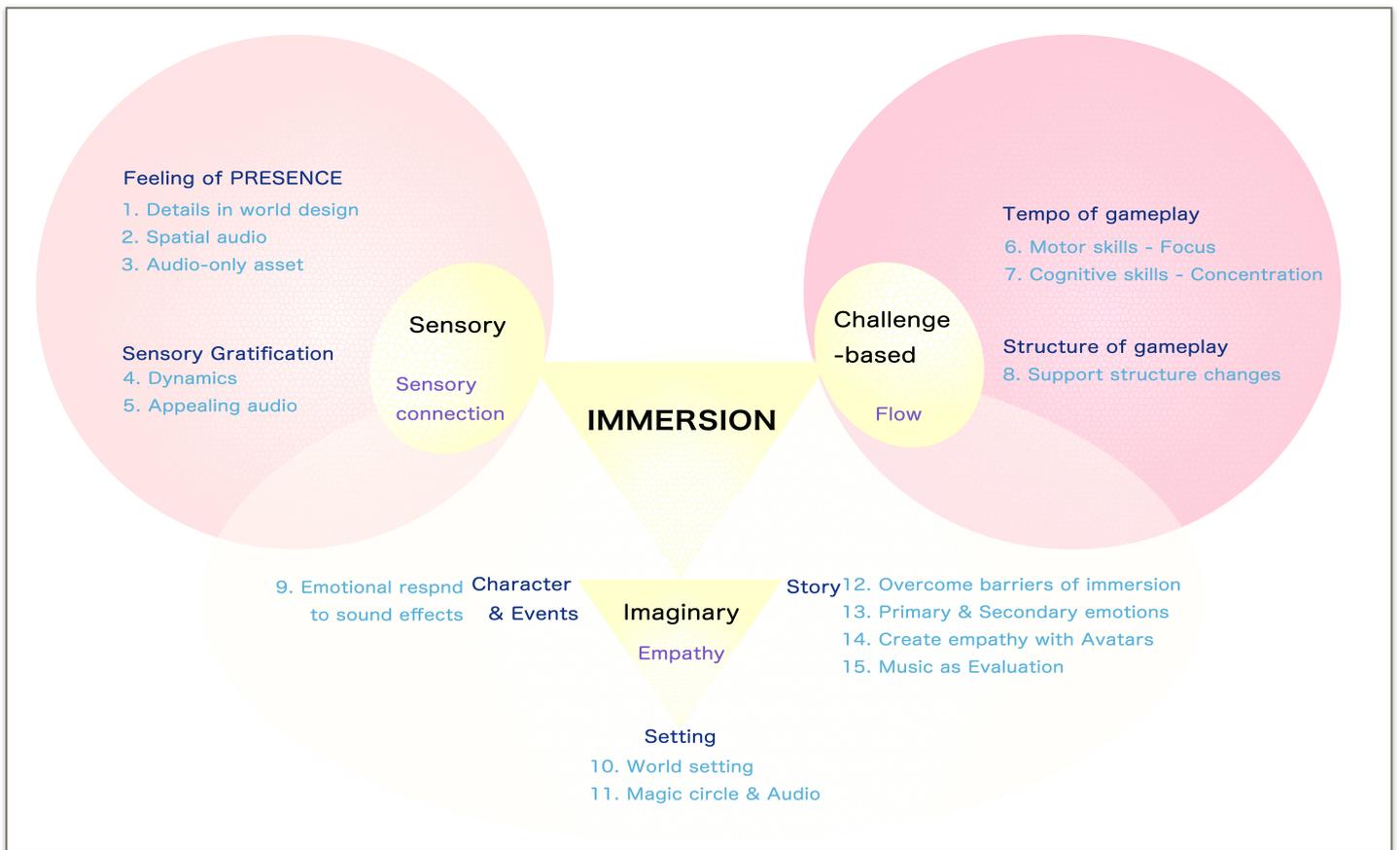


Figure 1. Overview of Audio Enhancement framework in SCI-model

inside world—the game mechanism, aiming to make players believe there is a “reality” which exists and functions well. While the lower part of this chart—the imaginary immersion—lays more emphasis on the emotional bonds, aiming for engaging players’ interest, even making them cry. More detail will be given in later chapters.

4. Literature review

4.1 Terminology of Immersion and Presence

Although there is little consensus on the definition of immersion within the field of game studies³⁸, Huiberts summarised three characteristics of immersion : teleportation, absorption and identification.

I. Teleportation : A feeling of being transported into the game world, or being surrounded by the game world.³⁹ This is supported by Janet Murray’s study which reveals the most accepted definition of immersion : the experience of being transported

³⁸ Huiberts, *op. cit.*, p.37-38

³⁹ Huiberts, *op. cit.*, p.39. More discussion can be seen at McMahan, 2003; Garneau, 2001; Dovey & Kennedy, 2006; Rollings & Morris, 2000; Pine & Gilmore, 1999.

to an elaborately simulated place is pleasurable in itself, regardless of the fantasy content.⁴⁰ According to Lombard and Ditton's 3-type classification on transportation⁴¹, this "teleportation" feature is very prominent in game—"the player can click on where they want to go on a map and they are instantly there".⁴²

II. Absorption : Absorption in activities is often mentioned in the definitions of immersion.⁴³ The player becomes thoroughly engrossed in the game world via the act of playing and reacting with different emotional effects. Huiberts observed that while the "loss of sense of self" is likely to occur in many forms of media consumption, the feeling of being absorbed 'in combination with intense concentration is a distinct property of the experience of game play'.⁴⁴

III. Identification : Immersion comprises the feelings of identification with the situation or a character of a game.⁴⁵ Based on Huiberts' investigation, interviewees generally feel far more satisfied or 'experience more fun' when they 'feel surrounded by the world' and 'being one' with the character. Once the player feels like being immersed and he/she perceives the game world as more realistic, then 'the game and its goals get more depth'.⁴⁶

As highlighted by Mel Slater, a distinction between immersion, presence, involvement, emotional response and degree of interest needs to be made.⁴⁷ It is especially common

⁴⁰ McMahan, Alison. (2003). Immersion, Engagement, and Presence—A Method for Analyzing 3-D Video Games . In Mark J.P. Wolf and Bernard Perron (Ed.) *The Video Game, Theory Reader* (pp.67-86). NY, US : Routledge, Taylor & Francis Group.

⁴¹ Based on Lombard and Ditton's studies, there are three types of transportation: (1) "You are there," (2) "It is here," and (3) "We are together." While the 'teleportation' I refer to here is the second type of transportation, the third type of transportation would be examined further in later chapters. see Lombard, M. & Ditton, T. (1997). At the Heart of it All: The Concept of Presence. *Journal of Computer-Mediated Communication*, 3(2). <https://doi.org/10.1111/j.1083-6101.1997.tb00072.x>

⁴² McMahan, *op. cit.*, p.77

⁴³ Huiberts, *op. cit.*, p.39. More discussions can be seen at Dansky & Kane, 2006; Dovey & Kennedy, 2006; Taylor, 2002.

⁴⁴ *Loc. cit.*

⁴⁵ Huiberts, *op. cit.*, p.39. More discussions can be seen at Brown & Cairns, 2004; Rollings & Morris, 2000; Rollings & Adams, 2003; Taylor, 2002.

⁴⁶ Huiberts, *op. cit.*, p.40-41

⁴⁷ Slater, M. (2003). A Note on Presence Terminology. *Presence Connect* 3 (3), 1-5. Retrieved from <https://scholar.google.com/citations?user=5gGSgcUAAAAJ&hl=en>

to see the term presence used synonymously with immersion, which simply adds more confusion, as observed by McMahan.⁴⁸ The analogy proposed by Slater with colour science can be used by us to separate these two concepts. According to Slater, the distinction between "immersion" and "presence" is like the difference between wavelength distribution and perception of colour. Immersion is more about how the system operates to arouse people's perception, while 'presence is a human reaction to immersion'. 'The more that a system delivers displays (in all sensory modalities) and tracking that preserves fidelity in relation to their equivalent real-world sensory modalities, the more that it is *immersive*', as analysed by Slater.⁴⁹

Hence, the standard of assessing a game's, that is the system's immersive level depends on how much it **preserves the fidelity to reality**. As Slater reminded us, presence is the human response to the displayed system. So, even if the same immersive system is given, different people may experience different levels of presence, and this could be justified through empirical studies. Nevertheless, it's still possible to discuss how well a computer game (the system) can emulate reality through its visual and audio design—to what extent is this "realness" close to its counterpart in the physical world, and can it elicit authentic emotions from players? In order to arouse the "feeling of presence" from players, how do music and sound design contribute to teleportation, absorption and identification—the three features of an immersive system? We'll delve into these questions step by step with examples in the following sessions.

4.2 Music in multimedia: Film and Game

The relationship between music and different media, and its different applications has been widely investigated. Sonnenschein listed four specific functions that classical film scores or incidental music serve in the emotional, technical and aesthetics aspects of film.⁵⁰

⁴⁸ McMahan, *op. cit.*, p.73

⁴⁹ Slater, *op. cit.*, p.1

⁵⁰ Sonnenschein, D. (2001). *SOUND DESIGN: The Expressive Power of Music, Voice, and Sound Effects in Cinema*. Studio City, CA : Michael Wiese Productions. (p.155-156)

① Emotional signifier : Regardless of genre, the music allows us to sense the invisible and inaudible, the spiritual and emotional processes of the characters portrayed. This is consistent with the study by Eladhari, Nieuwdorp and Fridenfalk, which revealed that music can be used to give the player an idea of what a character is like by hearing its affective process.⁵¹

② Continuity : Music can fill the gap either in the image or sound. The sense of continuity is maintained when music is played over spatially discontinuous shots. The research findings by Cohen also identified that music has at least eight functions in multimedia and the “provision of continuity” is the most prominent music gestalt—‘our capacity-limited brain is forced to “go with flow”. Hence, when the sound is organised in time, this organisation—music—helps to connect disparate events in other domains’.⁵²

③ Narrative cueing : According to Sonnenschein, attaching an emotional interpretation to the image, can cue the narrative that is to come, for example, advance knowledge of a threat or a setup for a joke. This is supported by Cohen’s study which discovered that the visual screen is often referred to as two-dimensional⁵³ with music adding the third dimension of emotion. In addition, Cohen noted that ‘communication of emotional meaning or meaning through association (e.g., of epochs, cultures and events) is particularly effective in situations that are ambiguous’.⁵⁴

Besides, Jørgensen reviewed available literature on film studies and game audio research and observed that the “diegetic sound” and “extradiegetic sound” both exist in films and games. The former is what the fictional characters are able to hear whereas the later can’t be heard by the fictional characters, only by the audience, aiming to communicate the mood or drama within the film. However, as highlighted by Jørgensen, the extradiegetic sound often has a different informative role which helps players not

⁵¹ Eladhari, Nieuwdorp, Fridenfalk, *op. cit.*, p.2

⁵² Cohen, A.J. (1999). The Functions of Music in Multimedia: A Cognitive Approach. *Paper presented at ICMPC: Proceedings of the Fifth International Conference on MUSIC PERCEPTION AND COGNITION, August 26-30, 1998, Seoul, Korea.* <https://musicog.discoveryspace.ca/sites/default/files/funcmusicmultimedia.pdf> (p.14)

⁵³ Cohen, *op. cit.*, p.15. See more discussions at Palmar, 1990 ; Rosar, 1994.

⁵⁴ *Loc. cit.*

only feel the mood but evaluate his/her choices of actions in the game world.⁵⁵ The usage of adaptive music in computer games is the most obvious clue that allows players to know when they should pick up the weapons to fight.

④ Narrative unity : Sonnenschein indicated that music structure can aid in the formal unity of the film narrative by utilising its compositional techniques, such as : repetition, variation, and counterpoint, thus supporting the narrative to proceed. This is consistent with Huiberts' findings, in which he proposed sounds of the affect domain, especially music, are very suitable to support time-based processes, 'as they share properties such as rhythm, timing, repetition and progression'.⁵⁶ Both viewpoints correspond to the early studies by Cohen, showing that 'visual and audio stimuli can be described in two dimensions : meaning and structural features. Structural features refer to aspects such as rhythm, contour, or patterning in time'. Even though the relation between the visual and the audio stimuli is not necessary, as Cohen explained, in order to make sense of the world, the human brain defines objects that produce correlated sensory patterns. That is, 'pattern from different modalities that have the same pattern of change, across time are likely to refer to a single object', as indicated by Cohen.⁵⁷

Based on Cohen's investigations, the study by Eladhari, Nieuwdorp and Fridenfalk highlighted several particular functions of music when it is applied to games, including how 'music can be used to direct attention to important features of the screen, to induce mood, to communicate meaning, to further the narrative, to enable the symbolisation of past and future events through the technique of leitmotiv, to heighten the sense of absorption, and to add to the aesthetics'.⁵⁸

Although film music and game music share almost the same functions, there is a significant difference between them—one is unconsciously heard and one is consciously

⁵⁵ Jørgensen, K. (2006). On the Functional Aspects of Computer Game Audio. *Proceedings of the Audio Mostly Conference 2006, Oct 11-12*. Retrieved from <http://bora.uib.no/handle/1956/6734>

⁵⁶ Huiberts, *op. cit.*, p.68

⁵⁷ Cohen, *op. cit.*, p.14-15

⁵⁸ Eladhari, Nieuwdorp, Fridenfalk, *op. cit.*, p.1

listened to. In Sonnenschein's opinion, a good music score for a film is not meant to be heard, 'at least not consciously, normally remaining subordinate to the dialogue and the visuals as the primary narrative vehicles'.⁵⁹ In contrast, the study by Jørgensen indicates that computer games should be regarded as dual ; that they are game systems as well as fictional worlds. That is to say, 'game audio has the overarching role of supporting a user system while also supporting the sense of presence in a fictional world', as Jørgensen clarified.⁶⁰ By providing specific information to the player about the state of the system, game audio (including music) has the role of easing the use of the system by offering players the information of the current usability, mood and atmosphere, orientation, control and identification.⁶¹ Only by doing so can players successfully complete the tasks one after another in the non-linear, hypothetical world of games where players are asked to believe in what they are playing. Film music and game music both attempt to intensify the psychological aspects of the scenario, it is either a linear story or a non-linear narrative. However, in contrast to the film audiences who only need to understand the context, the default for players is to feel the emotions of the scenes, reacting promptly to the changes in games or to make choices, especially in action role-playing games.

Therefore, when it comes to the functionality of music and audio design in games, I propose that the music features, indicated by Sonnenschein, could be reconsidered as—**Emotion signifier, Continuity, Narrative unity, and Narrative cueing for reactions**—in games. Further, the Atmospheric function and the Orientation function identified by Jørgensen should be also included into our analysis of game audio. The former is still regarded as one of the most central functions in games, inherited from the music tradition in films which works as a "mood enhancer" to "emotionally engage the audience" ; in most mainstream games, music is utilised to 'emphasise certain areas, locations and situations', as Jørgensen pointed out.⁶² The latter function is born for the

⁵⁹ Sonnenschein, *op. cit.*, p.156

⁶⁰ Jørgensen (2006), *op. cit.*, p.2

⁶¹ Jørgensen (2006), *op. cit.*, p.1

⁶² Jørgensen (2006), *op. cit.*, p.3

non-linear characteristics of game, especially for the players in an open-world action-role-playing game where they are able to freely explore but must face the menace coming from every angle at the same time. According to Jørgensen, 'the perhaps most obvious orienting function of sound is that it provides information about the presence of objects as well as the direction of sound sources'.⁶³ In general, the **Atmospheric function** focuses more on the implementation of music and soundscape, while the **Orientation function** mainly addresses the application of sound design in games.

Previous neurophysiological and psychological research has suggested that music activates independent brain functions that are separate from verbal and visual domains. That is to say, when 'there is music, more of the brain is active'.⁶⁴ Studies by Cohen discovered that 'increased activity of the brain may increase concentration on the primary focus and filter out distractions such as the frame around a monitor or screen, people in the vicinity, or the pressure from one's chair'.⁶⁵ Namely, music in multimedia lifts up the participants' concentration level, temporarily isolating them from the real world. This finding corresponds to one of the characteristics of immersion I mentioned before : **Absorption**, the feeling of being absorbed 'in combination with intense concentration is a distinct property of the experience of game play'.⁶⁶ The findings by Cohen and Huiberts were supported by Jørgensen's study. She successfully summarised the function of computer game audio that 'aims to combine usability with presence and immersion in the fictional game world and to support games—the user systems—operate smoothly by utilising music and audio design for informative purpose'.⁶⁷

In brief, as Cohen has demonstrated 'multimedia "works" when the multi sensory information is encoded, interpreted and stored such that the information can be acted upon'.⁶⁸ The entire process of receiving the information to forming the cognitive

⁶³ *Loc. cit.*

⁶⁴ Cohen, *op. cit.*, p.16

⁶⁵ Cohen, *op. cit.*, p.17

⁶⁶ Huiberts, *op. cit.*, p.39

⁶⁷ Jørgensen (2006), *op. cit.*, p.5 *The emphasis is added by this study.

⁶⁸ Cohen, *op. cit.*, p.13

understanding all take place in mind, which not only falls within the scope of cognitive psychology but is also greatly influenced by 'how music works within the multimedia context'.⁶⁹ Therefore, I will delve into the presenting context of 《NieR: Automata》 –the imaginary fictional world and the user system–through the Audio Enhancement Framework of SCI-model accordingly (see Figure1). I will begin with the Sensory immersion in chapter 5, later go through the Challenged-based immersion in chapter 6 and the Imaginary immersion in chapter 7. Finally, I will discuss the research results in chapter 8.

5. Sensory immersion

Based on the analysis framework of the SCI-model, the “Feeling of PRESENCE” and the “Sensory gratification” would be examined in this section. As explained previously, presence is a human reaction to immersion. Slater identified that 'presence is about form, the extent to which the unification of simulated sensory data and perceptual processing produces a coherent “place” that you are “in” and in which there may be the potential for you to act'.⁷⁰ In our discussion, the “coherent place” refers to the game, the immersive system which stimulates human sensory responses through its motor action systems. Further, Slater highlighted the need to distinguish presence, involvement, and emotions. He argued that both involvement and emotions are separable from presence. As in many situations of daily routine, one can be present but not involved. Meanwhile, one can be involved with no need to be present in the fiction world, such as watching a soap opera. Because 'being interesting, emotionally captivating, beautiful, fantastic – these are about content, not about the form'.⁷¹

Therefore, **Feeling of PRESENCE** could be seen as a FORM of human response to immersion–either as happening or not, cognitively understanding or physically reacting,

⁶⁹ *Loc. cit.*

⁷⁰ Slater, *op. cit.*, p.2

⁷¹ *Loc. cit.*

while **Sensory gratification** is the CONTENT of how we treat the music and sounds mentally and emotionally. Both aspects establish our sensory immersion in games.

5.1 Feeling of PRESENCE

Research findings by McMahan also points towards the idea that *presence* is the result of perceiving immersion, which is 'accomplished by blocking as many of the senses as possible to the outside world' ; while *immersion* refers to the 'user's mental absorption in the world'.⁷² The study by Schuemie et al. makes this point clearer, which assumes that the ability to interact with the mediated environment is the most important factor in the sense of presence.⁷³ According to Steuer, this 'mediated perception of an environment' should be called **telepresence**, as opposed to the term "presence" which refers to the 'natural perception of an environment'. In addition, Steuer explicitly indicated that these mediated environments include 'an animated but non-existent virtual world synthesised by a computer (for instance, the animated "world" created in a video game)'.⁷⁴ For Steuer, when one is forced to perceived two separate environments simultaneously through the communication technology, 'telepresence is the extent to which one feels present in the mediated environment, rather than in the immediate physical environment'.⁷⁵ This is consistent with the study by Lombard & Ditton which uncovers that presence is the 'perceptual illusion of nonmediation', namely, the extent to which 'a person fails to perceive or acknowledge the existence of a medium during a technologically mediated experience'.⁷⁶

⁷² McMahan, *op. cit.*, p.77

⁷³ Schuemie, M.J., van der Mast, C. A. P. G., Krijn, M., and Emmelkamp, P. M. G. (2002). Exploratory Design and Evaluation of a User Interface for Virtual Reality Exposure Therapy. In Westwood, J. D., Hoffman, H. M., Robb, R. A., Stredney, D. (Ed.), *Medicine Meets Virtual Reality*, IOS Press, 468–474. Retrieved from https://www.researchgate.net/publication/8257680_Exploratory_design_and_evaluation_of_a_user_interface_for_virtual_reality_exposure_therapy

⁷⁴ Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, 42, No. 4 (Autumn), 73–93. Retrieved from <http://papers.cumincad.org/data/works/att/27eb.content.pdf>

⁷⁵ *Loc. cit.* *The emphasis is added by this study.

⁷⁶ IJsselsteijn, W. A., Ridder, H., Freeman, F., and Avons, S.E. (2000). Presence: concept, determinants, and measurement. *Proc. SPIE 3959, Human Vision and Electronic Imaging V*, (2 June). Retrieved from <https://doi.org/10.1117/12.387188> *The emphasis is added by this study.

This 'perceptual illusion of nonmediation'—the telepresence, or the feeling of presence in a mediated environment—not only happens between physical and virtual surroundings, but also takes place between fictional locations, for example by clicking around the map one immediately arrives somewhere else within the game world. Similarly, Alluque`re Roseanne Stone also describes presence as the result of the unique persona within the physical body being transported to a mediated world.⁷⁷ This **teleportation** feature of immersion is essential for creating the 'perceptual illusion of nonmediation' which sustains the "feeling of presence" within the user system—namely the game. The feeling of presence could be built up through sounds in three aspects : details in world design, spatial audio, and the audio-only asset.

I. Details in world design

According to Huiberts, high quality auditory worlds in games can engage the player while masking the sounds in the user environment through two means : offering details that 'sucks the player into the game', and contributing to a stronger connection with the game. In Huiberts' survey, some interviewees mentioned the reason that makes them fond of the game is 'the smallest things had their own sounds...All sounds are very distinct and really make the game complete'.⁷⁸ Besides, the "auditory fitness" also makes the game world more convincing and pleasant. As one interviewee put it : ' [It] is not necessarily that adds something to the game mechanics, but really to the atmosphere. [...] Every building and character has its own sounds. [...] Because of good connection and consistency, the game has a nice busy atmosphere'.⁷⁹

In «NieR: Automata» ,10 primary places have their own location music in game which helps the player quickly get information about where he/she is in or in proximity to and who he/she is very likely to meet. For example, when the players hear the intro of <**3. Peaceful Sleep**>, they instantly know that they are approaching the Resistance Camp

⁷⁷ Stone, A. R. (1996). *The War of Desire and Technology at the Close of the Mechanical Age* (Cambridge, MA: The MIT Press), 83–92.

⁷⁸ Huiberts, *op. cit.*, p.57

⁷⁹ Huiberts, *op. cit.*, p.58

where they can not only gather information from Anemone, the alliance leader of the camp, but can purchase some weapons as well. Or, when they (tele)transport via Access Point⁸⁰ to Bunker—the space station which serves as the headquarter of YoRHa infantry—they will hear <**30. Fortress of Lies**> as soon as they arrive, and then receive the next mission from the Commander. The location music of Machine Village, <**17. Pascal**>, could be cited as a good example.

The Machine Village first shows up in gameplay 3 as the machine alliance to the players whose leader is a pacifist machine, Pascal. <**17. Pascal**> starts to play when the players arrive or get closer to the connected bridge, and the piece continues presenting Machine Village in this way at gameplay 4, 5, 11, 12, and 13. After the Machine Village has been destroyed in gameplay 18, <**17. Pascal**> is replaced by <**21. Wretched Weaponry: Medium/Dynamic**> when A2 fights against the fierce machine lifeforms inside the Machine Village. In other words, the players only hear the different location music of Machine Village when he/she has completed route **A** (gameplay 1-8), **B** (gameplay 9-15) and is heading for ending **C**.

Compared to other places in the game, Machine Village is the most peaceful place, full of innocent machine kids and machine philosophers⁸¹ without war. Okabe, the lead composer of 《NieR: Automata》, said, 'This is indeed the most idyllic piece among others ; however, I couldn't compromise with the order from Yoko, that is to use the mechanic sound of vocoder only. Thus, I have still put the original chorus, sung by the child, into the piece'.⁸² So, how did he successfully achieve this calm and pleasant atmosphere when combining these two different elements—the chorus of human and the vocoder sounds of the machine?

⁸⁰ The machine that the players could check map information, receive emails from other characters in game, and save the current record of gameplay (in case they might be killed in the coming battles). One place accommodates a few Access Points which is easily to find.

⁸¹ Such as: the village machine Satre and his followers, Make-up machine. In fact, Pascal also loves to read Nietzsche (see gameplay 18).

⁸² Saito, K. and Kaneko, M. (2017, March). 『NieR:Automata』 サントラ配信記念！音楽制作・MONACAスタッフインタビュー。Mora website. Retrieved from <https://mora.jp/topics/interview/nier-soundtrack/> (The citation is translated from Japanese by this study, the same hereafter).

Pascal

NieR: Automata

Keiichi Okabe

The musical score for 'Pascal' is presented in four systems. The first system is the 'INTRO' (measures 1-6), marked with a tempo of quarter note = 92 and a chord of C7sus4. The piano part features a melodic line in the right hand and a bass line in the left hand. The second system is 'VERSE A-1' (measures 7-12), also marked with C7sus4. It includes a piano part and a vocal line in the right hand. The third system is 'VERSE A-2' (measures 13-16), marked with C7sus4, featuring a piano part and a vocal line. The fourth system is 'VERSE A-3' (measures 17-20), marked with C7sus4 and a key signature change to A major (A ♮ M7). It features a piano part and a vocal line. Colored circles highlight specific melodic phrases: green circles highlight phrases in measures 6, 10, 12, 14, 16, and 18; orange circles highlight phrases in measures 8, 9, 11, 13, 15, 17, and 19. The score also includes dynamic markings (p, mp, mf) and performance instructions (clap + percussion, plus : piano + guitar + drum, plus : chorus of Machine, Chorus of Human kids (replaces the voice of machine)).

Figure 2. Score analysis of <17. Pascal>
(The coloured notes are added by this study)

As shown in Figure 2, the same chord, C7sus4, is used from INTRO to VERSE which functions as the base to unite different parts of the music. Even in the VERSE, different phases are assigned to different voices—either it's the vocoder, the human chorus, or just the instrumental melody. More importantly, the same theme (i.e., the melodic phrases circled in orange and green) appears almost every phase of the VERSE with slight variations—sometimes in the change of time interval (e.g., bar 9-12), sometimes having adjustments for the need of chord progression (e.g., bar 17-20).

According to the research by Sho Iwamoto, the programmer and the author of 《Grimoire of Game Design》, one prominent feature of the music at 《NieR: Automata》 is the vocals sung in every single soundtrack, which is very rare to see in

games. 'Compared to other games that use vocals only on battles or important events, NieR: Automata uses vocals to boost the emotions of the scene. That is, the vocal within the music is triggered by the progression of narrative, not just by the enemies'.⁸³ Iwamoto observed that 'either at the Resistance Camp or Machine Village, before the player is able to have any conversation with the characters of the place, they might feel suspicious about the people there. Once the main characters (such as Anemone, Pascal) introduce themselves with kindness, the vocals come in naturally within the ongoing music, which also represents the player's feeling of security'.⁸⁴ This is often applied to the dialogues happening in Machine Village. For instance, in gameplay 3, after Pascal has introduced himself, dubbed by a voice actress⁸⁵, his conversation with 2B and 9S immediately become silent and are replaced with subtitles. Instead of having a "sounded dialogue", the child's chorus comes in <17. Pascal> to accompany the activities that the players are going to complete in the village. By flexibly switching between voice acting or the vocoder/human chorus within the music, <17. Pascal> successfully makes the atmosphere harmless and creates a pleasing mood that encourages the players to think about profound questions thrown out by the characters, such as : 'Wait, [do] they [machine lifeforms] have genders?' by 9S, or 'Do you believe that existence precedes essence?' by Satre (see Appendix #1).⁸⁶

<17. Pascal> reveals the main theme directly with the visual—Pascal and the Machine Village at the beginning. Nevertheless, not all the location music follows the same rule. <35. Tower> shows exactly the opposite approach of presenting "place"—not only the "physical" mediated space, but also the "abstract" mental space of the characters. This track doesn't appear until gameplay 14, then it shows up in gameplay 16 and 20.

⁸³ Iwamoto, S. (2018, September 25). NieR: Automata—音楽によって「接続される」世界. [Web log post] note. Retrieved from <https://note.com/geekdrums/n/n9bbd292c74e1> (The citation is translated from Japanese by this study, the same hereafter).

⁸⁴ Iwamoto, *op. cit.*, p.11

⁸⁵ Although Pascal is set to be a male machine, he is dubbed by female voice actress both in Japanese and English versions.

⁸⁶ Both examples are quoted in gameplay 3, 33:40—33:50 for 9S, 34:10—34:35 for Satre.

The Tower

Figure 3. Score analysis

NieR: Automata

Keiichi Okabe

The musical score is presented in five systems, each with a piano part and a vocal part. The tempo is marked as $\text{♩} = 80$. The key signature is one sharp (F#), and the time signature is common time (C).

- System 1 (Measures 1-12):** Labeled "INTRO soundscape". The piano part features a complex, layered texture of chords. Chords are annotated as Bm7(9), C#m7(♭5)/B, Bm7(9), and E7/B. The vocal part, labeled "vocal 1 VERSE-A", begins with a *mp* dynamic. A purple circle highlights a specific chord in the vocal line.
- System 2 (Measures 13-23):** The piano part continues with chords GM7, Em6/G, GM7, GM7/F#, Em7(9), Em6, Em7(9), Em7(9)/D, CM7(9), and C#m7(♭5). Two purple circles highlight specific notes in the piano part.
- System 3 (Measures 24-27):** Labeled "vocal 2 VERSE-B". The piano part features chords F#7sus4 and F#7. A section is marked "[omitted here]". The vocal part is labeled "Synth sextuplets replace the rest note VERSE-A''".
- System 4 (Measures 28-29):** The piano part features chords C#m7(♭5)/B and Bm7(9). The vocal part continues with sextuplets.
- System 5 (Measures 30-34):** The piano part features chords Bm7(9)/A and GM7. The vocal part continues with sextuplets. A purple circle highlights a specific note in the piano part.

Figure 3 illustrates the main structure of <35. Tower> which is INTRO–VERSE A–VERSE B –A’–B’–A’’–OUTRO.⁸⁷ The whole piece is based on the electronic soundscape of INTRO, and there are more synthesis sounds coming in from the VERSE A’ to A’’. In gameplay 14, only the INTRO is shown as a loop for the hacking space of 9S’ memory, while in

⁸⁷ Verse B, B’, sung by vocal 2, and Verse A’ which has the same chord progression but with electronic sounds, are omitted in Figure 3.

The musical score for '35. Tower' is presented in a system of five staves, each containing a piano (left) and guitar (right) part. The score is in the key of D major (two sharps) and 4/4 time. The guitar part features a prominent sixteenth-note run in the upper register, with a '6' above the notes indicating a sixteenth-note figure. The piano part provides harmonic support with chords and bass lines. Chord changes are indicated by labels above the staves: Em6/G, GM7, GM7, Em7(9), Em6, Em7(9), Em7(9)/D, CM7(9), and C#m7(b5). Two specific sixteenth-note runs in the guitar part are circled in purple, with the label 'GM7/F#' written below the first circle and 'Em7(9)/D' below the second circle. The system concludes with a double bar line and a '8' below the piano staff, indicating an 8-measure phrase.

The musical score for '42' is presented in a system of two staves, piano (left) and guitar (right). The score is in the key of D major (two sharps) and 4/4 time. The guitar part features a chord progression from F#7sus4 to F#7. The piano part provides harmonic support with chords and bass lines. The system concludes with a double bar line and an '8' below the piano staff, indicating an 8-measure phrase.

Figure 3. Score analysis of <35. Tower>
(The coloured notes are added by this study)

gameplay 16, it replaces <2. City Ruins> as the location/combat music when 9S follows the (virus-infected) operator's order to fight against enemy in City Ruins, where the track plays from INTRO to the beginning of VERSE B. Interestingly, the real space "Tower" does not come into sight until the player continues to gameplay 18. So, when the

players hear part of <35. *Tower*> in gameplay 14 and 16, they have no idea what is going to come out. Instead, they hear “new” music in the character’s head and in the very familiar space where they’ve already listened to the same track from route A to almost the end of route B. This is a clever way to trigger the players’ interest, especially as it foreshadows that the coming rival is not the machine lifeform but the virus-infected operator who was part of YoHRa infantry. Furthermore, when it first presents the place, that is Tower, in gameplay 20, it’s still not the full piece (i.e., the VERSE A’ and A’’) ; that can only be heard if the player continues fighting against the Red Girl Boss (official name as “Terminal Alpha and Beta”) up to the very end of route C.

Compared to <**17. *Pascal***>, introducing the location/character in medias res, <**35. *Tower***> unveils its main purpose—presenting the location “Tower” and the relative character “Red Girl”—step by step, ‘like a brewing mist, gradually dissipating and giving away to the bright light’, as well-described by classical composer Ernst Toch, when he distinguished two types of beginning and ending of a composition.⁸⁸ In fact, this strategy is effective and supported by Gestalt psychology. According to Sonnenschein, the eight principles of Gestalt psychology not only apply to visual perception, it can also ‘find equivalents in aural perception’.⁸⁹ Sonnenschein stated that because the same timbre represents the same source, ‘the law of similarity states that similar sounds are grouped together even when separated by time’.⁹⁰ In the case of <35. *Tower*> , the pre-appearance of part of the piece in gameplay 14 and 16 functions as hints for when the players finally enter the Tower. Further, the best part of this score is it adds sextuplets of synthesiser sound in high registrar, which not only contrasts with the downbeat of the piano in low register but also creates an edgy mood to fight against the Red Girl—the immortal but self-evolving program, “the Network”.⁹¹ This tense feeling becomes more obvious when the rest notes are replaced by synth sextuplet in VERSE A’, as Figure 3 illustrates.

⁸⁸ Toch, E. (1948, 1977). *The Shaping Forces in Music*. (the Dover ed.). New York: Dover Publications, Inc.

⁸⁹ Sonnenschein, *op. cit.*, p.79

⁹⁰ Sonnenschein, *op. cit.*, p.82

⁹¹ Red Girl. [n.d.]. *NIER Wiki*. Retrieved from Nier Fandom website https://nier.fandom.com/wiki/Red_Girl

Briefly speaking, both pieces demonstrate the flexible music strategy of «NieR: Automata», which not only use suitable instrumentation to present the corresponding spaces in the game, such as, the kid's chorus and the synth sextuplet, but they also successfully embody the narrative through the core capacity of game audio—atmospheric function⁹²—supporting the players' activities in those spaces. According to Steuer, the richness of the visual and the relevant audio are exactly the "vividness", which he identified it as 'one variable property of media technologies that influences their ability to induce a sense of presence'.⁹³

II. Spatial audio

The vividness created by the soundtrack can not only induce the sense of presence in different spaces of the game, but the atmospheric sounds can also 'work indirectly to influence player action', as found by Jørgensen.⁹⁴ This is supported by the studies by Wirth et al.⁹⁵ and by Hamed and Perkis who focus on "Spatial Presence", a commonly accepted subtype of "Presence".⁹⁶ According to Wirth et al., 'labels like "sensory factors" (Witmer & Singer, 1998), "sensory richness" or "vividness" (Steuer, 1992) refer to media factors that contribute to enduring attention allocation'. Wirth et al. concluded that a continuous stream of highly detailed information could sustain users' involuntary attention allocation more effectively.⁹⁷ This is because the more sensory information a mediated environment emits to the participants, the more likely it is that the participants' attention allocation will persist. In other words, the user-participants' attention is likely to remain longer within the mediated surroundings, as they are "enveloped" by the

⁹² Jørgensen (2006), *op. cit.*, p.3

⁹³ Steuer, *op. cit.*, p.11

⁹⁴ Jørgensen (2006), *op. cit.*, p.3

⁹⁵ Wirth, W. & Hartmann, T. & Böcking, S. & Vorderer, P. & Klimmt, C. & Schramm, H. & Saari, T. & Laarni, J. & Ravaja, N. & Gouveia, F. & Biocca, F. & Sacau, A. & Jäncke, L. & Baumgartner, T. & Jäncke, P. (2007). A Process Model of the Formation of Spatial Presence Experiences. *Media Psychology*. 9. (pp.493-525). New Jersey : Lawrence Erlbaum Associates, Inc.

⁹⁶ Hamed, A. & Perkis, A. (2018). Spatial Storytelling: Finding Interdisciplinary Immersion. *11th International Conference on Interactive Digital Storytelling, ICIDS 2018, Dublin, Ireland, December 5–8, 2018, Proceedings*. Retrieved from https://www.researchgate.net/publication/329062184_Spatial_Storytelling_Finding_Interdisciplinary_Immersion_11th_International_Conference_on_Interactive_Digital_Storytelling_ICIDS_2018_Dublin_Ireland_December_5-8_2018_Proceedings

⁹⁷ Wirth, Hartmann, Böcking, Vorderer, Klimmt, Schramm, Saari, Laarni, Ravaja, Gouveia, Biocca, Sacau, Jäncke, Baumgartner, and Jäncke, *op. cit.*, p.500 *The emphasis is added by this study.

sensory stimulations within the simulated environment—their immersive experiences inside the game world continue carrying on.

Although the aim of this paper is not to distinguish different theories of Presence⁹⁸, the study by Hammed and Perkis on **Spatial Storytelling** worth noting here : ‘Spatial Storytelling works by spatially engaging a user inside a mediated environment whose discovery through exploration advances a non-linear narrative, and where space is the essential communication medium’.⁹⁹ This is supported by most cases of games, especially in 《NieR: Automata》, the correlation between the space and its corresponding characters is clear and strong—you’ll never meet Pascal in Resistance Camp (even they are alliance), and you would only fight against the Opera Boss (Beauvoir) in the Amusement Park. A study by Jenkins found that ‘spatial stories can evoke pre-existing narrative associations ; [...] they provide resources for emergent narratives’¹⁰⁰, which reveals that environmental storytelling can create preconditions for immersive narrative experiences.

As I mentioned before, the Machine Village is presented with its location music <17. *Pascal*> at the beginning (gameplay 3, 4, 5, 11, 12,13) but replaced by <**21-1. Wretched Weaponry: Quiet**> after it has been invaded by savage machine lifeforms at gameplay 18. Another example is the Resistance Camp. It is originally presented with <3. *Peaceful Sleep*> (see p.87) whose gentle style (bpm 64) is consistent with its functions as a friendly alliance and a supply station. However, after it has been destroyed by brutal machine lifeforms, the high-tempo combat music <**5. Birth of Wishes**> (bpm 166) starts to play at the Resistance Camp at the end of route A (gameplay 8). Both examples demonstrate well how the change of soundtracks can effectively subvert the mood of the spaces they’ve established before, in order to support the emergent narratives with the progression of game. ‘Games are usually characterized by spatio-temporal markers,

⁹⁸ The studies by Wirth et al. (2007) present a comprehensive view on the process model of Spatial Presence. For further discussions please consult the bibliography of this study.

⁹⁹ Hammed and Perkis, *op. cit.*, p.6

¹⁰⁰Jenkins, H., (2004). Game Design as Narrative Architecture. In Noah Wardrip-Fruin and Pat Harrigan (Ed.), *First Person: New Media as Story, Performance, and Game* (pp.118-130). Cambridge, Mass: MIT Press

that is, we point at a certain “thereness” (dungeon, lake, downtown library, etc.) to communicate how far we are in a game ; **space relays information on time**, as Hammed and Perkis concluded.¹⁰¹

Besides, Hammed and Perkis pointed out that the “experience of being enveloped by technology” could localise immersion to the **system’s sensorimotor contingency**, which is ‘to map and match the user’s proprioception and the range of information it affords the senses (visual, haptic, aural, etc.)’.¹⁰² This is supported by the study of IJsselsteijn, Ridder, Freeman and Avons which indicated that “the extent and fidelity of sensory information” and “the match between sensors and the display” are two determinant factors of “Presence”. The former refers to ‘the amount of useful and salient sensory information presented in a consistent manner to the appropriate senses of the user’—in other words, how visually and aurally detailed the demonstration of the world (system) is, while the latter, called **sensory-motor contingencies**, means ‘the mapping between the user’s actions and the perceptible spatio-temporal effects of those actions’.¹⁰³ So, in the following section, I will examine the spatial audio strategy of 《NieR: Automata》 to see how they implemented the sound design to increase levels of fidelity and create a system of high sensorimotor contingency. This can be divided to three parts : Simple 3D, Interactive verb, and Doppler effect.

Before proceeding to inspect each spatial acoustic technique utilised in the game, there are three points we should bear in mind. Firstly, 《NieR: Automata》 is not a Virtual Reality game nor is it recommended that one wears headphones. Hence, ‘creating a 3D sound atmosphere regardless of playback environment’¹⁰⁴ motivated the audio team to develop their own plugins which could express the depth through a normal stereo

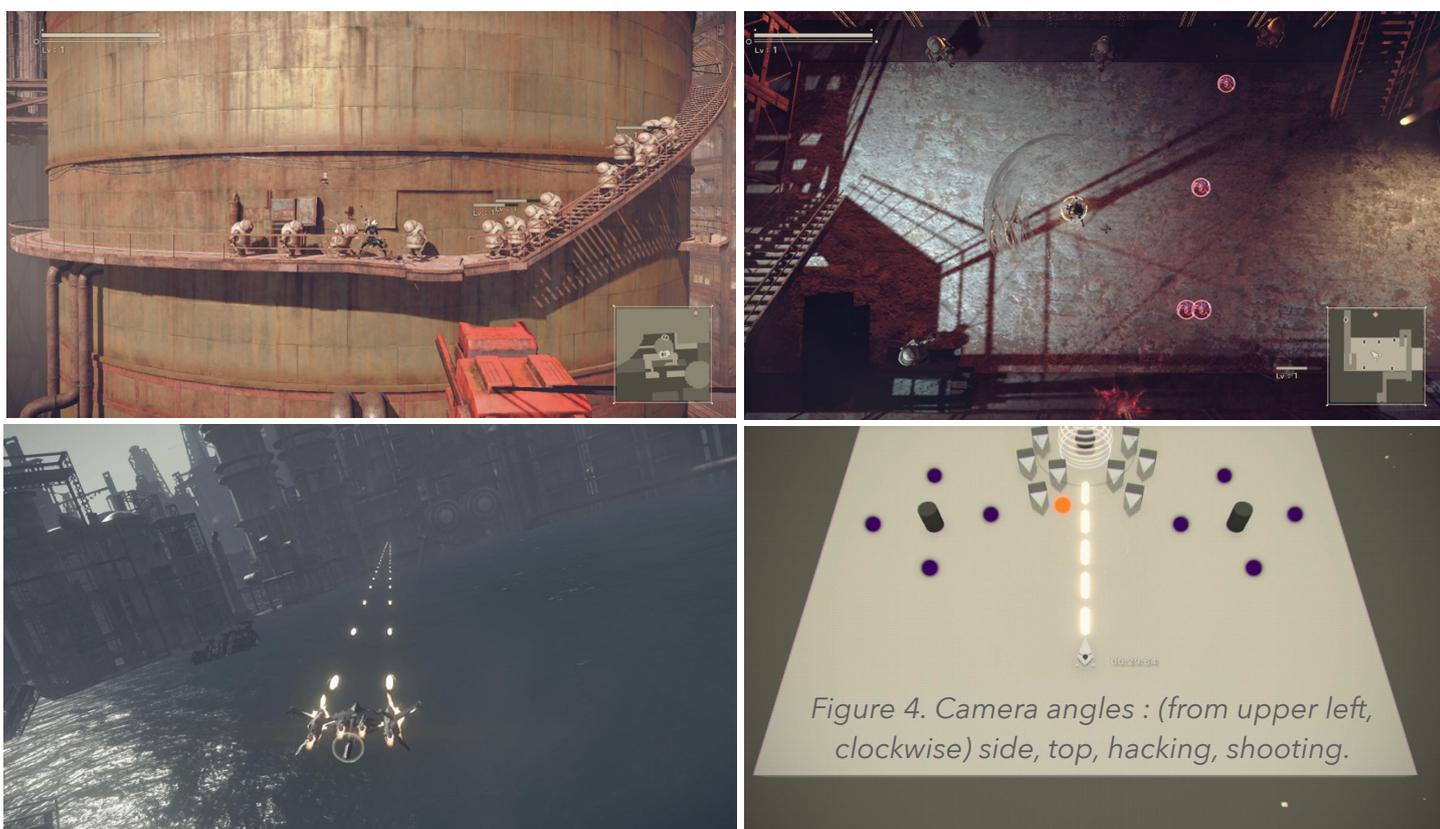
¹⁰¹Hammed and Perkis, *op. cit.*, p.6 *The emphasis is added by this study.

¹⁰² Hammed and Perkis, *op. cit.*, p.3 *The emphasis is added by this study.

¹⁰³ IJsselsteijn, Ridder, Freeman, and Avons, *op. cit.*, p.521

¹⁰⁴ Kohata, S. (2018). An Interactive Sound Dystopia: Real-Time Audio Processing in ‘NieR:Automata’. [Presentation] *Speech slide presented at Game Development Conference (GDC 2018-Vault)*. Moscone Center, CA.

speaker setup, instead of applying HRTF technique.¹⁰⁵ They also wanted to develop, 'effects with a low CPU cost that can be used on almost any sound'. According to Shindo and Kohata, the sound designer and audio programmer of 《NieR: Automata》, if there are too few specific sounds applied to the 3D effect, making them stand out, 'the player cannot experience the feeling of being surrounded by sounds in a 3D environment'.¹⁰⁶ Hence, it is necessary to keep the processing CPU of each sound as light as possible. Finally, it is worth discussing the 'effects carefully adjusted to preserve design and protect original audio', especially the sounds from the front. It is important to keep the original sound when applying these 3D effects, not only for the processing purpose, but for getting the whole balance as well. In short, all the spatial audio techniques are consciously utilised at 《NieR: Automata》 in order to make the audio more realistic for the purpose of increasing immersion and create surrounded effects. Furthermore, it enables the sound team to create a cohesive atmosphere for each space, helping them to blend the sounds together.



¹⁰⁵ Shindo, M. and Kohata, S. (2018, December 4). Part 1: The spatial acoustics of NieR:Automata, and how we used Wwise to support various forms of gameplay. [Web log post] Audiokinetic Blog. Retrieved from <https://blog.audiokinetic.com/the-spatial-acoustics-of-nierautomata-and-how-we-used-wwise-to-support-various-forms-of-gameplay-part-1/>

¹⁰⁶ *Loc. cit.*

II-1. Simple 3D and Subjective EQ

Because 《NieR: Automata》 is an ARPG in the third-person view, a third-person camera is set behind the player as default.¹⁰⁷ Given that there is a variety of gaming types within the game (e.g., shoot 'em ups, text adventure, etc.), it is essential to switch camera's angle between the top view, different side view, shooting scenes and hacking spaces, which can be seen in Figure 4.¹⁰⁸

Since "hearing" has an omni-directional character, study by Jørgensen found that a stereo system 'is not able to provide information on whether the source is located in front of or behind the player'.¹⁰⁹ Only the true surround systems can demonstrate 'significant possibilities for providing detailed information about the location of an offscreen source'. However, it was later shown by Huiberts that separate speakers can be used to 'situate the player in the middle of the action and—depending on the perspective—make the player aware of what is behind him'.¹¹⁰ In fact, the approach adopted in 《NieR: Automata》 is a great solution, as it proves that the surround effect is possible to achieve through stereo speaker setup.

In order to build up the sense of direction for each sound, the audio team integrated the **Simple3D** plugin to the program. As Figure 5 demonstrates, the original sound signal goes through four routes to generate the final output which is heard by the player. The *first* route allows the frequencies which are lower than 500 Hz to pass, while the *second* route allows those frequencies which are lower than 4K Hz to pass. The *third* route is the original input signal with a full range of frequencies that go directly to the mixing stage. The *fourth* route is specially designed for the sound source that goes behind by using the band pass filter (BPF). When there is a need to change frequencies, the sound designer adjusts the value of BPF on route 4 in order to raise or reduce the output

¹⁰⁷ Besides the changing camera view, it's also vital to set the best 'listener position' whose best location is found at the right side by the player. Since the distance between the listener position and the player is always fixed, it won't influence the hearing result much. For further discussion on its 'unique attenuations', please see footnote 272.

¹⁰⁸ *Loc. cit.*

¹⁰⁹ Jørgensen (2006), *op. cit.*, p.3

¹¹⁰ Huiberts, *op. cit.*, p.59

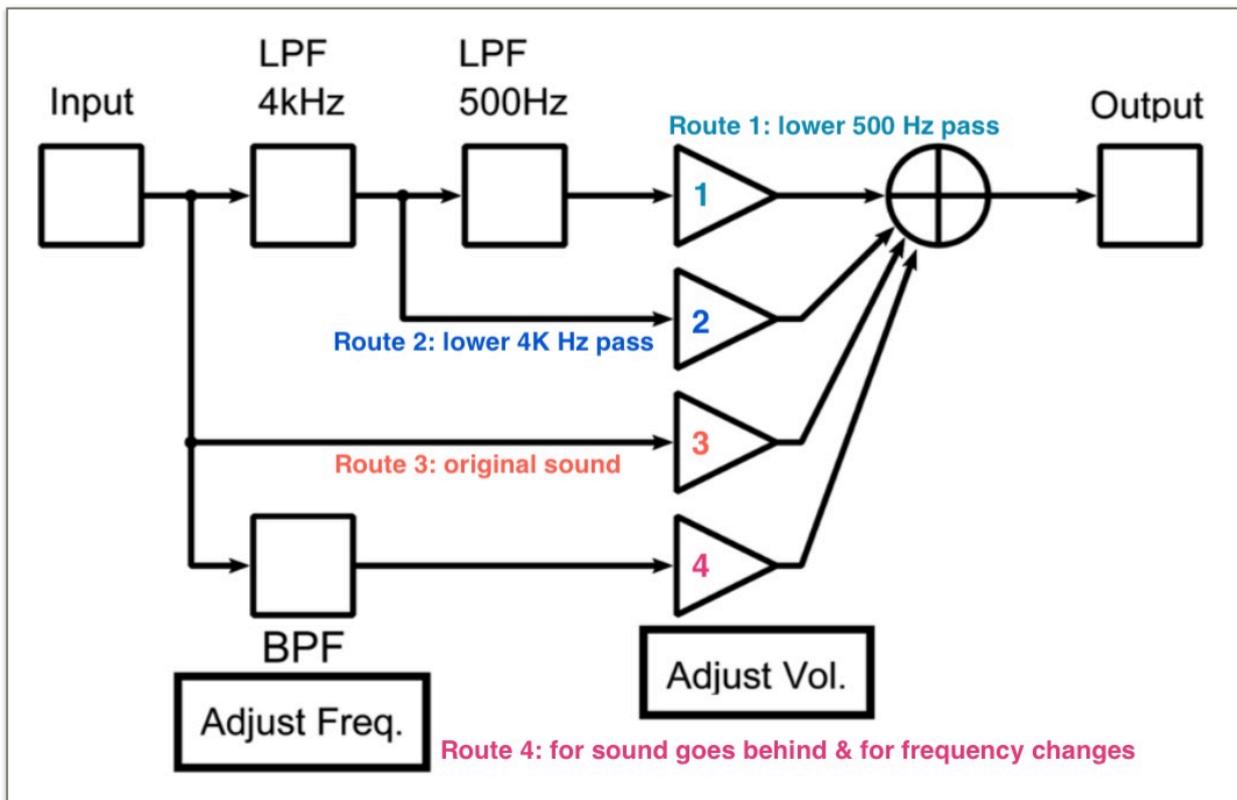


Figure 5. Simple3D DSP diagram [from PlatinumGames Inc.]
 (The coloured notes are added by this study, the same hereafter in Figure 6 and 8)

frequencies. The frequencies can also be increased or decreased by changing the volume ratios (shown as 'Adjust Vol.'). All these frequencies' adjustments depend on the angle of the listener to the sound source. The band pass filter (BPF) and volume ratio can change for each route, aiming to generate the sense of direction. Since the BPF of route 4 can continue adjusting the interface, it can emphasise the sound sources that go behind the player, meanwhile, it also collects the "noticeable sounds" which might be filtered out by other routes.

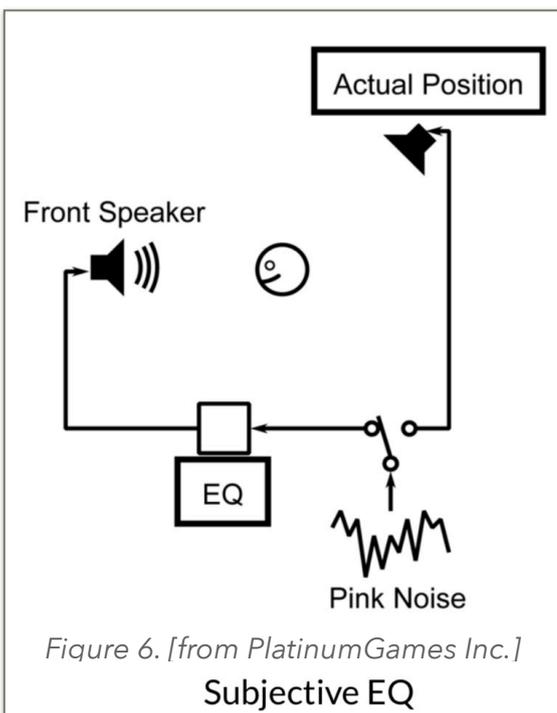


Figure 6. [from PlatinumGames Inc.]
 Subjective EQ

After the process of Simple3D, the mixed sound output goes through the next step : **Subjective EQ.** As illustrated in Figure 6, there are 2 speakers—one is always in front of the player while the other could be in any direction. The EQ functions on the front speaker while simultaneously the pink noise is applied to these two speakers by turns. The sound designer listens and compares the audio from these two speakers until both audio sound as close as possible. Then, the adjusted EQ is applied to that

specific sound of that specific direction. This process is done for all directions. Based on the EQ parameters of each direction, the sound designers can recreate the sound position through the front speaker. By going through the processes of Simply3D and Subjective EQ, 《NieR: Automata》 successfully strengthens the feeling of “sounds coming from the backside” by providing the high-detailed sound information, whose audio performance is smoother than those done through the traditional panning technique which only offers the limited directions of left and right.

II-2. Interactive reverb : Raycast, K-verb and Occlusion

The Interactive reverb consists of two parts : the Raycast system and the K-verb plugin. According to Kohata, the main goal of the audio team is continuously updating any change happening in the scenes in real-time. Therefore, automation is indispensable for maximising the interactivity of the game.¹¹¹ The general idea of their approach is to start by analysing the surrounding terrain through Raycast then adapting the reverb accordingly through K-verb. Finally, they use automation to continuously adjust reverb based on multiple parameters such as : direction, level, depth, the terrain composition, the real-time geographic changes, and so on.

① Raycast system

The Raycast system is used to detect the geometry of the space based on the player’s position. According to Shindo and Kohata, for each frame, several rays are casted in random directions from the player’s position, as shown in Figure 7. The Figure 8 presents how the raycast system is applied to different spaces in the game world. The green dots are the impact points where the rays hit the space, and each ray’s lifespan is

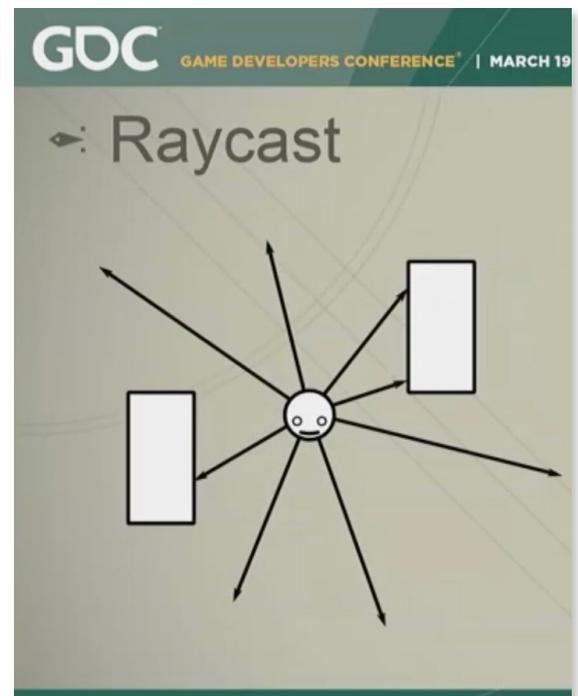


Figure 7. **Raycast** system (from the speech slide of Kohata, see footnote104)

¹¹¹ See footnote 104.

② K-verb DSP

After obtaining the terrain data from Raycast, the audio team begins to make the corresponding reverb with K-verb. Instead of following the player's position, the 'reverb should stay in world location'¹¹⁵, meaning that the main goal is to capture the reverb in relation to the origin source rather than which direction will be seen on the screen. As illustrated in Figure 10, it starts with two sounds—the dry level of the mix (dry component), known as the Input, and the wet level of the mix (wet component), labelled as AUX. Based on the standpoint of the listener (player), the audio team will break the wet component from 5 channels into 8 channels of the raycast's directions, which are on the absolute horizontal plane around the listener. Next, the loop in the middle is where the reverb processes. Data is extracted from the impact points from each direction to determine the delay duration, the level, and filter intensity, such as collision rate, acoustic impedance, and the low-pass filter (LPF) for different materials.

At this stage, the delay is considered to be accentuated too much and too much noise would also occur during this process. So, the reverb goes through the all-pass filter (APF) and Crossfeeds from each direction. Finally, the resulting reverb (the processed wet AUX) returns to 4 channels for the listener (player), and mixes with the dry component (original sounds) into the main output. Keeping the CPU usage minimum is also important for the whole process.

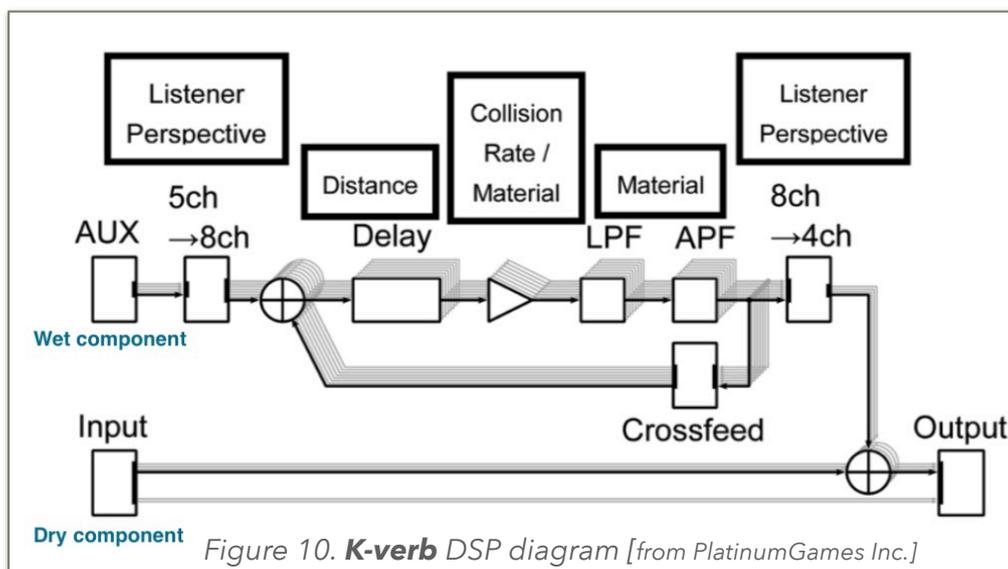


Figure 10. **K-verb** DSP diagram [from PlatinumGames Inc.]

¹¹⁵ Shindo and Kohata (2018), *op. cit.*, p.12-14

③ **Occlusion**

Generally, when the player enters into the building, the reverb is automated by calculating the occlusions within the sound area. However, this is only possible when the event happens inside the architecture, because here the depth is already known. In most situations, the audio team needs to adjust the reverb manually on a case by case basis. Furthermore, in order to make the immersive experience consistent with the narrative, some adjustments must be taken into account—even if it doesn't represent the sound performance in the reality.¹¹⁶ For instance, when the player goes somewhere that is supposed not to be found by the enemy, the audio team would lower the whole volume to match the obscure atmosphere.

Additionally, Shindo and Kohata have emphasised that their priority is producing “satisfying sounds” rather than “accurate simulation”. ‘The sound team focused on creating an immersive environment that would let players stop in their tracks to enjoy their surroundings’, was addressed by Shindo and Kohata.¹¹⁷ Therefore, all the sound effects need to be easily recognisable and pleasing instead of mere precise simulations.

II-3. Doppler effect : Flight unit, Emil

The study by Huiberts revealed that the Doppler effect is often added to moving objects in games in order to increase player engagement. For example, ‘the fantastic audio : rockets demonstrate the Doppler effect when they fly past your head’ in the shooter video game *Quake Live* (2009), or the ‘approaching vehicles with added Doppler effect, attract the focus of the user and make concentrating on the boosters more challenging’ in the audio game *Drive* (2002).¹¹⁸ So, what is Doppler effect ? According to scientist Poessel, the Doppler effect is often referred to as the Doppler shift.¹¹⁹ A common

¹¹⁶ Kamiyama, 【CEDEC2017】 『NieR:Automata』の世界を彩る効果音はどのように実装されたのか？デザインコンセプトとその仕組みについて, *op. cit.*, p.8-9

¹¹⁷ Shindo and Kohata (2018), *op. cit.*, p.2

¹¹⁸ Huiberts, *op. cit.*, p.60 and p.80

¹¹⁹ Poessel, M. (2011). Waves, motion and frequency: the Doppler effect. *Einstein Online*. Band 05–1001. Retrieved from <https://www.einstein-online.info/en/spotlight/doppler/>

example of Doppler shift is the **change of pitch** heard when a moving object (e.g., car, airplane, spaceship) is approaching the observer. In comparison to the frequency emitted from far, the received frequency would be higher when the moving object is approaching ; 'identical at the instant of passing by, and lower during the recession'.¹²⁰ This technique is also utilised in 《NieR: Automata》 , especially on the Flight unit and every time the character Emil makes an appearance.

There are two scenes in gameplay 9 which can be cited as clear examples that demonstrate this effect : One is in 16:37–16:40, when 9S comes out of the tunnel by the flight unit, flying toward the screen then leaving off the left side (see Appendix #2). The other is in 26:56–27:02, when 2B and the flight unit take off from the middle of the screen and fly to the left side. Both pitches of their recession sound are apparently lower than when the flight unit first appears and passes by the viewers (players).

Emil is a companion and major character in 《NieR》 , both in the human form and weapon form, whereas in 《NieR: Automata》 he appears as a machine head with no body. His main role in 《NieR: Automata》 is 'attached to a moving sales cart where he offers his wares for the roaming androids of YoRHa'¹²¹, as seen in Figure 11.¹²² In Kohata's public lecture at GDC2018, they confirmed that they added the Doppler effect on Emil's cart, where the player can buy new weapons or tools without returning to Resistance Camp. Whenever he shows up, the theme song of his sale cart, <**12. Emil's shop**>, plays **from his position** in the game world. If the player is fighting against the enemy in City Ruins, the pitch raising and dropping phenomena simulated by the Doppler effect will be heard by the player. Since Emil's cart moves freely by the game setting, the Doppler effect on the song is more obvious when Emil is within the player's sight. Nonetheless, Emil's cart often shows up unpredictably and out of the scene. Even so the player is still able to hear <12. Emil's shop> at a soft volume, which is what we're

¹²⁰ Doppler effect. *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Doppler_effect

¹²¹ Emil (Character). [n.d.]. *NIER Wiki*. Retrieved from Nier Fandom website [https://nier.fandom.com/wiki/Emil_\(Character\)#Human%20Form](https://nier.fandom.com/wiki/Emil_(Character)#Human%20Form)

¹²² KOS-MOS. (2017). "Emil's Shop :3". *STREAM-NieR:Automata™*. Retrieved from <https://steamcommunity.com/sharedfiles/filedetails/?id=928235607>

going to discuss, as the last factor which establishes the "Feeling of PRESENCE"—the Audio-only asset.



Figure 11. Emil's shop in City Ruins with 2B (left) and 9S (right)

III. Audio-only asset

Huiberts identified that the 'audio is very suitable for presenting time-based information with the advantage that the information is perceived even though the player is not immediately next to the source'.¹²³ This is consistent with the findings of Jørgensen, which show that the orienting functions of game audio can 'provide information about the presence of objects as well as the direction of sound sources. [...] before the player has actually seen it'.¹²⁴ In 《NieR: Automata》, most battles begin after the enemy has appeared on the screen. The only exception is when Emil's shop shows up in City Ruins.¹²⁵ According to the lead composer Okabe, the concept of composing <12. Emil's shop> is based on the director's request which intended to create Emil's cart as a "propaganda truck". Therefore, when the music was implemented in the game, composer Ueda added an effect on it which makes the music sound like it's playing from

¹²³ Huibert, *op. cit.*, p.61

¹²⁴ Jørgensen (2006), *op. cit.*, p.3

¹²⁵ Besides the main 5 endings (ABCDE), it is possible to fight with Emil Boss when the player reaches level 99. In that case, Emil and his cart appears on the screen for the player to attack.

Emil's Shop

NieR: Automata

Kuniyuki Takahashi, Kakeru Ishihama

VERSE-A Emil's vocal (treble clef) + Bassoon / Tuba (bass clef)

♩ = 116 F

INTRO: Trumpet + Snare drum

Figure 12. Score analysis of <12. Emil's shop>
(The coloured notes are added by this study)

a megaphone. 'Since Emil's truck can move anytime in any speed, when it travels rapidly in high speed, the Doppler effect on the music is evident', said by Okabe.¹²⁶

The instrumentation of <12. Emil's shop> also follows the image of "propaganda truck", as Figure 12 demonstrates. The parts of "VERSE A and INTRO' " repeat three times in <12. Emil's shop>, which can be seen as the most iconic section of the whole piece. It mainly uses brass to express the excitement of Emil whenever he appears to sell his products ; the bright trumpet and snare drum (the yellow and orange circles) used in the INTRO and INTRO' (functioning as "bridge" within the piece when it repeats) successfully creates a "marching atmosphere". This marching feel is sustained in the verses that follow as well, through the regular quarter notes of the bassoon/tuba, which are low register sounds in brass family. When it comes to verse, not only do the delightful vocals of Emil show his hospitality to welcome the clients (players), but the melody phrases in high and low pitches successfully represent their dialogues from bar 2 to bar 12, just like they are bargaining with each other.

In most cases the player can't see Emil within their sight immediately but can hear <12. Emil's shop> playing somewhere. Depending on the level of volume the player heard, he/she can estimate how far Emil's shop is located and decide whether he/she will go closer to purchase weapons or animal feed to ride on the boars in City Ruins. In truth, the Doppler effect is very helpful for players to locate Emil, especially when they can't distract themselves from their ongoing battles.¹²⁷ In addition to this, some players pointed out that 'the 3D sounds seamlessly change to the 2D sounds' when the shop's menu opens (i.e., the screen view changes from the battle field to the 2D interface where is full of products for players to choose from).¹²⁸ This is consistent with the experience that the audio team were attempting to offer. According to Kohata, they also mixed some Lo-fi noise in Emil's music. After the player has talked to Emil and entered

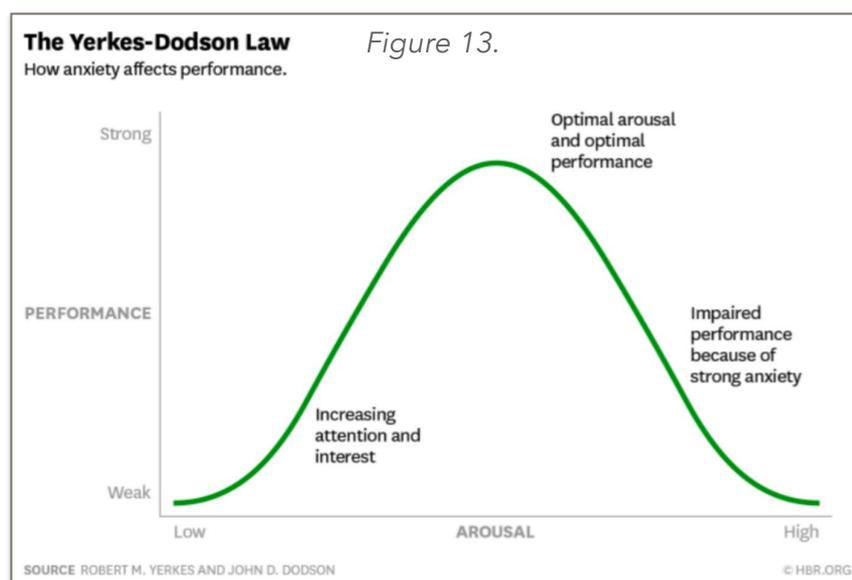
¹²⁶ Nagayoshi, H. (2017, April). 『ニーア』 サウンドを生み出す、音の職人たち. 2083 web. Retrieved from <https://www.2083.jp/contents/201704nier/> (The citation is translated from Japanese by this study, the same hereafter.)

¹²⁷ GDC Vault - An Interactive Sound Dystopia/ Real-Time Audio Processing in 'NieR/Automata'. *Interactivemusic website*. Retrieved from https://scrapbox.io/interactivemusic/GDC_Vault_-_An_Interactive_Sound_Dystopia:_Real-Time_Audio_Processing_in_'NieR:Automata' *The citation is translated from Japanese by this study.

¹²⁸ *Loc. cit.*

the shop, the Lo-fi parameter turns off and the soundtrack switches to stereo.¹²⁹ We'll examine the Lo-fi effect further in chapter 7.

The study by Coen reveals the significant relationship between sounds and the player's performances. He lists three aspects that we should consider seriously when using sound as a tool to improve—or worsen—a player's ability in games. The first and the most obvious one is that 'Audio can alert the player to things outside of their field of view'. We have already shown the necessity of having the "audio-only asset" through the discussion of *<12. Emil's shop>*. The second aspect is that the 'Auditory reaction time is significantly faster than Visual reaction time'. It is based on several studies conducted on athletes during 2007–2010 which showed that the test subjects 'scored 40-50ms faster with auditory as opposed to visual reaction time'.¹³⁰ The last aspect is that 'Audio produces a stronger physiological response than Visuals alone'. As highlighted by Coen, sound is not just for the purpose of immersion or enjoyment, because having music or sounds go along with the visual can also make the receivers more arousing or stressful. He noted that 'according to the Yerkes-Dodson Law, a small amount of stress or anxiety helps us complete complex tasks more effectively, but once we reach a certain level of stress we begin to perform worse', as presented in Figure 13.¹³¹



¹²⁹ See footnote 104.

¹³⁰ Coen, O. (2018, June 5). Player Success: How to Help or Hinder It with Sound. [Web log post] Audiokinetic Blog. Retrieved from <https://blog.audiokinetic.com/player-success-how-to-help-or-hinder-it-with-sound/>

¹³¹ *Loc. cit.* *Figure 13 is also from the study by Coen.

As Coen reminded us, the more sounds happening at once, the higher the player's cognitive load. Therefore, even when we are designing the "audio-only asset", it is essential to take the visual and gameplay into account, 'because they add to this load as well'.¹³² Otherwise, the resulting information overload might reduce the players' performances or delay their reactions, which leads to an unpleasant experiences and easily drawing them out of the mediated environment, as they will seek rest in the real world. Namely, the immersion in the game crashes due to the fatigue.

In fact, the issue of cognitive load is related to our next topic—Dynamics—which has the potential to be one of solutions to the "information overload" problem. Before proceeding to the next section, we can see the "Feeling of PRESENCE" is successfully established in 《NieR: Automata》 by offering details in world design through its diverse music approaches, building up a system of high sensorimotor contingency via the effective spatial audio design, and rather minimal audio-only asset which shows the presence of the object to support their main activities, combatting in the game.

5.2 Sensory gratification

Besides the "feeling of presence", the other pillar that enhances the sensory connection in sensory immersion is "sensory gratification". According to Huiberts, the sensory gratification can be stimulated through dynamics and appealing audio.

IV. Dynamics

When designing an immersive world of game, Huiberts indicated that 'preventing the player from becoming tired of sound by keeping the world dynamic can make the gameplay appealing for a longer time'.¹³³ As mentioned before, Iwamoto discovered that every single soundtrack of 《NieR: Automata》 has a vocal element which is very rare to see since 'human is very sensitive to human voice and tend to focus on it once there is someone singing'.¹³⁴ That is also the reason why most of games only add human voice

¹³² *Loc. cit.*

¹³³ Huiberts, *op. cit.*, p.64

¹³⁴ Iwamoto, *op. cit.*, p.2

on combat music or vital events. 'When the player continues hearing the soundtracks with vocals singing one after another, it is easy to get tired soon', was observed by Iwamoto. However, this listener fatigue actually doesn't happen in 《NieR: Automata》. Iwamoto commented that instead, 'all the vocals within the music melts into the game world without feeling uncomfortable'.¹³⁵ Through the previous analysis on <17. Pascal>, we've seen how the vocal/chorus within the music blends into the game through the narrative progression. Nonetheless, that doesn't completely solve the possible listener fatigue of "playing/listening for a long time", especially when the main battle fields usually play the same track, such as <2.City Ruin>. It presents City Ruins almost throughout the whole game, only shortly replaced by <35.The Tower> in gameplay 16. Essentially, the player would hear the same tune playing at this main location from route A to route B (i.e., gameplay 1 to 15) without any change. So, in order to encourage players to continue playing, the crucial question becomes this : how does one reduce the listener's fatigue and avoid them to be fed up with repetition?

① **Vertical Remixing : Multi-layers, Multi-functions**

According to Michael Sweet, the audio director for more than 100 award-winning video games, there are two methods of interactive composition applied to games—Horizontal re-sequencing and Vertical remixing. The former is 'typically used when one idea has finished and a new one needs to begin'¹³⁶ whilst the latter is when multiple layers of music are added or taken away to create levels of intensity and emotion'.¹³⁷ As Figure 14 exhibits, the "horizontal re-sequencing" method shows two different music pieces appearing one after another. Given that each music cue has its own idea and musical theme (e.g., tempo, rhythm), rapid changes are discouraged in this method as they would easily distract the player from his/her current level or state. There are three techniques could be applied to smooth this process : crossfading, transitional and

¹³⁵ Iwamoto, *op. cit.*, p.3

¹³⁶ Sweet, M. (2015). *Writing Interactive Music for Video Games : a composer's guide* (2nd ed.). Indiana : Pearson Education. Inc. (p.206)

¹³⁷ Sweet, *op. cit.*, p.219

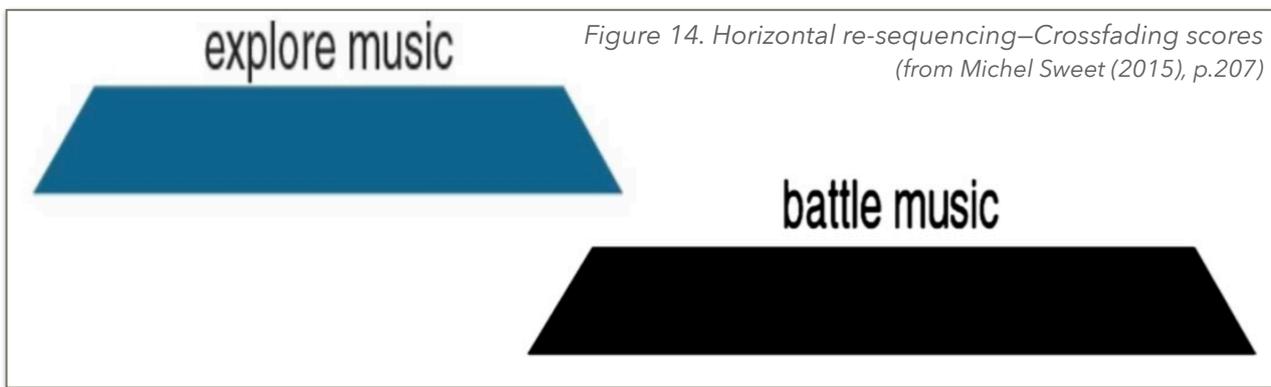


Figure 14. Horizontal re-sequencing—Crossfading scores
(from Michel Sweet (2015), p.207)

branching scoring models. But in general, horizontal re-sequencing is ‘better suited for longer sequences (more than 20 seconds per cue)’, as highlighted by Sweet.

Vertical remixing, on the contrary, is a single music piece that is divided into several layers targeting various emotional states within the game.¹³⁸ Figure 15 illustrates the general concept of this method¹³⁹: **Layer 1** is the ambient drone or soundscape, starting when the player firstly enters the level. When he/she sees danger in a distance, the music engine of **Layer 2** fades in to increase the tension via percussion or rhythmic elements. Then, when the player encounters the enemy, the string ostinato or melodic element of **Layer 3** fades in on top of that. After the battle is ended, Layer 2 and Layer 3 will fade out accordingly. Since these different layers originate from the same cue, there is musical continuity between different states. ‘This allows [the] music engine to fade in layers easily and quickly without the player becoming distracted by changes to the music’, was analysed by Sweet.

According to the investigation by Iwamoto, 《NieR: Automata》 uses the ‘vertical remixing’ method to implement all the soundtracks into the game. Firstly, he observed that except for a few specific boss battles, there is no alternative combat music for lesser

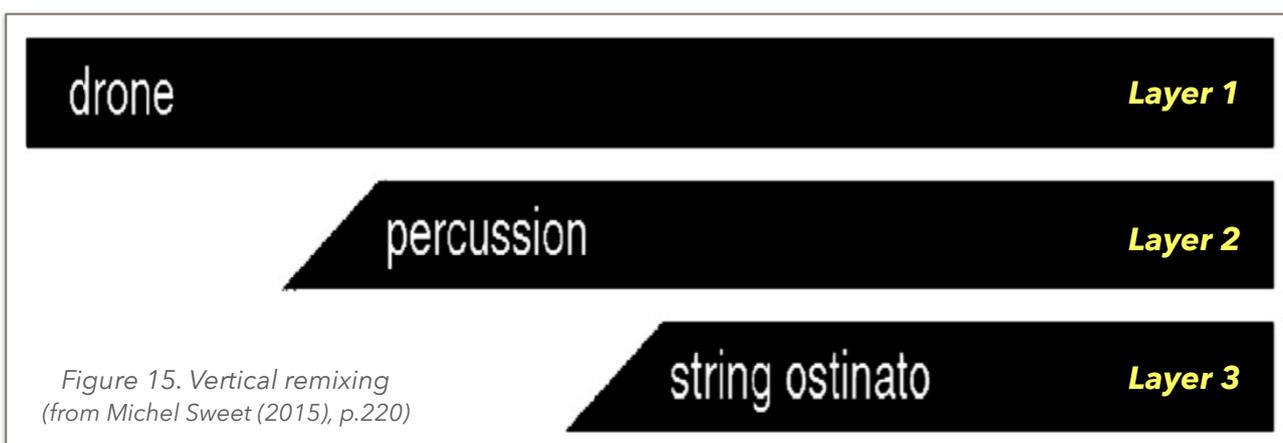


Figure 15. Vertical remixing
(from Michel Sweet (2015), p.220)

¹³⁸ *Loc. cit.*

¹³⁹ Sweet, *op. cit.*, p.220 *The colour notes are added by this study.

enemies (i.e., enemies of lower level). The way to inform the player of the approaching danger is via taking good advantage of vertical remixing, which is 'useful when the composer needs multiple quick changes to intensify the score'.¹⁴⁰ Secondly, taking <2.City Ruin> as an example, Iwamoto pointed out that even there is no event occurs, the ambience (Layer 1) and the rhythmic elements (Layer 2) still appear by turns when the player is exploring the place. He identified that these variations aim to keep the players' interest, not feeling bored. Thus, these kinds of arrangements are deliberately made on each soundtrack.¹⁴¹ In other words, the composer intentionally creates these varied arrangements in one piece, with the aim of making it application to different situations—fights, explores or nothing happens. Put it differently, apart from a few specific tunes for specific boss battles, the music of 《NieR: Automata》 is multi-functional through its clever implementation of vertical remixing.

Sweet highlights that, the most effective score of vertical remixing starts off 'being an excellent piece of music that has dynamics, swells in intensity, and offers harmonic changes' before splitting into layers.¹⁴² That is, it is important that each layer is never just an instrument loop that remains stagnant, but should 'still have a complete musical thought containing dynamics and emotional level'.¹⁴³ This inability to sync harmonic or melodic changes to game events is actually the biggest disadvantage of vertical remixing, because all layers belong to the same music cue, 'the harmonic framework is static. [...] tempo map is also static', as Sweet stated.¹⁴⁴ Even though it might be at the cost of flexibility, vertical remixing can still be used to add a cohesive quality to the music without making people bored. Figure 16 demonstrates how <2.City Ruin> gets this job done. The ambient soundscape of Layer 1 (N.C., i.e., No Chord) starts at the beginning and continues playing till the end. Layer 2 is the most dynamic part because it accommodates three types of rhythm. Since Rhythm 1 (INTRO and OUTRO) is quite

¹⁴⁰ Sweet, *op. cit.*, p.219

¹⁴¹ Iwamoto, *op. cit.*, p.15-16

¹⁴² Sweet, *op. cit.*, p.229

¹⁴³ *Loc. cit.*

¹⁴⁴ Sweet, *op. cit.*, p.220

City Ruins-Rays of Light

NieR: Automata

Keiichi Okabe

N.C. soundscape $\text{♩} = 100$ *ad lib. p*

INTRO Rhythm 1 **VERSE A-1** Rhythm 2 *mf*

[omitted here]

*Red. **

12 **G \flat M7** **G \flat M7** **B \flat m7(9)** **B \flat m7(9)** **G \flat M7** **G \flat M7**

18 **VERSE A-2** Vocal comes in + Rhythm 2 **B \flat m7(9)** **B \flat m7(9)** **G \flat M7** **G \flat M7** **B \flat m7(9)** **B \flat m7(9)** **E \flat /B \flat**

25 **E \flat /B \flat** **G \flat M7** **G \flat M7** **Fm7** **Fm7** **E \flat m7(9)** **Fm7(9)**

32 **VERSE B** Vocal continues + Rhythm 2 **B \flat m7(9)** **B \flat m7(9)** **B \flat m7(9)** **B \flat m7(9)** **G \flat M7** **G \flat M7** **B \flat m7(9)**

Figure 16. Score analysis of <2. City Ruins>
(The coloured notes are added by this study)

39 **B \flat m7(9)** **E \flat /B \flat** **E \flat /B \flat** **G \flat M7** **G \flat M7** **Fm7**

45 **Fm7** **E \flat m7(9)** **Fm7(9)** **B \flat m7(9)** **B \flat m7(9)**

INTRO' Rhythm 2 continues **VERSE B' G \flat M7**

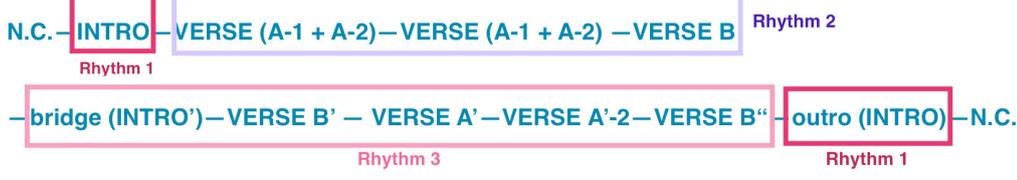
[omitted here] *mp*

VERSE B' No vocal + Rhythm 3 (arpeggio)

52 **G \flat M7** **Fm7** **Fm7** **E \flat m7(9)** **E \flat m7(9)** **D \flat M7** **D \flat M7** **C \flat M7**

60 **C \flat M7** **B \flat m7(9)** **B \flat m7(9)** **Gm7(\flat 5)** **C7(\flat 9)** **F7ses4** **F7**

[Structure]



similar to Rhythm 3, I put them in the similar colours—red and pink. Compared to the obvious changes between different rhythms on Layer 2, the melodic elements of Layer 3 almost remain the same, that is, the harmonic framework is relatively fixed. For example, the INTRO and INTRO' are totally omitted here because they have the same chord progression as VERSE A (including A-1 and A-2). But there are three ways to prevent the "motionless movement" which happen in this tune. **Firstly**, the combination of alternative rhythms makes it sound different, which are shown in the purple group (with Rhythm 2) and the pink group (with Rhythm3). **Next**, the order of the different parts of VERSE is free. For instance, VERSE B' comes before VERSE A' in the pink group. And when it reprises, it can only repeat VERSE A'-2, instead of all parts of VERSE A'.

Finally, <2.City Ruin> has Layer 4—the vocal, which appears in VERSE A-2, VERSE B in purple group, whereas it doesn't always show up at its counterpart in the pink group—you won't hear the vocal singing at VERSE B', but you can hear it again at VERSE B". These variations add more flavour into the same harmonic movement, which solidifies the idea that an excellent score of vertical remixing will 'all change and evolve over time'. This is because it not only allows the composer to 'leverage the emotional musicality of the piece and its dynamics, but also allows the music to later be enhanced by the player's actions in the level', as indicated by Sweet.¹⁴⁵

The multi-layer structure of <2.City Ruin> enables itself to adopt changes in games and keep the player's interest. Nevertheless, the general fade-in/fade-out rule, suggested by Sweet, doesn't completely apply to the actual implementation at 《NieR: Automata》. Instead, in order to avoid listener fatigue, a clean mix for battles and dialogues at the same time is fundamental. The study by Coen revealed that cognitive loading is affected not only by the audio but also the visual. Thus, he proposed that to keep the player's focus on the sounds and the visuals that matter the most, 'a very difficult section of a game, or something with a lot of rapid visuals, may benefit from a reduced number of sounds and a quieter mix overall'.¹⁴⁶ This general principle can be seen in lots of

¹⁴⁵ Sweet, *op. cit.*, p.229

¹⁴⁶ Coen, *op. cit.*, p.5

situations in 《NieR: Automata》. Take gameplay 5 as an example, from 1:04:40 to 1:09:35, where 2B and 9S not only search for the missing Child machine to fulfil the requested mission, but also need to fight against the lesser enemies spontaneously appearing in the City Ruins. When they find the Child machine, they have to persuade him to come back to Machine Village to apologise to his mother. Since the City Ruins is also the main location Emil drives his cart to sell goods, the player can hear <12. Emil's shop> from time to time, even when they're still in a fight. So, there are many audio components going on in these 5 minutes: dialogues between 2B, 9S, and Child machine, several weapon sounds during the battles, the foley sounds when they run back to the Machine Village (e.g., the sound of 2B's high heels, Child machine's jumping sounds), the environmental sounds (e.g., birds' tweet, waterfall) and Emil's song. It would be unreasonable to add Layer 2 (the strong rhythmic elements) of <2.City Ruin> as this would further burden the player during the lesser enemy fights. Instead, the audio team chose to stay in Layer 1 (the ambience) and set the volume very soft. This is especially done based on the equal-loudness curves since, 'humans react more quickly to sounds with rapid changes in volume than slow changes, even if both sounds reach the same peak volume'.¹⁴⁷ By doing so, the most important sounds of weapons and the attack feedback can be heard clearly without hinderance.

But what about the random appearance of Emil's song? By making the attention-grabbing sounds (i.e., high-transient sounds) the priority, we solve the issue of focusing on the battles. However, with the exception of gameplay 4 where Emil only shows up as a machine head without his shop, his theme song <12. Emil's shop> always goes along with <2.City Ruin> in City Ruins; this happens every time he pops up, in gameplay 5, 6, 7, 16, 17, 18, 19 and 20. How does this work? Kutay identified that this **sonic sludge** could be prevented by providing a priority indexation that describes which sounds are important and at what points they need to be noticed.¹⁴⁸ Kutay emphasised that 'the key to preventing sonic fatigue is to create sound effects that **vary in volume and**

¹⁴⁷ *Loc. cit.*

¹⁴⁸ Huijberts, *op. cit.*, p.64

frequency in relation to each other'.¹⁴⁹ In truth, that is exactly the solution for playing <12. Emil's shop> parallel to <2.City Ruin>.



¹⁴⁹ Kutay, S. (2006, May 1). Bigger Than Big: The Game Audio Explosion. [Web log post] gamedev.net. Retrieved from <https://www.gamedev.net/articles/audio/music-and-sound-fx/bigger-than-big-the-game-audio-explosion-r2317/>



Figure 18. Frequency analysis of <12. Emil's Shop>
 [From top : INTRO (no vocal), VERSE A (with vocal)]

Figure 17 exhibits the frequency distribution of <2. City Ruin>, including the part of N.C. (Layer1), INTRO (Layer 1+2, without vocal) and VERSE A-2 (Layer 1+2+3, with vocal) while Figure 18 presents the frequency distribution of <12. Emil's Shop>, from the INTRO (without vocal) to VERSE A (with Emil's vocal and the brass family as rhythmic elements). Given that the brass family's frequencies occupy from 200Hz to 10KHz, especially instruments in high register, such as the trumpet, used as the iconic intro of <12. Emil's Shop>, it is evident that <2. City Ruin> keeps its main frequencies below 1K Hz, whether or not the vocal is within the music. That is to say, whenever Emil's cart appears on the screen, or even if <12. Emil's Shop> is playing somewhere out of sight, there is always the room—the frequency space (2K Hz–10K Hz)—for it to be heard

naturally. 'When volumes and frequencies are selectively assigned, the sound effects will breathe and compliment each other regardless of when they play', concluded Kutay.¹⁵⁰

From my point of view, not only does the appearance of <12. *Emil's Shop*> successfully avoids the "sonic sludge" via frequency allocation, but its unpredictable emergence also restrains the **emotion sludge** to some extent. Huiberts observed that 'depending on the game concept, incorporating the effect sounds of agents that are not visible (yet) can be used in order to create tension or surprise, which can make the game more stimulating'.¹⁵¹ Although it might be risky to play two different tunes with different moods at the same time¹⁵², I'd like to propose an alternative view from an aesthetics perspective. Counterpoint is a frequently used compositional technique in the context of Western classical music. Composer Toch suggested that we should shift the emphasis to the "counter" part of the term which could broaden the interpretation from "the point of contrast" to "**the various happenings in the point of contrast**".¹⁵³ For Toch, the nature of counterpoint is opposition, contradiction, and fighting, which is 'one of the most powerful shaping forces in music (and art altogether)'.¹⁵⁴

Even though the counterpoint is usually discussed in relation to singular pieces of music, Toch's pioneered insight of counterpoint could greatly help us understand the necessity of having two different forces happening simultaneously. According to Toch, counterpoint means the simultaneous presence of two (or more) contrastingly moving voices, or we might say, melodic lines.¹⁵⁵ In gameplay 19 of *«NieR: Automata»* (1:20:06-1:29:30), which is very close to the Ending C (gameplay 20), Ending D and E (gameplay 21), the protagonist 9S has been confronting the deep grief of losing his beloved partner 2B (gameplay16) and is going to hack into the last destination—the

¹⁵⁰ *Loc. cit.*

¹⁵¹ Huiberts, *op. cit.*, p.61-62

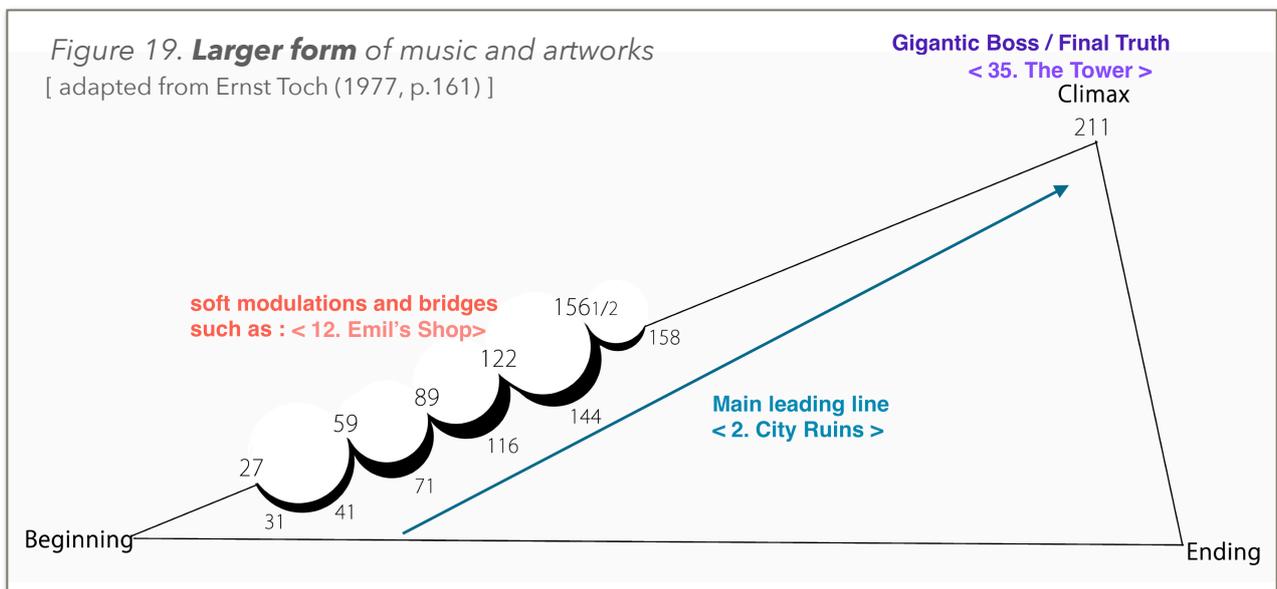
¹⁵² A video clip uploaded by Callard 22 on Youtube whose title is 'Emil ruins the mood' pointed out that there are indeed some players feel the contradiction between the ongoing narrative and Emil's appearance. The video was uploaded on March 31 in 2017. Retrieved from <https://www.youtube.com/watch?v=wDmxuOGh7QI>

¹⁵³ Toch, *op. cit.*, p.134-135 * The emphasis is added by this study.

¹⁵⁴ Toch, *op. cit.*, p.136

¹⁵⁵ *Loc. cit.*

Figure 19. **Larger form** of music and artworks
 [adapted from Ernst Toch (1977, p.161)]



Tower. Before he arrives at City Ruins to approach the Tower, he has just fought against one of his former allies—operator 210, and has been eager to kill the other protagonist A2. In short, the complex emotions of 9S, as well as of the player, will have accumulated and reached a pretty intense state. How does one keep the player sustained in this strong emotional state without feeling too heavy, and still be willing to move on to the cruel truth at the end? Of course, curiosity could be a motivation. Nevertheless, when it comes to the physical condition of the players, as Yerkes-Dodson Law told us, once we reach a certain level of emotional load, the performance is impaired (Figure 13).

Okabe said, 'Director Yoko told me that NieR:Automata is a game of mourning. The world is full of sorrow, where not many happy moments would happen'.¹⁵⁶ Generally speaking, the final treasure/triumph/truth is supposed to be obtained at the end of the game narrative. No game designer wants their players to give up halfway. So, how can one keep motivating the players to fight till the end, even if they start to perform worse? Music can help. If we consider the whole game plot as a music piece, the "various happenings in the point of contrast" could serve as an antidote as they add the dynamic to the whole work. As illustrated in Figure 19, Toch used Wagner's *Meistersinger Prelude* as an example, 'the larger musical form, even if still in one movement, will provide, by subdivisions, for resting points and breathing spaces. [...] the long ascending line, ascending as a whole, will show curves, notches, retarding moments, similar to those of

¹⁵⁶ Nagayoshi, *op. cit.*, p.8

the melodic line'.¹⁵⁷ In 《NieR: Automata》, the melancholic mood of <2. *City Ruin*> is the main ascending line while the delightful contrasts made by <12. *Emil's Shop*> serves as the breathing spaces. With their parallel placement, the tension gets chances to slacken for a while, therefore the whole piece could ensure the resumption of the chase is more effective.¹⁵⁸ 'A tragedy will not loosen its grip after the climax ; but before it, on the long way up to it, provision might well be made for some temporary, refreshing laughs. They will help the reader or spectator to brace himself for the ever tightening grip of the plot', concluded Toch.¹⁵⁹ In other words, the simultaneous presence of two different melodic lines helps the players maintain the intense mood without feeling overload, and prepare themselves to embrace the last gigantic climax to come.

② **Silence** : *Seamless, Decisive moments*

"Listening begins with being silent."—Joachim-Ernst Berendt

'In NieR:Automata, our goal is to connect several wide areas seamlessly, aiming to let the players explore different locations freely and enjoy the natural scenery—especially in City Ruins, the centre of the game world', said Kaji Yasuyuki, the environmental artist at PlatinumGames.¹⁶⁰ It is clear that "seamless" was indeed one of the chief principles when designing the game, both visually and aurally. According to Okabe, for each soundtrack, there are 3 versions—Quiet, Normal, and Dynamic—which are created to adopt different situations in the game. In the quiet version especially, the director Yoko was pretty demanding and requested them several times to remove "too much sound" from the piece. Composer Hoashi observed that 'the requests from Yoko sometimes would make the track so it "doesn't sound like music" which made me a bit worried ; however, this modified "quiet music" amazingly matched NieR's world when I heard

¹⁵⁷ Toch, *op. cit.*, p.161 * The numbers refer to bar code of "Meistersinger", but in terms of game, we can regard them as the boss fights of different levels here. The coloured notes and texts are added by this study.

¹⁵⁸ Toch, *op. cit.*, p.138

¹⁵⁹ Toch, *op. cit.*, p.160-161

¹⁶⁰ Taura, T. (2016, August 4). 『NieR:Automata』 プラチナゲームズ開発現場ご紹介#02. [Web log post] PlatinumGames official website. Retrieved from <https://www.platinumgames.co.jp/dev-nier-automata/article/27>

them inside the game'.¹⁶¹ Okabe added that 'these "quiet version" are actually very useful when changing from one tune to another, which makes the whole game much more connected and more seamless while playing. This is probably the reason why Yoko insisted that sounds were "as less as possible".¹⁶² Following the previous discussion, these music of "quiet version" could be understood as the ambience layer (Layer 1) that usually begins to play when the player is approaching a location and gradually fades out when the player is leaving for the next place.

Considering there is always full track or ambient music in the background, incorporating silence into the game could be the best way to create contrast. Huiberts found that 'the use of silence can also improve focus and concentration, as a sudden silence can make our hearing very alert (Murray Schafer, 1977, p.259)'.¹⁶³ For example, at the last Boss fight with Eve in gameplay 8, which is supposed to be the toughest battle right before the ending A, 2B and 9S only successfully cut Eve's right hand during the combat music <**25. Depending Weakling**> (21:54-25:53). About 20 seconds later, the new combat music <**37. The Sound of the End**> starts to play (26:14), signalling that the battle is not over. Along with its multi-layer composition, a female vocal was singing softly while 2B was determined to chop Eve's head. Suddenly, the music ceased the moment 2B slashed Eve's neck (28:54-29:07). Thus, the silence shows that the battle with Eve is truly finished. During this 7-minute long battle, the mere 20-seconds of lasting silence is undoubtedly the most powerful part of the scene (see Appendix #3).

Besides using silence for such particular moments as shown above, there are other specific occasions in 《NieR: Automata》 where the seamless soundscape is broken down by silence, including defeated in the battles, making a wrong decision and dying, and all the campaign stories of the Machine Lifeforms, Bosses and Pods. Based on the gameplay videos (1-21) by Shirrako, the protagonists from gameplay 1-8 lead to Ending A is 2B, while with almost the same route, 9S becomes the protagonist in gameplay 9-15

¹⁶¹ Saito and Kaneko, *op. cit.*, p.5

¹⁶² *Loc. cit.*

¹⁶³ Huiberts, *op. cit.* p.65

where the player can use "hacking", the exclusive attack skill of 9S till the Ending B. Then, in gameplay 16–21, the players switch between 9S and A2 to play, heading for the rest main endings : C, D, and E. This is supported by Kamiyama's study which demonstrated the framework of 《NieR: Automata》 : When it comes to the second round [route B], not only does the hacking attack of 9S make the game fresher, but the memory campaigns of machine lifeforms that appear at significant points also greatly help solve the puzzles of the story. Then, all the hints/secrets that have been foreshadowed in route A and route B are revealed in the third round [route C], such as : the true reason of building YoHRa project, and how these protagonists react once they know their destiny, etc.¹⁶⁴

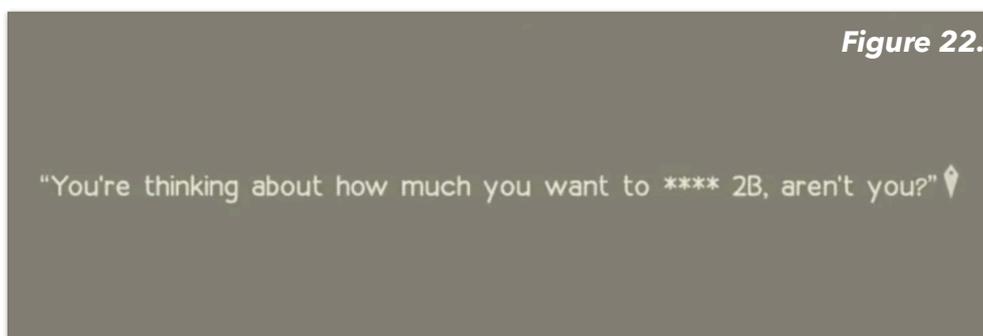
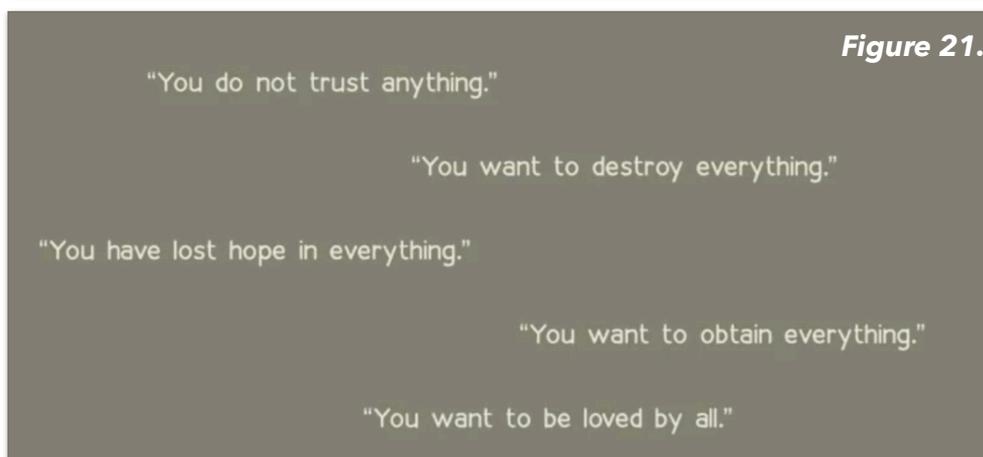
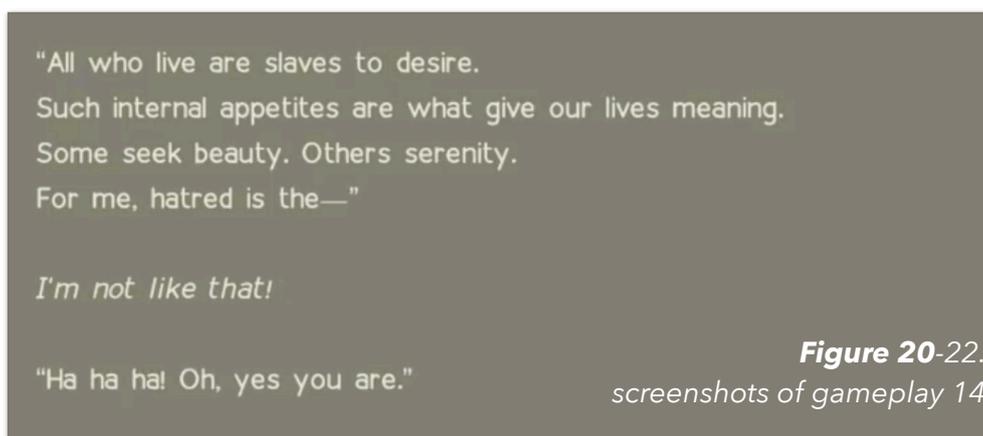
In my opinion, this is an ingenious way of using silence to make these campaign stories stand out. **Firstly**, the silence separates these new-introduced aspects of characters well from the main storyline which consists of continuous battles in the seamless soundscape. **Secondly**, the silence enlarges the contrast between two parties—the androids of justice and the machines of evil. The former takes place in the world which comprises music, rich sounds in battles and environments of high resolution, the players can freely move as per his/her will. Whereas, the latter often shows in the form of 2-dimensional theatre¹⁶⁵, where the player becomes a spectator who can only watch in monochrome style, without music or any other sounds. They passively listen to the narrator telling the nostalgic memories of the machines. Even in the case of Bosses' campaign stories, which often take place in the same environment as the main storyline, the absence of music increases the player's focus on the monologue of the character (e.g., Beauvoir in gameplay 11) or the interactions between characters (e.g., Adam and Eve in gameplay 11, 13, 14, 15).

Through incorporating the silence, new perspectives are recognised by the players, expanding the storyline to a greater scale. The rich narratives not only animate the

¹⁶⁴ Kamiyama, D. (2017, June 17). ニーアオートマタにおける垂直型インタラクティブミュージックの一例. [Web log post] Nine Gates Studio. Retrieved from <http://blog.nine-gates.com/1307/>

¹⁶⁵ There are 10 campaign sessions for Machine Lifeforms, except for 3 stories in gameplay 12-13, other 7 stories are presented in the form of 2D theatre. Please see Table 6 at p.107

characters, giving the story depth, but they're also able to immerse the players into the scene. In gameplay 14, for instance, Adam Boss is hacking into 9S' consciousness, challenging him with provocative questions (31:50-35:00). The ambient music of <23. *Broken Heart*> (Layer1) continues playing till Figure 20 and Figure 21 (-33:40), while 9S keeps denying what Adam says. However, all of sudden, the music is interrupted unexpectedly, then Figure 22—the most intriguing question—comes out. There is no denial from 9S like before but the silence stays till the next frame (33:41-34:25).



Apparently, putting the silence here for this **** question was on purpose—it alerted the players' attention and indeed successfully triggered quite a number of discussions about what does the **** refer to. Below, in the gameplay video uploaded onto

Shirako's Youtube channel, more than one person mentioned that 'the 4th wall is breaking here'¹⁶⁶ (commented by Roman Kurbanov, IERServer)¹⁶⁷, showing that Adam's question somehow sticks out in the audience's (the players') mind. For example, Chitanda Eru even replied to the thread, saying that 'Adam please, Everyone wants to **2B'. Furthermore, people started to imagine what this **** could be. MidnightDawn augured that 'apparently the "*" was actually meant to be 'kill'...I guess it was censored because Yoko Taro wanted us to think it was something sexual, and then ** our minds over when we realised the truth'. Similarly, Mary Nicholson added that 'well, in our defence there are a large number of words that "****" could have been. Love, hate; kiss, kill. Also the "*" is often used to censor curse words so he made it very easy for players to jump to that conclusion'.¹⁶⁸ Since there is no official answer from PlatinumGames or director Yoko, I believe this conversation would continue whenever there is a new comer to the game. Perhaps, no answer is the best answer to engage the players, allowing them to give their own interpretations, which is what director Yoko believes a good game can offer.¹⁶⁹

This section has reviewed the main techniques used in 《NieR: Automata》 to create dynamics. They used vertical remixing method to compose multi-functional music whose multiple layers could fade in/out freely based on game events. Meanwhile, the audio team took good advantage of silence to create contrast, intensifying the depth of storytelling. Therefore, by adopting contrast and silence, 《NieR: Automata》 not only effectively prevents the listener fatigue and emotional sludge within the mediated

¹⁶⁶ The fourth wall is a performance convention in which an invisible, imagined wall separates actors from the audience. The convention assumes that the audience can see through the 'wall' while the actors couldn't. 'Breaking the fourth wall' can be done through either directly referring to the audience, the play as a play, or the characters' fictionality. It's widely adopted in drama, video games and books. Retrieved from https://en.wikipedia.org/wiki/Fourth_wall

¹⁶⁷ Shirako. (2017, February 23). NieR Automata - Gameplay Walkthrough Part 14 - EMP Boss Fight & Story Secrets (PS4 PRO). [Youtube] Retrieved from <https://www.youtube.com/watch?v=L7-YN90D3k0&list=PLCLLeSTzz6trYrOxS1anPprphUxjuOXBAA&index=15&t=0s>

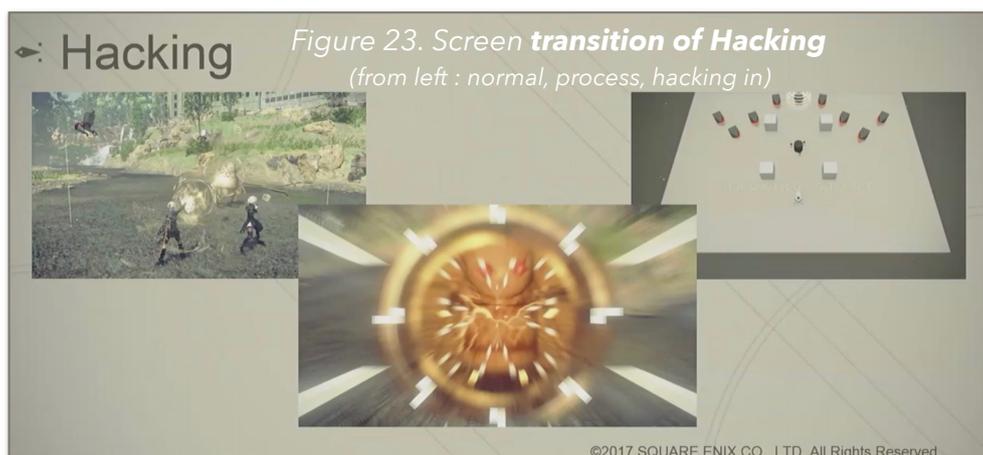
¹⁶⁸ *Loc. cit.*

¹⁶⁹ When the interviewer asked Yoko why he used many philosophers' names for the characters in NieR, he said it would be much more fun for players to find their answers, and that's part of the entertainment of a good game. 歪力. (2017, April). 《尼爾：自動人形》製作人揭2B設計真相：眼罩+黑衣是本體. 4GAMERS. Retrieved from <https://www.4gamers.com.tw/news/detail/31949/nier-automata-producers-interview>

environment, but also immerses the players to think over the plot and characters' intents which extends to further discussions in non-mediated environment—the reality.

V. Appealing audio—8-bit music, Tone Filter

Besides dynamics, the study by Huiberts shows that the appeal of specific sounds and music tracks engages in players' participation in games.¹⁷⁰ The appealing audio sometimes doesn't contribute to the feeling of presence, but instead enhances the "vibe" of the game which 'makes the player want to play more, while at the same time making him/her happy'.¹⁷¹ Kohata mentioned in GDC 2018, that in order to fully present the world of 《NieR: Automata》, they added some electronic effects to enhance the whole atmosphere, including music transition during the hacking sequences and the Lo-Fi effects.¹⁷² In this section, we'll take a close look at the sound implementation on hacking, while the Lo-Fi effect will be elaborated in chapter 7.



For Ueda Masami, the implementation director of 《NieR: Automata》, the most important thing was to make the music flow naturally, and continue to be pleasant for the players no matter how many times these tracks have repeated.¹⁷³ When the players start the route B, they play the role of 9S, who is the scanning model of YoRHa and defeats the enemies by hacking into their electrical circuits. Thus, when the hacking

¹⁷⁰ Huiberts, *op. cit.*, p.66

¹⁷¹ Huiberts, *op. cit.*, p.67

¹⁷² See footnote 104.

¹⁷³ Taura, *op. cit.*, p.9

section begins, the visual changes from the 3-dimensional environments to the 2-dimensional screen, as shown in Figure 23.¹⁷⁴ Following the visual change, the soundtrack gradually devolves into an 8-bit chip version when entering the hacking sections. Instead of using cross-fading, 'the original track is gradually bit-crushed into square waves to smooth the transition', said by Kohata. This "bit-crushed" process is made by **Tone Filter**, the plugin developed by Ueda who was inspired by the previous game's soundtrack <Legend of Nier: 8-Bit Heroes> and therefore proposed to use 8-bit sounds in the hack scenes.¹⁷⁵

The process begins with converting the chosen music pieces from stereo into mono for the DSP processing. As Figure 24 illustrates, a 48-tone range (four octaves) filters out the mono signal as a **sine wave**. At this stage, any very low, very high and subtle tones are cut out. Next, they adjust the level of each tone and apply the distortion to this range, aiming to change its waveform into a **square wave** (a harsher sound associated with classic game consoles).¹⁷⁶ Then, they drop out noise that occurred during the previous stage. Finally, this processed tone range is mixed back in with the original track, after which other spatial effects (e.g., reverb) are added accordingly. There are 2 points that Kohata and Ueda emphasised. **Firstly**, the Tone Filter system generally operates one

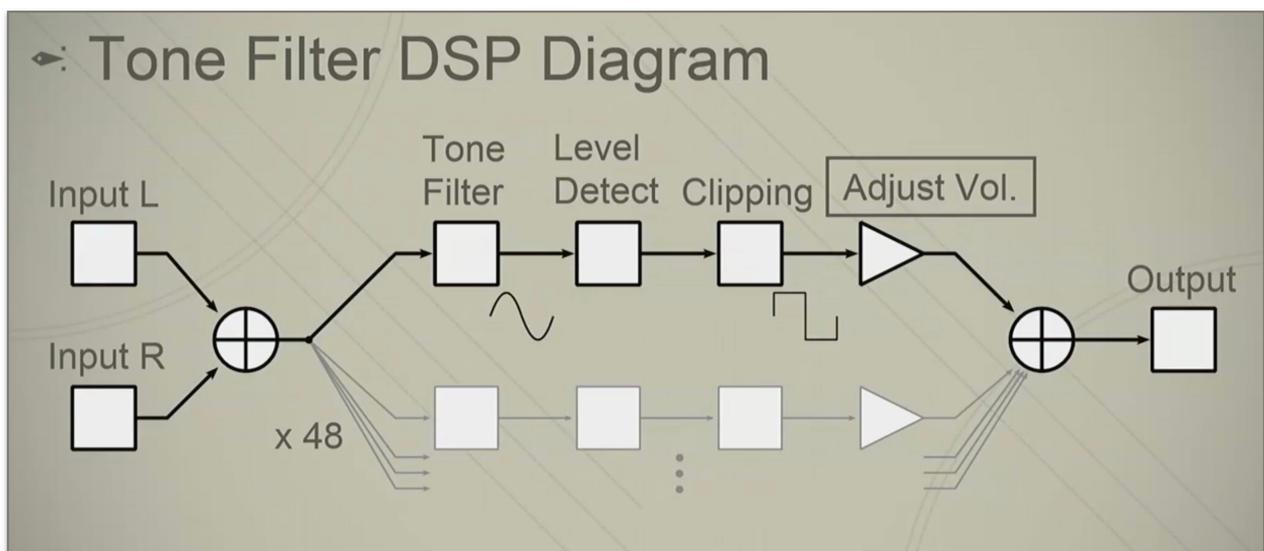


Figure 24. **Tone Filter** DSP diagram (from the speech slide of Kohata (2018), see footnote 104)

¹⁷⁴ See footnote 104.

¹⁷⁵ Ueda, M. (2017, July 25). 『NieR:Automata』 BGM実装の開発裏話. [Web log post] PlatinumGames official website. Retrieved from <https://www.platinumgames.co.jp/dev-nier-automata/article/155>

¹⁷⁶ *Loc. cit.*

track at one time, and the aim is not to generate different sounds but to repurpose the track into square wave by using band-pass-filter in high intensity.¹⁷⁷ Only by using the deconstructed tones directly, is the natural sounding transition possible. **Secondly**, the Tone Filter system may not exactly be a solution to making 8-bit music, but it focuses more on ‘handling several consecutive changes’¹⁷⁸, which corresponds to the chief principle of the game design—seamlessness—for 《NieR: Automata》 .



Figure 25. Transition by Crossfading only (without Tone Filter)



Figure 26. Transition with Tone Filter

Compared to only using cross-fading technique (Figure 25), the sounding effect adopting in the Tone Filter ‘softens the switch over to 8-bit, which adds to the impression that 9S is trying to hack, but doesn’t know whether he’ll succeed. It also makes for a nicer connection to the 8-bit version when the hack is successful’, explained Ueda.¹⁷⁹ Normally, this transition only occurs for 2 or 3 seconds during smoothing the process which is not easy to be noticed. But in gameplay 19, when 9S tries to hack into the Tower for the final boss (1:27:00–1:27:27), the transition lasts 27 seconds. The original soundtrack <**31. Song of the Ancients–Atonement**> is obviously distorted

¹⁷⁷ See footnote 104.

¹⁷⁸ Ueda, *op. cit.*, p.3

¹⁷⁹ *Loc. cit.* *Figure 25, 26, 27 are translated from the same article by Ueda (2017).

before 9S can successfully hack in, which adequately informs the player of the difficulty and the uncertainty of breaking in (see Appendix #4).

However, not every track has an 8-bit version. Since ‘we couldn’t make 8-bit versions of every single piece from scratch, in some cases, we had to use some automatically-generated music’, described Ueda. The automatically-generated music goes through the same Tone Filter DSP until the hack succeeds. The only noteworthy difference is after the hacking, as shown in Figure 27. Ueda indicated that ‘instead of fading completely out, the regular version of the song lingers in the background while the filtered version continues to play over it. [...] We felt that it sounded best with the filtered version at 80% with a bit of a stereo delay’.¹⁸⁰ After this power-consuming DSP stage, the sounds still needed ‘a lot of small improvements [to be], tweaked and tested and built up, to create a quality musical experience’, noted Ueda.



Figure 27. Using Tone Filter for no corresponding 8-bit music

Actually, the idea of using 8-bit music to present 9S’ hacking scenes didn’t appear until Ueda’s proposal. Nonetheless, it turned out to be a very wise decision for 3 reasons.

Firstly, it demonstrates the character’s unique skill well by incorporating suitable sound effects. This especially takes place in rout B, where the players have already known what is going to happen. The new sound effects correspond to the different protagonists of the similar route, which gives the game a brand-new perspective to continue. **Secondly**, the 8-bit music is the new appeal among the normal soundtracks. A study by Reid found that the bit-crushed timbre of chiptune has a specific charm which ‘may adhere to the stylistic conventions of video game soundtracks of the 1980s and 1990s.’¹⁸¹ Thus, not

¹⁸⁰ *Loc. cit.*

¹⁸¹ Reid, G. (2018). Chiptune : The Ludomusical Shaping of Identity. *The Computer Games Journal*, 7, 279–290. <https://link.springer.com/article/10.1007%2Fs40869-018-0070-y>

only is a chiptune's "technostalgia" summoned from the players' side, but the advanced yet retro atmosphere also makes these high-tech android (i.e., 9S, 2B, A2) more "humanised". **Thirdly**, the 8-bit music brings the game narrative to the next level—metaphysical and reflective. Reid discussed the chiptune's "technostalgia" by examining Mike D'Errico's paper with the insight from Jacques Derrida's notion of "hauntology".¹⁸² This is originally a philosophical concept, referring to the return or persistence of elements from the past, and has been applied to cultural studies on electronic music by cultural theorist Mark Fisher and music journalist Simon Reynolds since 2000. They use **sonic hauntology** to describe a music aesthetic preoccupied with pastiche cultural memory, temporal disjunction and **the nostalgia for "lost futures"**—'the anticipated future never actually arrived', noted Fisher.¹⁸³ If the players have completed all the main endings, at least up to the Ending C (gameplay 20), they would see that 9S has gone mad and yelled out the meaning of their "life", "being", or "existence"—YoHRa was created to perpetuate the lies, all of them are sacrificial lambs—no future for them will come. The last point might need some context to reflect upon, nevertheless, we can affirm that the bit-crushed sound quality makes these androids' grief, love, and hate more convincing. 'This notion suggests that the interfaces and sonic qualities of chiptune's microchip technologies impact identity through their agency—one that is non-human', concluded Reid.¹⁸⁴

Having discussed how to construct Dynamics and create Appealing audio that increases the sensory gratification, I'd like to sum up this chapter—Sensory Immersion—with the help of scientific findings by Coen. Previous research has shown that instead of distracting their attention, the game soundtracks can help the players effectively focus. A study done in 2010 and 2012, focused on the action/adventure game *Legend of Zelda: Twilight Princess* (Nintendo EAD, 2006), hypothesised that players would "perform best with the in-game audio on". They tested this with 2 experimental groups : one with the

¹⁸² Reid, *op. cit.*, p.280

¹⁸³ Fisher, M. (2013). The Metaphysics of Crackle : Afrofuturism and Hauntology. *Dancecult : Journal of Electronic Dance Music Culture*, 5(2), 42–55. <https://dj.dancecult.net/index.php/dancecult> (p.45)

¹⁸⁴ Reid, *op. cit.*, p.280

in-game audio off, and the other group with the in-game audio off with the unrelated music playing in the background. They found that 'the average player actually scored highest when playing the game with unrelated music in the background! [...] mute [...] only applied to players who were unfamiliar with the franchise or had less experience with games', pointed out Coen.¹⁸⁵ For the less experienced players, they have to focus more on learning the controls, familiarising themselves with unexpected scenes coming one after another, and so forth. Thus, 'the in-game audio ended up acting as a distraction' when they were already experiencing 'a high cognitive load from the visual', this is the listener fatigue we have examined before. However, for the more experienced players, Coen identified that they can 'effectively take advantage of the additional information given to them by the in-game audio without experiencing an excess cognitive load'.¹⁸⁶ Furthermore, this "additional information" **heightens the players' arousal level**, not only physically supporting them to perform better, but also increasing 'the intended impact of the multimedia experience, making the experience seem real'.¹⁸⁷ The early study by Cohen has already suggested that 'heightened arousal temporarily suspends contact with "real" reality and facilitates belief in the virtual reality'.¹⁸⁸

Slater indicated that 'a sign of presence is when people behave in a VE [virtual environment] in a way that is similar to what their behaviour would have been in a similar real life situation'.¹⁸⁹ That is, the more a mediated environment is experienced as real, the more a player feels present in it. In brief, music and the in-game audio are capable of facilitating the sensory connections through "sensory gratification" and the "feeling of presence", therefore, they productively enhance the sensory immersion of games.

¹⁸⁵ Coen, *op. cit.*, p.2-3

¹⁸⁶ *Loc. cit.*

¹⁸⁷ Cohen, A.J., *op. cit.*, p.17

¹⁸⁸ *Loc. cit.*

¹⁸⁹ Slater, *op. cit.*, p.3-4

6. Challenged-based immersion

“Philosophers from Plato to Sartre have remarked that people are most human, whole, free, and creative when they play.”—Mihaly Csikszentmihalyi

The second dimension of the Audio Enhancement framework, challenged-based immersion, enhances the players' engagement with the activities. Huiberts suggested that 'audio in games can be a valuable constituent for supporting and challenging the player in this process, since both activity and sound progress in time (Cf. Buxton, Gaver, & Bly, 1991)'.¹⁹⁰ This time-based feature of music and game audio could 'result in a continuous experience of deep engagement'—that is **flow**¹⁹¹, a concept proposed by Mihaly Csikszentmihalyi in 1997.

According to Csikszentmihalyi, 'flow denotes the holistic sensation present when we act with total involvement. [...] 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'.¹⁹² Simply put, "that was fun" or "that was enjoyable" is often heard when describing such a flow experience. The study by Jenova Chen also adopted Csikszentmihalyi's flow concept to examine the Dynamic Difficulty Adjustment (**DDA**)¹⁹³ in video games. He referred to Csikszentmihalyi's flow as 'the feeling of complete and energized focus in an activity, with a high level of enjoyment and fulfilment (Debold, 2002)'. Based on Csikszentmihalyi's theory, he developed the award-winning game 'fIOW'¹⁹⁴ with Nicholas Clark, aiming to provide the experience that enables players 'to

¹⁹⁰ Huiberts, *op. cit.*, p.68

¹⁹¹ *Loc. cit.*

¹⁹² Csikszentmihalyi, M. (2014). *Flow and the Foundations of Positive Psychology— The Collected Works of Mihaly Csikszentmihalyi*. Claremont, CA : Springer.

¹⁹³ Dynamic Difficulty Adjustment (DDA) is the process of automatically changing parameters, scenarios, and behaviours in a video game in real-time, based on the player's ability. The goal of DDA is to keep the user interested (not feeling bored or frustrated) from the beginning to the end, providing a good level of challenge.

¹⁹⁴ 'fIOW' is an independent video game created by Jenova Chen to accompany Chen's master thesis. The Flash version of 'fIOW' (PS3) had been played over 3.5 million times by 2008, and won the Best Downloadable Game award at the 2008 Game Developers Choice Awards. Retrieved from [https://en.wikipedia.org/wiki/Flow_\(video_game\)](https://en.wikipedia.org/wiki/Flow_(video_game))

lose track of time and to forget all external pressures'.¹⁹⁵ Indeed, Csikszentmihalyi's extensive research showed a wide spectrum of play-forms, and he identified that 'flow seems to occur only when persons face tasks that are within their ability to perform. This is why one experiences flow most often in activities which have clearly established rules for action, such as rituals, games, or participatory art forms like the dance'.¹⁹⁶ Despite the high correlation between "flow" and "immersion", we should be careful not to confuse these two different ideas. **Firstly, flow is not equal to play.** Csikszentmihalyi's research was done by interviewing 'a variety of people who have invested a great deal of time and energy in play activities', including rock-climbers, explorers, marathon swimmers, chess masters, composers of music, modern dancers, and inveterate gamblers. Even though the most typical "flow experience" occurs during playing, as Csikszentmihalyi explicitly expressed, 'play is not synonymous with flow'.¹⁹⁷ Instead, the flow experiences depend on flow activities, where 'in each case the person is able to use some skills in acting on a limited area in his or her environment'. And this environment could be a mediated environment, such as a computer game of virtual reality, or a non-mediated environments, like a mountain, beach, casino, gym, working studio or a dancing hall. The "immersive experience" we have examined so far has focused on its three features—teleportation, absorption, and identification—within the mediated-environment, the computer game. Apparently, the "teleportation" is not necessary for the flow experience when it refers to the activities in the non-mediated environments. So, we can consider "flow" as a broader concept than "immersion".

Secondly, flow is not equal to immersion, even though they have high correlation. Csikszentmihalyi defined the "flow experience" as the condition that 'we came to call the well-ordered, fully functioning dynamic state of consciousness. [...] in each case the complexities and contradictions of the world are filtered out until only a limited set of

¹⁹⁵ Chen, J. (2006). Flow in Games. (Master thesis). Available from Jenova Chen's website https://www.jenovachen.com/flowingames/Flow_in_games_final.pdf

¹⁹⁶ Csikszentmihalyi, *op. cit.* p.138 *The emphasis is added by this study.

¹⁹⁷ Csikszentmihalyi, *op. cit.* p.137

well-ordered goals and means are left in awareness'.¹⁹⁸ This definition shows that *Flow* shares the same "absorption" feature as *Immersion*. However, people who have experienced flow, often describe it as "self-forgetfulness", as a "loss of self-consciousness" and even a "transcendence of individuality". 'What is usually lost in flow is not the awareness of one's body or of one's functions, but only the self-construct, the intermediary which one learns to interpose between stimulus and response', highlighted Csikszentmihalyi.¹⁹⁹ This is consistent with the study by Chen which distinguished the flow experience in games from other flow activities through a metaphor often used by Csikszentmihalyi : Flow as driving a small boat in parallel to the current. Being able to drive freely gifts a sense of control over micro action, and being carried by the current offers a sense of control over the macro activity, therefore evokes Flow. [...] In video games, not only can players gain control from the progression, **they can also earn it through driving the boat, which is in fact making meaningful choices.**²⁰⁰ If the player loses his/her ego, without any opportunity to reflect on himself/herself during the gameplay, it's impossible for him/her to go through one hard challenge after another tougher one. Especially in «NieR: Automata» , there are some tasks that are not directly related to successfully killing the enemy, but instead providing moral options for players to choose from.

Having demonstrated the difference between these two concepts, I propose that instead of discussing the subject experiences which differ between individuals, it will be more beneficial analysing the objective infrastructures which enable these feelings, as presented in Table 2.

Environment	Objective	Subjective
Non-mediated & Mediated	Flow Activities	feeling of 'Flow'
Mediated	Immersive Systems	feeling of 'Presence' (supported by the fidelity to reality)

Table 2. Comparison between Flow and Immersion

¹⁹⁸ Csikszentmihalyi, *op. cit.* p.216

¹⁹⁹ Csikszentmihalyi, *op. cit.* p.141 *The emphasis is the original text.

²⁰⁰ Chen, *op. cit.*, p.13 *The emphasis is added by this study.

In spite of their distinctions, flow activities actually have the same structure in computer games, the mediated systems which intend to provide immersion. Csikszentmihalyi observed that 'a flow activity not only provides a set of challenges or opportunities for action but it typically also provides a system of graded challenges, able to accommodate a person's continued and deepening enjoyment as skills grow'.²⁰¹ Similarly, Chen also pointed out 'it is obvious that gamers value video games based on whether or not those games can provide Flow experiences (Holt, 2000)'.²⁰²

Moreover, Chen found that in order to make the game experience "fun" for everyone, 'to find exactly the right amount of challenge to engage with the exact abilities is the only way to access Flow'.²⁰³ This is supported by Ermi & Mäyrä's immersion research on their SCI-model, which revealed the challenged-based immersion is fundamentally built up by interactions and this feeling is 'at its most powerful when one is able to achieve a satisfying balance of challenges and abilities'.²⁰⁴ Both Chen and Ermi & Mäyrä's studies supported the necessity of providing diverse challenges for players to break through, either it is "complexed but negotiable" or it requests the motor skills to proceed in the game. Therefore, this chapter will be divided into two sections—the Tempo and the Structure of the gameplay—to analyse how the soundtracks of 《NieR: Automata》 help immerse the players into the challenged-based system, engaging them to make their choices while maintaining the flow experience in games.

6.1 TEMPO of gameplay

In Huiberts' research, many participants mentioned the positive influence of game audio on challenged-based immersion when corresponds to the player's action and supports the gameplay.²⁰⁵ Sonnenschein identified that music can alter listeners' perception of time by slowing down or speeding up the scenes via particular styles. Even though the

²⁰¹ Csikszentmihalyi, *op. cit.* p. 244

²⁰² Chen, *op. cit.*, p.3

²⁰³ Chen, *op. cit.*, p.7

²⁰⁴ Ermi and Mäyrä, *op. cit.*, p.8

²⁰⁵ Huiberts, *op. cit.*, p.69

objective time can be measured with clock, 'a brisk, repetitive march will quicken things, while romantic or New Age music tends to soften stress, relaxing the passage of time, even making it stand still,' as seen through research carried out by Sonnenschein.²⁰⁶ Not only does music affect players' subjective sense of time, but the emotional responses to music can influence their performances as well. Yannakakis, Paiva, Karpouzis and Hudlicka indicated that 'games can elicit player emotional responses which, in turn, may affect changes in the player's physiology, reflect on the player's facial expression, posture and speech, and alter the player's attention and focus level'.²⁰⁷

'Ermi and Mäyrä state that challenge-based immersion comprises both sensorimotor abilities and cognitive challenges. The distinction between the two is also found in Adams' classification of immersion', found Huiberts.²⁰⁸ Thus, we will adopt these two categories suggested by Huiberts to analyse different types of challenges, which demand different skills. That is, the **motor challenges**, such as reacting rapidly to specific events, and the **cognitive challenges**, including strategic thinking or logical problem solving in a puzzle game, for example.

Besides, Hevner's classic studies have demonstrated that 'mode, tempo and rhythm were the determinant cues for emotions'.²⁰⁹ This emotion perception in music is 'constituted by the main effects of the individual cues' and 'could be predicted with the linear combination of the cues,' analysed Juslin and Lindström.²¹⁰ That is to say, when the music fades in, the players are not only immediately aware of the emotional elements that it carries, but they channel their current mood swiftly in order to prepare for the coming events with the company of the whole composition. Furthermore, the

²⁰⁶ Sonnenschein, *op. cit.*, p.91

²⁰⁷ Yannakakis, G., Paiva, A., Karpouzis, K. and Hudlicka, E. (2011). *Emotion in Games. Proceedings of the 2011 Affective Computing and Intelligent Interaction Conference; Emotion in Games (EmoGames) Workshop*, Springer, 2011 Retrieved from http://yannakakis.net/wp-content/uploads/2014/07/Emotion-in-Games_CameraReady.pdf

²⁰⁸ Huiberts, *op. cit.*, p.69

²⁰⁹ Hevner, K. (1935). The affective character of the major and minor modes in music. *Am. J. Psychology*. 47, 103–118. doi: 10.2307/1416710 and Hevner, K. (1937). The affective value of pitch and tempo in music. *Am. J. Psychology*. 49, 621–630. doi: 10.2307/1416385

²¹⁰ Juslin, P. N., and Lindström, E. (2010). Musical expression of emotions: modelling listeners' judgements of composed and performed features. *Music Analysis*. 29, 334–364. doi: 10.1111/j.1468-2249.2011.00323.x

study by Eerola, Friberg and Bresin has shown that it is essential to analyse each music cue as a whole, instead of distinguishing specific elements from the whole composition. They explained that, 'certain important musical cues such as harmony, melodic contour, or dissonance could be of equal relevance for attributing emotions to music and were included within the music structure of our design without any systematic manipulation'.²¹¹ Therefore, I will regard each soundtrack as an independent music structure and examine them for the motor challenges and the cognitive challenges as follows.

I. Motor challenges

According to Huiberts, the motor challenges often appear in a high-paced games which need players to focus on the activity itself, or stimulate players to react rapidly, like the continuous threat of new approaching balls in the game *Zuma Deluxe* (Oberon Media, 2003).²¹² In such cases, players would much benefit from a brisk, speedy tune, which helps them concentrate on the fast-changing stimulus. Otherwise, the 'moderate music might cause the player to be disturbed by real-world thoughts, with a consequent decrease in concentration and performance', elaborated Huiberts.²¹³

Table 3 exhibits the BPM list (tempo) of combat music used in 《NieR: Automata》, in comparison with the tempo markings of Western classical music, based on the studies by Dr. Brian Blood.²¹⁴ The only exception is <3. *Peaceful Sleep*> which is categorised as location music in Table 1. Although it's indeed mainly used for presenting the Resistance Camp, there are a few tasks inside the Camp that the players are requested to complete. Thus, I included this tune on the list. 'According to Murray Schaefer (1977, p.227) this tempo is related to human heart beat ; a normal relaxed heart beat is 60 to 80 beats per

²¹¹ Eerola, T., Friberg, A. & Bresin, R. (2013). Emotional expression in music: contribution, linearity, and additivity of primary musical cues. *Frontiers in Psychology*, volume 4 (July 2013). doi: 10.3389/fpsyg.2013.00487

²¹² Huiberts, *op. cit.*, p.70

²¹³ Huiberts, *op. cit.*, p.72

²¹⁴ Blood, B. (2017). Music Theory Online : Tempo. *Dolmetsch Online*. Retrieved from <https://www.dolmetsch.com/musictheory5.htm>

minute'.²¹⁵ In other words, once the speed exceeds 80 bpm, any human would feel not relaxed, but instead tensed. Besides the slowest two tracks <**3. Peaceful Sleep**> and <**34. Alien Manifestation**> , all the combat music's speeds are above 100 bpm, locating at the quick range of expressions from Andante to Vivacissimo. This supports Huiberts' hypothesis that rapid tune helps the players succeed in the challenges which need motor skills.

For instance, <**5. Birth of a Wish**> is used at the battle with Adam & Eve (gameplay 2,4,10), Father Servo (gameplay 5, 11,12) and High-speed machine (gameplay 4,12). <**25. Dependent Weakling**> played during the final boss fight with Eve (gameplay 8,14,15), while <**37. The Sound of the End**> always appears during the most tough boss fights, like the battle between Engel and Pascal (gameplay18), or the heartbreaking moment when the virus-infected 2B had her last battle (gameplay16). The study by Collins discovered that 'the more senses pleased, the lesser external (real-life) impulses are noticeable', showing the "environmental functions" of combat music 'which contribute to a stronger connection with the game environment'.²¹⁶ Hence, the combat music <**10. Grandma–Destruction**> will be investigated in this section—"motor challenges", while the soundtracks of <**3. Peaceful Sleep**> and <**34. Alien Manifestation**> will be examined in the following discussion of "cognitive challenges".

Soundtrack	BPM (beat per minute)	Tempo markings (definition)
5. Birth of a Wish	166	Vivacissimo (very quick, faster than vivace) : no specified, but the higher level 'prestissimo' is 184-240 bpm.
10. Grandma–Destruction	148	Presto (very quick) : 100-152 bpm
25. Dependent Weakling	132	Allegro (quick, lively and bright) : 84-144 bpm
31. Song of the Ancients–Atonement	128	
21. Wretched Weaponry : Medium/Dynamic	116	
37. The Sound of the End	104	
34. Alien Manifestation	72	Andante (walking pace) : 56-88 bpm
3. Peaceful Sleep	64	

Table 3. Tempo of combat music

²¹⁵ Huiberts, *op. cit.*, p.70

²¹⁶ Huiberts, *op. cit.*, p.72

Grandma-Destruction

NieR:Automata

Kuniyuki Takahashi, Keiichi Okabe

INTRO Rhythm1

$\text{♩} = 148$

Chords: Em, Em, Em, B7

Rhythm 2 + Vocal VERSE A-1

Chords: Em, B7, Em, B7, Em

Melodic Fulcrum

Chords: D/F#, G, A dim, B7, CM7, D7

Rhythm 3 + Vocal VERSE A-2

Different chord progression, compared to A-1

Chords: B7/D#, Em, B7/F#, Em/G

1.

Chords: E7/G#, Am, B7, C

Figure 28. Score analysis of <10. Grandma-Destruction>
(The coloured notes are added by this study)

The image displays a musical score for the track <10. Grandma-Destruction>, consisting of four systems of piano and bass staves. The score is annotated with various elements:

- System 1 (Measures 24-26):** Features a piano melody starting with a B7 chord (circled in red) and a bass line with a B7sus4 chord (circled in red). A green arrow points to the chord progression from B7 to C.
- System 2 (Measures 27-30):** Labeled "VERSE A-3" with a "(with repeat)" symbol. The piano melody uses chords Em, B7/F#, Em/G, and E7/G#. The bass line continues with a consistent rhythmic pattern.
- System 3 (Measures 31-34):** Includes a first ending bracket (1. 3.) and a second ending bracket (2. 4.). The piano melody features Am, B7, C M7, and B7 sus4 chords. The B7 sus4 chord is circled in red.
- System 4 (Measures 35-38):** Features a second ending bracket (2. 4.) and a B7 sus4 chord (circled in red). The piano melody concludes with a B chord, labeled "transition to VERSE A'". The bass line includes a triplet of notes.

Figure 28. Score analysis of <10. Grandma-Destruction>

The track <10. Grandma-Destruction> appears at different scenes, including the cave where Adam first shows up (gameplay 2, 10), City Ruins (gameplay 3), Outer space (gameplay 14) and when the player flies the unit shooting above the ocean (gameplay 16). It also has a very distinguishable sound—a spinning sword—at bar 5, which comes in

right before the music goes to VERSE A-1, signalling that the war has begun. It also shows up at several particular boss battles, such as : Engels (gameplay 11), Auguste (gameplay 19) and the 2B model fight after 9S has broken through the Tower (gameplay 20). Interestingly, this track doesn't have VERSE B, even though it appears at so many occasions. The piece's main structure is **INTRO**→**VERSE A** (A-1+A-2+A-3)→**VERSE A'** (A-1*8 times)→ Repeat to **INTRO**→**VERSE A**→**VERSE A''** (A-3)→**VERSE A'** (A-1*4 times)→ Back to **INTRO** as the ending. However, this track never bores you and even allows you to be unaware of the repetition. How can it achieve such effect ?

Composer Toch discovered that 'the close recurrence of an expression affects us somehow unpleasantly and we prefer to substitute a synonym. The unpleasant effect, however, disappears as soon as we feel the repetitions to be planted with deliberation and purpose'.²¹⁷ Actually, the MONACA team indeed utilised many "synonyms" for this combat tune. Firstly, after the impactful 8 bars of **INTRO**, the chorus comes in immediately and plays through the whole **VERSE A** (A-1+A-2+A-3). Then, the chorus shows up at **VERSE A'** randomly as other instruments do (e.g., piano). This is the first varied flavour they added to distinct these two sections. Secondly, for the 8-time repeating VERSE A-1 at **VERSE A'**, the melody goes between high octave and middle range alternatively, accompanied by different harmonic choices of the same chords. Thirdly, it integrates three different rhythms in **INTRO** and **VERSE A**. Toch took the "Fate" motif of *Beethoven's 5th Symphony* and elaborated that 'such iterances, as we may call them, are of partly tonal, partly rhythmic nature'.²¹⁸ These three rhythms all use staccato on every eighth note which give the coherent atmosphere of quick and continuous strikes. Meanwhile, the different position of accents also makes them distinguishable from each other. The first rhythm at **INTRO** emphasises the 1st, 2nd, 4th and 8th beats, while the 3rd, 4th, 7th and 8th note are rest in the second rhythm at VERSE A-1. Then, when it comes to the third rhythm at VERSE A-2 and A-3, only the 1st, 4th and 7th beats are accented, as illustrated in Figure 28.

²¹⁷ Toch, *op. cit.*, p.120-121

²¹⁸ Toch, *op. cit.*, p.121

In terms of the functionality in games, the combat music must be "loopable" in gameplay ; other in-game music for menu screens, and final screens must also be "loopable", 'because it is difficult to predict how long a particular screen will be active, music with the capability of continuous repetition works great', explained the game composer Aaron Marks.²¹⁹ Except for the practical reason of saving the RAM processing power, "loopable music" is not really a difficult demand for composers. Marks indicated that 'the only real secret is finding the perfect spot after a bar or measure to abruptly cut off the music, enabling it to seamlessly begin again'.²²⁰ The <10. *Grandma–Destruction*> actually reveals the secret—using the **melodic fulcrum** at the end of each section.

According to Toch, 'the melody issues from a focal pitch (frequently but not necessarily the dominant) and returns to it repeatedly, moving either in one direction or in opposite directions, which makes the fulcrum more pivot or an axis'.²²¹ In <10. *Grandma–Destruction*>, the **dominant chord B7** (bar 8, 16, 24) and **B7sus4** (bar 26, 34, 36) function as the melodic fulcrums—the former makes the progression which tends to come back to Em, the beginning chord while the latter not only has the same chordal function as B7, but also refreshes the organised patterns by its ascending scale of sixteenth note. Furthermore, as marked by orange circles in Figure 28, the dominant chord B7 appears in diverse forms each time—from staccato to quarter note—which gives the phrase colourful transitions whenever it's going to repeat. Hence, melodic fulcrums help the composition proceed and keep the players in the fighting mood, by reducing the risk of tedium, thus even the VERSE continuously repeats. According to the inspection by Marks, 'any type of music can be looped, from orchestral to techno, with agreeable results. Longer loops are best, especially during gameplay'.²²² Indeed, the duration of <10. *Grandma–Destruction*> is 5 minutes 32 seconds. Even if the player can't finish the challenge in time, the exact same tune plays again 5-6 minutes later,

²¹⁹ Marks, A. (2009). *The Complete Guide to Game Audio*. (2nd ed.). Oxford : Focal Press (Elsevier Inc.).

²²⁰ Marks, *op. cit.*, p.235

²²¹ Toch, *op. cit.*, p.124

²²² Marks, *op. cit.*, p.235

instead of a 15-second loop for whole section. 'To enhance the player's immersion, music can be used to help the player through the stages of the game that are time-consuming or effortful', Huiberts concluded.²²³

Apart from the composition itself, Hevner also reminded us that 'any emotional evaluations are dependent on the context established by the musical materials in question'.²²⁴ As mentioned before, this tune is obviously used for crucial battles. Nonetheless, it is not merely a tune for combat. Following the request from director Yoko, who intended to build up a sorrowful world for 《NieR: Automata》, the MONACA team had to create sombre tunes in order to 'make the players feel sad whenever and whatever they do in the game'.²²⁵ Thus, 'even in the fierce fighting scenes, I would put some mournful elements into the music to accompany the battles', elaborated Okabe. He indicated that his general composition principle for NieR was to "describe the situation + the character's mood" ; 'especially this time, I put this **sense of sorrow** (哀愁感) in almost every tune, which is also my interpretation of the **Nier-like atmosphere** (ニ－アらしさ)'.²²⁶ In fact, <10. *Grandma–Destruction*> is the adapted version of <*Grandma*> which appears in the previous game 《NieR》, used against the super strong boss whose tactic is 'to distract and weaken you by making you remember some of your saddest and most painful memories'.²²⁷

For example, when <*Grandma*> plays, Kaine the protagonist of 《NieR》, confronts the toughest boss who pretends to be her grandmother. And 《NieR: Automata》 inherits the same context. The <10. *Grandma–Destruction*> plays when 9S confronts the enemies whose appearance is a crowd of 2B model—who was his dearly loved partner

²²³ Huiberts, *op. cit.*, p.70

²²⁴ Eerola, Friberg, and Bresin, *op. cit.*, p.10

²²⁵ Nagayoshi, *op. cit.*, p.8

²²⁶ *Loc. cit.*

²²⁷ Napolitano, J. (2010, May). Deep into NieR: Interview With Vocalist and Lyricalist Emi Evans. *OSV (original sound version) website*. Retrieved from <http://www.originalsoundversion.com/deep-into-nier-interview-with-vocalist-and-lyricist-emi-evans/comment-page-1/>

and had been killed in front of him. So, when we come back to <10. *Grandma–Destruction*>, we are supposed to consider it a tune written in E-minor scale, instead of G-major scale. The melodic fulcrum, chord B7, is the dominant chord (**V**) which harmonically tends to return to the tonic (**I**), chord Em. Meanwhile, chord B7sus4 as a suspended chord not only refreshes the rhythm via its ascending melodic minor scale, but it also creates the open sound through the lack of minor or major third—namely, the lack of full resolution.²²⁸ Therefore, the unresolved state is prolonged as the painful battle persists. Even the <10. *Grandma–Destruction*> doesn't include the vocalist, Emi Evans, who always puts in 'as much anguish and melancholy' as she could in her singing²²⁹, the music tonality has already given the clue, especially when the piece takes place in such heartbreaking context. Hence, we can clearly see the first distinct feature of the music in 《NieR: Automata》—that is, **combat with the sense of sorrow**.

Secondly, the other distinct feature of 《NieR: Automata》 is the **singing while killing** mode (or Uta-Utai mode in Japanese, ウタウタイモード), which comes from the previous game 《NieR》 as well. The so-called "singing while killing" mode actually originated from another action role-playing video game *Drakengard 3*, known as *Drag-on Dragoon 3* (or *DOD3*) in Japan, whose director was Yoko Taro and the composer was Keiichi Okabe—the same as 《NieR》 series. This mode suggests that the player can enter into the most powerful state to kill the bloody enemies as long as they are singing with the background music (BGM). So, this mode is also a type of weapon which increases the player's speed and attack capability and can be triggered when the player has been fighting in high-tension.²³⁰ Iwamoto observed that when 'aiming to keep the tension and not to interrupt the ongoing atmosphere, at some point, the game design would adjust itself to match the dynamics of music. For example, the <*Song of the Ancients–Devola*> and <*Song of the Ancients–Popola*> in previous game 《NieR》 and the featured setting

²²⁸ Suspended Chords. *MeAmBobbo's Guitar and Recording Wiki*. Retrieved from MAB Wiki website http://foobazaar.com/wiki/index.php?title=Suspended_Chords

²²⁹ Napolitano, *op. cit.*, p.5

²³⁰ SYSTEM. *Drag-on Dragoon 3*. Retrieved from Square Enix official website <https://www.jp.square-enix.com/dod3/sp/system/index.html> *The citation is translated from Japanese by this study.

of “Uta-Utai” mode in *DOD3*, I believe that the director Yoko must have understood the power of “song” very well, therefore continued to use human voices in most of the 《NieR: Automata》 soundtracks.²³¹ Indeed, all the combat music listed on Table 3 has a singing part, either vocal or chorus, including the 2 tracks (bpm 81) mentioned by Iwamoto which are adapted into the faster version <**31. Song of the Ancients–Atonement**> (bpm 128) for battle scenes in 《NieR: Automata》. In fact, not only does the combat music accompany the vocal singing, but all the soundtracks were requested to have vocals within the music as well. Okabe elaborated how he had dealt with the requests from director Yoko : ‘When we produced the previous game, Yoko had already told me that he wants to have vocals in every tune. However, I was not sure if it’s a good idea to apply to all the music. I even tried to persuade or coax him into pure instrumental versions.. [...] After we had finished 《NieR》, Yoko still persisted with his “all-vocal policy”. So, I had no choice but to “use all human voice in the tune” and twist my arrangement accordingly. For instance, using the pure instrumental composition in the main melody but with very soft chorus singing in the background. That’s how I fulfilled his orders to compose for 《NieR: Automata》’.²³² So far, we’ve seen the significant role of the human voice in the music for 《NieR: Automata》 –to keep the tension high during battles–especially during the motor challenges which demand that players focus on continuous in-game activities and speedy movements. Having demonstrated the two distinct features of the music in 《NieR: Automata》 –combat with the sense of sorrow, and singing while killing– I will now examine whether Okabe’s strategy goes well with the opposing style of missions–the cognitive challenges.

II. Cognitive challenges

Compared to the high-paced motor challenges in games, Huiberts notes that ‘games that mainly require motor skills offer fast music, accompanied with very direct sound

²³¹ Iwamoto, *op. cit.*, p.26

²³² Tada, T. and Sagako. (2016, August). 『ニア』を彩る音楽が生まれる地“MONACAスタジオ”ツアー。『オートマタ』のヴォーカル曲数は……全部!?. *Dengeki Online website*. Retrieved from <https://dengekionline.com/ele/000/001/341/1341859/> (The citation is translated from Japanese by this study, the same hereafter.)

signals, such as *Super Mario Bros* (1985) [...] [while] games that principally require cognitive skills, on the other hand, these tend to feature music with a more relaxed mood and more subtle sound design, making the experience more reflective and allowing the player to **concentrate on strategic planning**'.²³³ These cognitive challenges include strategic thinking and logical problem solving in a puzzle game. He mentioned the 3D artillery tactical game *Worm 4 Mayhem* (Team17, 2005) as an example, where the player can fire the weapon only once or twice in order to win. And it is accompanied by 'rather timid, almost ambient music, which supports concentration on thinking and determining the most successful weapon'.²³⁴ Therefore, the music for cognitive challenges usually has a relatively slow tempo that matches the human heart beat—60 to 80 beat per minute—in order to create a smooth and serene atmosphere where players calm down and stimulate reflections.

According to an investigation by Rona, 'a slow build tempo can add tension and energy to a scene in which something is gradually becoming more significant, such as danger'.²³⁵ As presented in Figure 29, the andante speed of <**34. Alien Manifestation**> (bpm 72) enables players to reasonably slow down their pace and start to pay attention to the environment with hidden threats. The first 8 bars of INTRO where there are very few elements in the music—rhythm in percussion and only one chord in electronic timbre, especially suggests that something is going to happen and players should stay alert. Although Figure 29 doesn't illustrate the full score, the dominant chord D7 of G-major scale doesn't show up. Instead, the **dominant chord B7**, suspended chord **B7sus4**, diatonic chord **Bm**, and the tonic chord **Em** of E-minor scale and its variations (e.g., Em7, Em add9, E) always occupy the beginning and the ending of a section. Hence, the tonality of this tune should be considered the same as <*10. Grandma—Destruction*>—they are both written in E-minor scale.

²³³ Huiberts, *op. cit.*, p.71 *The emphasis is the original text.

²³⁴ Huiberts, *op. cit.*, p.69

²³⁵ *Loc. cit.*

Alien Manifestation

NieR: Automata

Keiichi Okabe, Keigo Hoashi

INTRO Rhythm (percussion) at 2 octave below + Electronic chord at R.H → Trouble is brewing.

♩ = 72 E (omit 3)

mp (Crescendo at 2nd time)

4

Chorus comes in as melody.

E(omit 3)

D/E

7

C/E

Bm/E

Em add9

10

D add9/E

C/E

Bm/E

13

VERSE A-1 Rhythm remains + Theme1 in staccato at R.H + No chorus (melody)

Em7

Am/E

Figure 29. Score analysis of <34. Alien Manifestation>

(The coloured notes are added by this study)

15 **Em7** **Am/E** **Em7**

18 **Am/E** **Em7** **Am/E**

VERSE A-2 Rhythm remains + Theme1 in staccato at L.H (2 octave below) + Chorus (melody)

21 **Em7(9)** **D add9/E** **C add9/E**

24 **Em7(9)** **Em7(9)** **D add9/E**

27 **C add9/E** **Em add9** **Asus4** **Gsus4** **Csus4** **Em7(9,11)**

f Rhythm remains + Chorus (melody) + Theme2 in staccato at R.H (one octave above)

Figure 29. Score analysis of <34. Alien Manifestation>

30

32

35

38

Em7 (9,11) (Crescendo at 2nd time)

Am7(13)/E mp

Em7 (9,11)

CM7(13) GM7/B Am7 (9,11)

Am7 (9,11) F add9 F add9 F m6 /A b

VERSE B-1 Rhythm (bell-like) remains soft + No theme (staccato) + Focus on Melody (chorus) in quarter note.

VERSE B-2

[omitted here]

Pitched rhythm follows harmony.

Table 4 provides the outline of the structure of <34. Alien Manifestation>. By and large, the foremost rhythm begins at INTRO and plays through the end. There are few elements in INTRO, except the distinctive electronic sound at right hands, which helps build up a sense of impending doom. This sound actually reoccurs in later sections, both

in VERSE A and in VERSE B (circled in blue), suggesting that the enemy hasn't been defeated yet—there's still trouble looming on the horizon.

Next, compared to the high-speed and busy composition of <10. *Grandma–Destruction*> (bpm 148), the strong beats of staccato only show up in VERSE A to emphasise the themes (circled in orange and purple) in <34. *Alien Manifestation*> (bpm72). Meanwhile, VERSE B accentuates the expression of melody without the aid of staccato but with remaining rhythm which has changed from percussions to the bell-like sounds in soft volume.

Thirdly, Okabe elaborated how he had adopted this “all-vocal policy” from director Yoko: the chorus (human voice) appears as melody every 8 bars in general, then the melody goes by itself without chorus around 8 bars for example INTRO-1 to VERSE A-3, and VERSE B-1 to BRIDGE, with the intension of adding variations for the long battle.

Section	Rhythm	Staccato	Melody (with/out chorus)
① INTRO-1 (bar 1-8)	percussion at 2 octaves below	X	electronic chord at R.H. (X)
② INTRO-2 (bar 9-16)	at 1 octaves below	X	O (chorus=melody)
③ VERSE A-1	remains	Theme 1 in Staccato at R.H.	X
④ VERSE A-2	remains	Theme 1 in Staccato at L.H. (2 octave below)	O (chorus=melody)
⑤ VERSE A-3	remains	Theme 2 in Staccato at R.H. (1 octave above)	O (chorus=melody)
⑥ VERSE B-1	<ul style="list-style-type: none"> Bell-like sounds keep rhythm in soft volume. 	X	Focus on Melody (O) in quarter note
⑦ VERSE B-2	<ul style="list-style-type: none"> Pitched rhythm follows harmony. 	X	Focus on Melody (X)
⑧ VERSE B-3		X	Focus on Melody (O)
⑨ BRIDGE	A-1 + A-3		No chorus
⑩ VERSE B'	B'-1 + B-2 + B-3		
ENDING	A-1 + one extra bar ends in chord Em (the tonic) as a prolonged whole note.		

Table 4. Composition structure of <34. *Alien Manifestation*>

In truth, most of the battles in 《NieR: Automata》 are “motor challenges”. The “cognitive challenges” are fewer in comparison and most of the requests/tasks from other characters are accompanied by location music, such as finding “lost girl” for Big Sister machine in Machine Village (<17. *Pascal*>, gameplay 3), bringing the Child machine to the Mother machine at City Ruins (<2. *City Ruins*>, gameplay 5) and to fulfil the request of “family squabble” (gameplay 4). Nonetheless, it’s still essential to have combat music whose composition is moderately looser than the speedy and tense music of motor challenges. The other senses can only come in when there is a moment of rest during listening. Take <34. *Alien Manifestation*> as an example, this is the first piece of music that plays when players go through Menu and enter the game. As an opening title, the music should help players quickly grasp the story setting and what they should do. Therefore, in the first 7 minutes of the beginning, only the **INTRO-1, BRIDGE, VERSE B’ and ENDING** (i.e, the sections’ background in light blue at Table 4) of the tune play while the protagonist 2B fights against different types of battles in different spaces. For example, flying the unit and shooting above the ocean, shoot ‘em ups in the tunnel, and using sword for the short attack after she fell into the Abandoned Factory (gameplay 1–Prologue). That is quite a lot of information for players to comprehend in the first 7 minutes of the game, so the accompanied combat music should be as simple as possible, but still able to keep the mood on edge. That’s the first timing this tune shows up—without chorus, only the rhythm and soft melody with the electronic sounds.

The second timing is after 2B has the first rendezvous with 9S just after she’s finished fighting against a giant enemy (accompanied by <31. *Song of the Ancients-Atonement*>). Then, the same part of <34. *Alien Manifestation*> accompanies 2B’s exploring and her continuous fighting with endless enemies coming from the factory’s outside, stairwell, to inside (gameplay 1–Prologue, 9:14–20:54). But this time, it is with a stronger melody (**VERSE B’**) and random choruses—only when 2B enters one specific room inside the factory, will the **VERSE A-2** show up. Except that, there is no chorus during these long and non-stop battles, which is a necessary setup, because it’s very important for players to understand the relationship between 2B and 9S from the beginning. So, the comparatively loose composition not only creates room for the

weapons sound effects during the battles, but also allows the players to focus on their brief conversations in order to grasp protagonists' personality quickly. At this moment, 9S still politely called 2B "ma'am" and 2B always responds "emotions are prohibited" to any of 9S' question or say, excitement. The same part of music also plays in the gameplay 9 (17:00-22:07), the beginning of route B, when 9S and 2B "first meet" each other.

The third timing is when 2B first calls 9S by his nickname "Ninze" during their battles inside the castle of the Forest Kingdom (gameplay 5 of route A and gameplay 13 of route B). This is a considerably big move in their relationship. 9S had at this point previously tried to encourage 2B to call him by his nickname "Nines" at the Amusement Park (gameplay 3 & 11), but he was refused by 2B. Here, <34. Alien Manifestation> plays in full track and loops for around 10 minutes (gameplay 5, 41:11-51:19), following the composition order from **INTRO**→**VERSE A (A-1 to A-3)**→**VERSE B**→**BRIDGE** to **VERSE B'**. Given that the players have listened to the no-chorus version (i.e., section ❶→❷→❸) many times, it is reasonable to play the chorus parts (i.e., section ❹ to ❺) to strengthen the combat vibe through variations and more importantly, to enhance the emotional progression.

Finally, the last timing of playing <34. Alien Manifestation> is when 9S finds the TOP SECRET of the YoHRa project, which is also the climax of the game, right before the ending C, at gameplay 19 (35:23-45:32). It's worth noting that this 10-minute loop has a distinguished implementation to other tunes. The scene takes place inside the Resource Unit where 9S must hack into all the blocking devices and get the information. So, the first 8 minutes is all about hacking which has music distorted by the Tone Filter to match the 2D screen of hacking. However, the loop goes like this : **INTRO**→**VERSE A**→**VERSE B**, then repeats back to the chorus parts: **VERSE A**→**VERSE B**, without distorting any melody or chorus. Only the ambience, **Layer 1** of the tune, has been processed by Tone Filter. So, the players still hear some 8-bit sound in the background at the same time as the clean melody and original chorus during the hacking. This is the only part in the whole gameplay where it is implemented like this (see Appendix #5).

Furthermore, only the no-chorus section of **VERSE A-1** is playing when 9S reads the Top Secret at the hacking screen, which uncovers the cruel truth that the YoRHa black box circuitry is made from cores of the machines. Undoubtedly, this was a subversive secret to 9S and the players, which means the players can't help but rethink the logic of the plot setting—YoRHa infantry is built to protect humans who are still alive on the moon, thus YoRHa units must destroy the notorious machine lifeforms, created by the aliens, to regain the sovereignty over earth. Hence, the last 2 minutes of the loop need to stay simple to make room for this rather radical cognitive activity occurring inside the players' head. To sum up, this type of combat music is not just for the cognitive challenges but in most of the cases in 《NieR: Automata》, it is used to prepare the players for the vital moments during the battles, in order to proceed the narrative and let them reconsider the greater construction of the game design.

In fact, there is indeed a cognitive challenge in the Resistance Camp—the request of “Sorting Trouble” from the manager (gameplay 4). As Figure 30 exhibits, the player has to find a path between these stocks in order to reach the package (where the upmost arrow is located). It requires the player to have an overview of the dark space and figure out which path would be the easiest way to move those very heavy things. Therefore, a relatively calm tempo will support the player with the strategic planning. Here, **<3. Peaceful Sleep>** (bpm 64) plays. However, as shown in Figure 31, the most iconic sounds—the bagpipes (circled in green) and the vocals don't appear in this section.



Figure 30. Cognitive challenge—transport the package

Peaceful Sleep

NieR: Automata

Keiichi Okabe

INTRO Chant + percussion + Shaker (bar 4)

VERSE A Melody (Bagpipes for R.H.) +
Chant (bar 12 & repeat) + percussion

VERSE B

VERSE B Main melody (Lyrics in half note & whole note) + Melody (Bagpipes in quarter note) + percussion remains

Figure 31.
Score analysis of <3. Peaceful Sleep>
(The coloured notes are added by this study)

Instead, only the left-hand instruments (i.e., the bass clef) are playing from INTRO, VERSE A to VERSE B, accompanied by soft percussion to add some rhythm. Before the players are actually able to enter the storage and start this challenge, they have to pass through the music box to the Manager which they needed to obtain in Amusement Park

(gameplay 3). Once the player gives Manager the music box and finishes the conversation, the "music box" version of <3. Peaceful Sleep> will begin to play, meaning that the music has already filtered out other instruments and human voice—only the dreamy bell-like timbre represents the music. The "music box" version continues playing during their dialogue with the Resistance Member who demands the "Sorting Trouble". Then, the original music smoothly fades in once the player enters the room and starts the challenge. This extra "music box" version of <3. Peaceful Sleep> makes the transition very natural, which not only connects the section of cognitive challenge inside the storage at the backyard well, but it also successfully conjoins with the main ground of Resistance Camp. The Resistance Camp functions as the information centre where the players can gather several hints about other areas or receive tasks from the resistance members or Anemone, the leader of the camp. So, the music itself needs to leave room for the upcoming information. The players highly benefit from the slow tempo with the slackened composition when they stay in Resistance Camp.

Compared to <10. Grandma–Destruction> and <34. Alien Manifestation>, Figure 31 shows that <3. Peaceful Sleep> is a tune in the G-major scale. The **diatonic chord of G-major scale** gradually descends to the dominant chord D7 in VERSE A. Then, the diatonic chord of G-major scale gradually ascends again in VERSE B, like tides gently hit the shore. The tender and soothing atmosphere not only helps the players get rest and information, but it also enables the players to think profoundly about their dialogues with different resistance members. For instance, look at the following line by the weapon trader (gameplay 1) 'Although sometimes I wonder...what if my weapons are just making my friends die all the faster?', and this line by the supply trader who insists to keep his broken "original leg" in order to preserve his "true self" (gameplay 2 & 9), 'But this leg? This left leg? It's never been touched. So if I go and replace that, what happens? I mean, would I even be ME anymore?'. Briefly speaking, along with engaging players in crucial moments that proceed narratives, this "calming type" of combat music also supports them to succeed the cognitive challenges and obtain information. The space within the composition also enables the players to reflect on themselves which mingles their will with the characters' identities, choices and actions.

Before proceeding to the next session, we conclude this section on “TEMPO of gameplay” by noting the general composing strategy of Okabe. As seen in previous chapter, 《NieR: Automata》 uses vertical remixing method to compose multiple layers of music as one tune to adopt different situations in one area. Even when the players confront enemies and start fighting, the ongoing location music won’t suddenly change to another “battle tune” with a totally different structure and tempo. Instead Okabe indicated that with the compositions his biggest challenge was, ‘how to **keep the same tempo but provide different emotions**—fierce and moderate—at the same time. However, based on the premise of “not to change tempo”, what kind of new elements could be added to create such diverse effects also fascinated me as a composer’, was indicated by Okabe.²³⁶ However, he also mentioned the pros and cons of adopting this method : ‘On the one hand, people might think it sounds desultory when there is mainly one tune [remaining the same tempo] for one location ; but on the other hand, it would be also refreshing for players when they come back next time and discover the new variations of the same tune. That is part of the relish during the gameplay which I hope the players will experience’.²³⁷

6.2 STRUCTURE of gameplay

According to the results in Huiberets’ player questionnaire, the anticipation of music can increase the feeling of immersion. Several players pointed out that the background music ‘gives you a prepared feeling’ and ‘when the music changes pace you know something is coming’.²³⁸ In 《NieR: Automata》, <34. *Alien Manifestation*> has demonstrated this effect well. Additionally, Huiberts found that there are two audio techniques which enhance the challenged-based immersion—musical loops vary in time, and music follows the game progression. In truth, Okabe’s composing strategy proves the effectiveness of the first technique—musical loops vary in time. Huiberts observed

²³⁶ Sugihara, T. and Komori, D. (2016, August). 『NieR:Automata (ニール オートマタ)』の音楽はこうして作られる！ MONACAのスタジオに潜入、2バージョンのテーマ曲も公開(2/2). ファミ通 (*famitsu*) website. Retrieved from <https://www.famitsu.com/news/201608/17113095.html?page=2> (The citation is translated from Japanese by this study, the same hereafter.)

²³⁷ *Loc. cit.*

²³⁸ Huiberts, *op. cit.*, p.73

that 'the soundtrack uses several musical loops that vary in time and (the triggering parameter) threat causes a variation in the background music'²³⁹; this is a typical feature of the "vertical remixing" method in game music. As illustrated in Figure 15, the multiple layers of composition are additive and individually controlled. The Layer 1 (ambience/ drone) requires the fewest control inputs and generally plays from the beginning to the end of the section. Then, other layers have their own control input, 'allowing the layers to be faded in and out on their own—based on a specific gameplay scenario, such as running state, flying state, engage in combat, controlled by low health', analysed Sweet.²⁴⁰ Adding alternative layers to make the same tune sound different has two merits : **elevating positive user experiences** and **supporting the action flow**.

According to Huiberts' inspection, one player mentioned that 'music with spirit, with that X-factor, has a very addictive effect on me. I notice that I tend to play games with inferior gameplay more often when the music is good, just because the game experience in general is positively influenced'.²⁴¹ Also, because the same tune plays the whole time, 'there is little change in the music and there are only slightly different patterns used for variation, the music doesn't distract from the gameplay, while still supporting the flow of the player's actions', identified Huiberts.²⁴² The vertical remixing implementation creates consistent and appealing background music which assures players of the steadiness of battles and cohesion with the game plot.

The second audio technique to enhance challenge-based immersion is that music can follow the progressions of the game scenario with the aid of rhythmic/irregular sounds or repetition. For instance, when the player arrives in a new space, they start to explore the environment. Then, either the threatening opponents appear or some assignments are given within the space. Lastly, the most difficult challenge—the boss fight— shows up 'at the end of levels in action games which requires maximum focus of the player (cf.

²³⁹ Huiberts, *op. cit.*, p.74

²⁴⁰ Sweet, *op. cit.*, p.224

²⁴¹ Huiberts, *op. cit.*, p.75

²⁴² *Loc. cit.*

Rollings & Adams, 2003, pp. 306–308).²⁴³ Gameplay 7 of *«NieR: Automata»* can be cited as a good example. As Table 5 exhibits, when 2B arrived at Copied City a new place appears in game for the first time, she started to explore the place, trying to find any clues about 9S' whereabouts. However, she found nothing but many corpses of YoRHa infantry within the space and confronted the enemy—Adam. Thus, the player changes their interaction from exploration to combat mode. The previous two fights with Adam are accompanied by the location music **<20. Copied City>** with slight variations whereas in the final boss fight, the soundtrack is replaced by the combat music **<16. End of Unknown>** after 2B (the player) has seen that 9S has been crucified by Adam, which inspires a stronger motivation to beat the boss.

Player's actions	Game plot (time code)	Music
Explore	2B arrives in Copied City, looking for 9S	
Combat 1	2:28 Adam shows up and declaims : 3:13 WARNING 1 <i>"Some desired love! Others family! Only then did I realise the truth... The core humanity...is conflict.</i> <i>They fight. Steal. Kill. THIS is humanity in its purest form!"</i> 2B: "You know nothing about humanity."	< 20. Copied City > INTRO + VERSE A-1 (keep looping)
Combat 2	4:58 WARNING 2 6:18 2B: "I don't have time for you... I need...to find 9S ! "	< 20. Copied City > VERSE A-2 (with chorus)
Watch the cut scene	Adam presents 9S as being crucified for long time. 7:09 Adam: "I mean, after all...We all need something to fight for, don't we?"	No music
Combat 3	7:29 WARNING 3 10:11 2B beats Adam with sword. Music fades out.	<16. End of the Unknown>

Table 5. Scenario progression in Gameplay 7

It's worth noting that before the end of Combat 1, the background music doesn't change but keeps looping those 8 bars of **<20. Copied City>** from INTRO to VERSE A-1, as presented in Figure 32. Huiberts' study unveiled that certain repetition of sounds can make players comfortable, as players' attention is drawn to new things since the

²⁴³ Huiberts, *op. cit.*, p.73

Copied City

NieR: Automata

Keigo Hoashi

INTRO

♩ = 59

Musical score for the Intro of 'Copied City'. It consists of two staves (treble and bass clef) in a 6/8 time signature. The key signature has three flats (B-flat, E-flat, A-flat). The tempo is marked as quarter note = 59. The piece starts with a piano (*p*) dynamic. The first measure is marked with an 8-measure rest and the chord Fm7 (9,11). The melody in the treble clef features eighth-note patterns, while the bass clef provides a steady accompaniment.

VERSE A-1

Musical score for Verse A-1 of 'Copied City'. It continues from the Intro. The first measure is marked with an 8-measure rest and the chord Fm7 (9,11). The second measure is marked with the chord C^b6 / G^b. The third measure is marked with the chord Fm7 and is circled in yellow. The fourth measure is marked with the chord D^bM7. The dynamic is marked as mezzo-piano (*mp*). The bass clef features a rhythmic pattern with a '2x' marking, indicating a double-measure rest.

VERSE A-2 Chorus comes in

Musical score for Verse A-2 of 'Copied City'. It continues from Verse A-1. The first measure is marked with the chord Fm7 and is circled in yellow. The second measure is marked with the chord A^bM7(9) / B. The third measure is marked with the chord Fm7 and is circled in yellow. The dynamic is marked as mezzo-piano (*mp*).

Musical score for the end of Verse A-2 of 'Copied City'. It continues from Verse A-2. The first measure is marked with the chord D^bM7. The second measure is marked with the chord Fm7 and is circled in yellow. The third measure is marked with the chord A^bM7(9). The piece ends with a double bar line and the text '[omitted here]'.

Figure 32. Score analysis of <20. Copied City>
(The coloured notes are added by this study)

repeated things are already familiar (circled in yellow on Figure 32).²⁴⁴ Similarly, 'they will become more receptive to other inputs on an emotional level', identified

²⁴⁴ Huiberts, *op. cit.*, p.75

Sonnenschein.²⁴⁵ In fact, Adam had already shown his interest in "unravelling the riddle of humanity" when the player fought against him (and Eve) at gameplay 4. Since it's important to let players hear clearly what Adam found out about "humanity" during the battle, only the rhythmic sounds keep looping in Combat 1 without chorus. The chorus does not come in the tune until Adam has finished his declaration in Combat 2. Nonetheless, since too much repetition may cause irritation or boredom driving the player away from the scene, the accompanied music changes to another new tune <**16. End of the Unknown**> in the last battle—the most difficult one—to defeat the boss Adam. This example proves well how decisive the role of music can be within the dramatic storytelling ; 'there [must] be sufficient direction and evolution to keep the audience alert', highlighted Sonnenschein.²⁴⁶ In brief terms, in order to reinforce the structural changes of gameplay or to support the division levels of storylines, the music should 'correspond to the "peaks and valleys" of gameplay'.²⁴⁷ These "peaks and valleys" should not be limited to one scene but should take the whole structure of gameplay into account. A study by Lindley indicated that 'continuity of action within the representation of the story world of a film, theatre play, or computer game, representing causal interconnections within the diegesis, or story world [...] is a primary technique for the construction of the central conflict form at a detailed level'.²⁴⁸ Thus, when a causal interconnection happens, or say, a subversive change takes place in the game world, the music should denote these developments as well.

The Resistance Camp, for instance, has been always represented by the location music <**3. Peaceful Sleep**> (bpm 64) from its first appearance in gameplay 1. However, it is invaded by ferocious machine lifeforms in gameplay 8, which is very close to the ending of route A. Hence, the combat music <**5. Birth of a Wish**> (bpm 166) substitutes for the original location music—as it is not "peaceful" anymore. Besides, 'in computer action

²⁴⁵ Sonnenschein, *op. cit.*, p.116

²⁴⁶ *Loc. cit.*

²⁴⁷ Huiberts, *op. cit.*, p.75

²⁴⁸ Lindley, C.A. (2002). The Gameplay Gestalt, Narrative, and Interactive Storytelling. In Frans Mäyrä (Ed.), *Proceedings of Computer Games and Digital Cultures Conference* (pp.203–215). Tampere: Tampere University Press

games, the central conflict form usually has a recursive structure. The overall restorative three-act model is applied to the action game experience as a whole, and the dramatic arch is completed when the user finishes the game', noted Lindley.²⁴⁹ Although the "three-act model" does not really apply to the diegesis of 《NieR: Automata》, it does have a recursive structure which utilises three protagonists (2B, 9S, A2) in different routes (A, B, C) with alternative perspectives to present the narrative gestalt as a whole.²⁵⁰

The background music of Resistance Camp changing from <3. *Peaceful Sleep*> to <5. *Birth of a Wish*> is merely one plot progression in the first round of the gameplay (route A). As Lindley says, 'gameplay and the narrative gestalt formation processes are more fundamentally integrated. To achieve this, gameplay must be more than a repetitive interaction mechanism **for progressing through a larger scale but fixed and linear narrative structure**'.²⁵¹ Therefore, the "peaks and valleys" should be considered within this "larger scale" of narrative structure which is to say they should be 'typically constructed according to the conflict-driven model of dramatic narrative'. The best example is <17. *Pascal*> which not only functions as the location music of Machine Village but also presents the pacifist nature of Pascal. As mentioned earlier, the location music of Machine Village is substituted for combat music <21. *Wretched Weaponry : Medium/Dynamic*> in gameplay 18 because the machine lifeforms have fed on each other and A2 has come to save Pascal. Furthermore, when the perspective changes to another protagonist 9S in gameplay 19, <17. *Pascal*> doesn't appear to present the Machine Village, instead, the <21-1. *Wretched Weaponry : Quiet*> plays in the background when 9S arrives and has a conversation with Pascal whose painful memory has been deleted by A2 in gameplay 18. It's a logical and indeed appropriate musical choice, not only because it follows the scenario progression, but in a sense, the amnesiac Pascal is certainly not the previous Pascal. Both the original Pascal—the pacifist

²⁴⁹ Lindley, *op. cit.*, p.205

²⁵⁰ Lindley, *op. cit.*, p.204-205 * 'three-act model' includes ❶ a central protagonist, a conflict involving a dilemma of normative morality, ❷ a second act propelled by the false resolution of this dilemma, ❸ and a third act in which the dilemma is resolved by an act that reaffirms normative morality.

²⁵¹ Lindley, *op. cit.*, p.210 *The emphasis is added by this study.

machine lifeform who loved to read Nietzsche's philosophy—and the child machine lifeforms he helped to raise and educate, don't exist anymore. In other words, only after the players have finished the linear narrative of route **A** (gameplay 1-8) , route **B** (gameplay 9-15) and head for route **C** (gameplay 16-20), they are capable to see the entire story structure and experience the characters' emotions in depth.

In computer action games, the game level is designed for the pursuit of progressing the player through the recursive structure in order to puzzle out the big picture of the game gestalt. According to Lindley, 'a gestalt may be understood as a configuration or pattern of elements so unified as a whole that it cannot be described merely as a sum of its parts'.²⁵² The **narrative gestalt** unifies the game experiences into a coherent narrative structure, building up the comprehensive understanding in players' mind while the **game gestalt** could be understood as a pattern of interaction. So, being able to proceed to higher levels of the game could be seen as performing the (game) gestalt well. To some degree, they compete with each other for players' perceptual, cognitive, and motor effort. Even though there is the tension between the narrative and the gameplay, the investigation by Lindley revealed the significant relationship between these two gestalts. He discovered that 'gameplay gestalts may be highly demanding, and therefore highly immersive, but tend to be very shallow thematically, and performatively repetitive. These are positive values for many game players, frequently leading to addictive playing. There is nevertheless a widespread desire for greater thematic depth, even among dedicated gamers'.²⁵³

As far as the "thematic depth" is concerned, it is highly relevant to the setting of characters within the narrative gestalt—especially in the action role-playing game (ARPG) —which we will examine shortly. Having demonstrated how music supports the structural changes in games and how its tempo reinforces players' skills when overcoming various challenges, we can infer that an immersive computer game not only makes players feel "presence" in the mediated environments, but also continuously engages players in the

²⁵² Lindley, *op. cit.*, p.215 *The emphasis is added by this study.

²⁵³ Lindley, *op. cit.*, p.210 *The emphasis is added by this study.

action flow through corresponding music and sounds. As this section has shown, the audio can effectively help players succeed in the game mechanism—the gameplay, ‘a combination of resource management, strategic and tactical decisions’²⁵⁴. I would like to delve into how the soundtracks establish the emotional bonds between players and the game world, even building up the connection between unknown players, to make them not only “believe” but also “cry” when the imaginary immersion occurs.

7. Imaginary immersion

“Our bodies have an amazing plasticity and polymorphism that is often brought out precisely in our relations with technologies. We are bodies in technologies”— Don Ihde

According to Huiberts, the key for establishing the imaginary immersion is the player’s empathy with the game, and it could be enhanced with game audio through three aspects : Characters and events, the Setting and the Story.²⁵⁵ Thus, this chapter will be divided into three sections ; I’ll adopt an embodied cognition approach with philosophy theory to examine the “Characters and events”. Then, the importance of seamless ambient sounds will be discussed in the “Setting” section. Finally, in the “Story” section, five main endings will be scrutinised through different perspectives, such as music variations, structural functions and the narrative gestalt.

7.1 Characters and Events

Study by Brown and Crains found that empathy with the game character is very important during the deepest stage of immersion, Total Immersion.²⁵⁶ This is supported by Lindley’s research which argued that ‘the complexity and performative demands of a gestalt must lie within a particular range for a specific person in order for a game to be engaging and immersive’.²⁵⁷ This so-called “specific person” could be the protagonist, the characters, in an action role-playing game (ARPG). Sonnenschein discovered that

²⁵⁴ Lindley, *op. cit.*, p.208

²⁵⁵ Huiberts, *op. cit.*, p.82-83

²⁵⁶ Huiberts, *op. cit.*, p.82

²⁵⁷ Lindley, *op. cit.*, p.207

'our ability to listen is multi-focused, which means we can glean information through several different psychological and perceptual perspectives'.²⁵⁸ Therefore, he proposed the **referential listening** as the fourth listening type to complement Michel Chion's three other listening modes : reduced, causal, and semantic. 'Referential listening consists of being aware of or affected by the context of the sound, linking not only to the source but principally to the emotional and dramatic meaning', indicated Sonnenschein.²⁵⁹

Compared to being a passive spectator of film, being a player in the context of game demands much more information, helping players locate the enemies and know the plot progression at the same time. The former can be achieved by spatial audio design, while the latter requires the corresponding music to hypnotise the players into a plausible make-believe world. 'In all types of films, rather than supporting the realistic image on screen, the music allows us to sense the invisible and inaudible, the spiritual and emotional processes of the characters portrayed', highlighted Sonnenschein. This is consistent with the study by Yannakakis, Paiva, Karpouzis, and Hudlicka, which showed that 'characters provide a rich medium to express emotions, trigger emotions and adapt to the emotions of players'.²⁶⁰ In contrast to the distance between films and spectators, games 'let the players get under the skin of characters as we get to know their motivations, ambitions, sorrows, and fears'.²⁶¹ Hence, it seems promising to implement narratives and diverse themes into games to 'create deeper involvement through characters and emotions (Bizzocchi, 2008, p.4-5)'.²⁶² Furthermore, the research by Henry Jenkins distinguished the uniqueness of games among other narrative media (e.g., film, novel) : 'Narrative comprehension is an active process by which viewers assemble and make hypotheses about likely narrative developments on the basis of information drawn from textual cues and clues [...] spectators test and reformulate their mental maps of the

²⁵⁸ Sonnenschein, *op. cit.*, p.77

²⁵⁹ Sonnenschein, *op. cit.*, p.78 *The emphasis is added by this study.

²⁶⁰ Yannakakis, Paiva, Karpouzis, and Hudlicka, *op. cit.*, p.15

²⁶¹ Jørgensen (2010), *op. cit.*, p.315

²⁶² As cited in Jørgensen (2010), *op. cit.*, p.315

narrative action and the story space. **In games, players are forced to act** upon those mental maps, to literally test them against the game world itself'.²⁶³

Besides, the usage of characters is not only 'the driving narrative force in computer games'²⁶⁴, but can also 'work as tools to create emotions in games'.²⁶⁵ Since 'creating empathy between players and characters is important for this emotional bond', game designer David Freeman proposed that the goal of a game should be to 'move the player through an interlocking sequence of emotional experiences'.²⁶⁶ Lindley's study demonstrated that there is a strong tendency of such emotional bonds in ARPG : 'One inspiration for the *Purgatory* gameplay experience is live action role playing (LARPing).²⁶⁷ [...] The experience sought by LARPer is one of deep characterisation and intensity of emotional experience in character'.²⁶⁸ By offering 'more varied, subtle, and deeper emotional experiences' to players, the **bleed effect** would naturally occur between players and the roles they play in a role-playing game (Jeepforum n.d., Montola 2009, Waern 2010).²⁶⁹ 'When role-playing, the players will never be able to fully separate between the two identities even when they intend to be "in character"; the real world thoughts and emotions will increasingly bleed into the role identity, and vice versa, thereby making the distinction between player and role more and more transparent (Waskul & Lust 2004, p.349)', analysed Jørgensen.²⁷⁰ When the player's self-identity has intertwined with the role they play, the player is not just immersed into the world setting of the game, but is submerged in the mental world of the characters as well. In the audio realm of games, this phenomenon is represented through two aspects : Emotional responses to sound effects, and Practicing empathy via sounds.

²⁶³ Jenkins, *op. cit.*, p.126 *The emphasis is added by this study.

²⁶⁴ Sheldon, *op. cit.*, p.31

²⁶⁵ Freeman, D. (2004) *Creating Emotion in Games: the Craft and Art of Emotioneering*. Berkeley, CA: New Riders. (p.38)

²⁶⁶ As cited in Jørgensen (2010), *op. cit.*, p.317

²⁶⁷ A LARP is a form of improvisational theatre without an audience. *Purgatory* is a horror role-playing game, released in 2016. Retrieved from <https://rpgmaker.fandom.com/wiki/Purgatory>

²⁶⁸ Lindley, *op. cit.*, p.210

²⁶⁹ Jørgensen (2010), *op. cit.*, p.319

²⁷⁰ *Loc. cit.* *The emphasis is added by this study.

I. Emotional responses to sound effects : Lo-Fi (noise) & 2B's final battle

Huiberts' study showed that players consider sounds as an important information resource which indicates the state of the avatar (character), either it is a threat that might scare the players, or it is associated with the vitality of the avatar. Since the protagonists of 《NieR: Automata》 are all androids, the audio team utilises the "Lo-fi" plugin to add noise effect in different levels to present the avatars' current circumstances. It is generally used when the player encounters problems with the transmitter voice or when the player's senses are being jammed, such as : being hit by specific status or effects.²⁷¹ Given that the Lo-fi noise is essential for this game, this effect is inserted into many Audio Buses and the Actor Mixer, combined with other effects in the Multi effect plugin, such as : distortion, filtering and flanging.²⁷² However, to keep hearing the noise without making players feel uncomfortable is difficult. Hence, not only did the audio team apply this effect very lightly in most of the scenes, but also utilised the real-time audio for the sound-crafting.

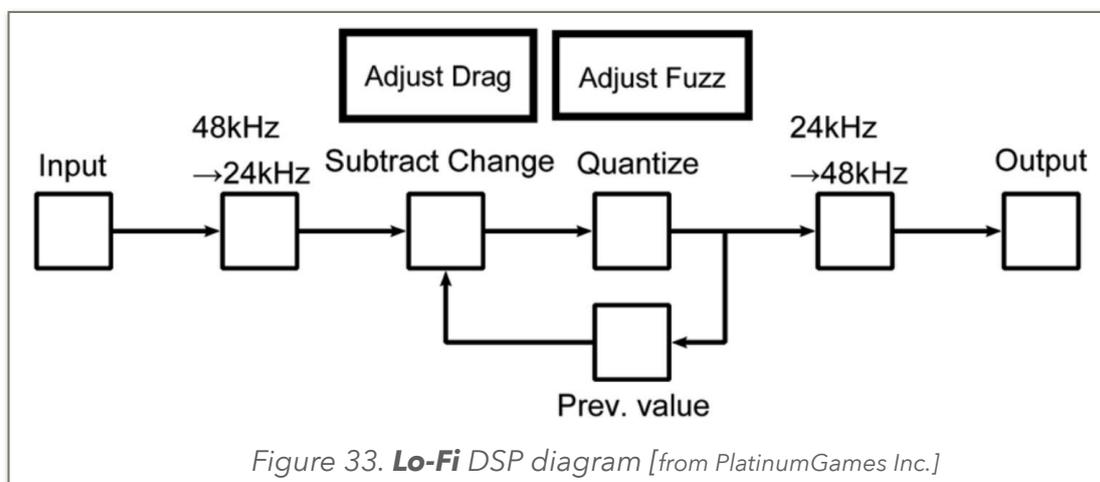


Figure 33 only exhibits the mono setup for the Lo-fi plugin processing, but it will duplicate based on how many channels are needed. The sound-crafting process starts from lowering the sampling rate to less than 50% or 24kHz, because Lo-fi sound doesn't need high sampling rates and this procedure can largely reduce the processing load.

²⁷¹ See footnote 104.

²⁷² Shindo, M. and Kohata, S. (2019, January 9). Part 2: The spatial acoustics of NieR:Automata, and how we used Wwise to support various forms of gameplay. [Web log post] Audiokinetic Blog. Retrieved from <https://blog.audiokinetic.com/the-spatial-acoustics-of-nierautomata-and-how-we-used-wwise-to-support-various-forms-of-gameplay-part-2/>

Figure 34. Noise & Screen flicker present **2B's vitality** (left: menu ; right: virus contamination rate 90 %)



Secondly, it makes the value closer to the previous value, and then is quantised. The more the quantising note is lost, the more buzzing is heard, so, this step helps prevent excessive noise and roughness in sound after the quantising. Finally, the sampling rate is turned back and we get the output. 'Even a standard process like this can make the game more immersive if you add an additional step or two', stated Shindo and Kohata. When it comes to the real implementation, the final battle of 2B in gameplay 16 (55:52-1:10:01) can be cited as the best example of how the amount of noise represents the vitality of the character. This scene takes place after the Bunker and Commander have been invaded and destroyed by the enemies. 2B put 9S into the escape route and entered herself into the combat zone. However, she had already been infected by circuit virus which was rapidly progressing in its own algorithm inside her body and was impossible to remove.

When the virus contamination rate was 16%, she can fight with countless machine lifeforms, and the player could hear the soundtrack clearly with the environmental sounds (e.g., waterfall, the heel's reverberation in the underground passage) when 2B ran inside the sewer. After the contamination rate rose to 21%, the noise effect got stronger and the screen slightly flickered. Interestingly enough, even a professional player like Shirrako, had to pause the gameplay twice in order to check if there is any problem with the hardware (1:00:04). Along with contamination rate increasing from 32% to 57%, the noise effect was heard more clearly while the soundtrack was playing meanwhile, the screen was full of the flickers and glitches making it harder and harder to recognise situations. The left screenshot of Figure 34 shows the menu interface after the contamination rate has reached 70%. It shows the attempt of the player (Shirrako) who

was trying to control but apparently failed. Then, the right screenshot is the final alert from Pod 042–2B’s circuit system has been totally corrupted and all of a sudden, she self-destructs herself on the bridge (1:06:56). Noises and flickers were shown on the screen, because of ‘the abnormality detected in visual sensor’ as Pod 042 said (see Appendix #6). Then, the player only sees the black and white screen without soundtrack or any environmental sound before A2 arrives to execute 2B’s infected body.

Aside from the technical part, it is noteworthy that the soundtrack *〈37. The Sound of the End〉* continues playing in the background from 2B’s landing at Flooded City to her automatic explosion (55:52–1:06:55). Even if the volume of noise keeps raising up, the music loop doesn’t cease but accompanies the player (as 2B) to go through the whole contaminating process. Considering there are countless fights happening during these 10 minutes, such as other sounds from the weapons, the changing environments, the noise, the dialogues between 2B and Pod 042—should one consider that too much stimuli with heavy information to players? Why did they still decide to reserve the room for music? Or in other words, what kind of functions of music are necessary for such busy scenes? What do they want to realise through the music, when the protagonist is approaching her end? Lastly, how do they achieve their aims and avoid the cognitive load at the same time?

The soundtrack *〈37.The Sound of the End〉* is probably the most essential combat music which not only appears at Goliath battles through 5 main routes (i.e., prologue, gameplay 6, 8, 9, 11, 14, 15), but also accompanies all the final battles of principal characters, including Pascal’s battle with Engles before his memory is erased (gameplay 18), the last fights between 9S and A2 (gameplay 20, 21) and the death of 2B (gameplay 16). In most cases, the score would play from INTRO, VERSER-A to VERSE-B²⁷³ and so on whereas, only here (gameplay 16) it sounds distinct from others. This is because it only plays **VERSE-C** and continues its 24 bars as a loop till 2B launches the self-destruction. The original tempo (bpm 104) has already set this combat tune to a “quick and lively”

²⁷³ In light of the space limit, the INTRO and VERSE-B are omitted in Figure 35. However, they’re in similar structure which consists of continuous sixteenth notes at high clef and accented eighth notes in Staccato at bass clef, as VERSE-A.

The Sound of the End

NieR: Automata

Keigo Hoashi

INTRO

VERSE-A
Em7(9,11)

VERSE-B
Em7(9,11)

VERSE-C
Em7(9,11)

CM7(9,#11,13)

[continues for 24 bars, omitted here]

Figure 35. Score analysis of <37. The Sound of the End>

(The coloured notes are added by this study)

pace, as shown in Table 3. The ceaseless sixteenth notes, performed by strings at VERSE-A and VERSE-B, make the score sound more edgy, meanwhile the constant eighth-note percussions in staccato strengthen the masculine spirits of the music, as if the fearless army can fire on enemies anytime. Nonetheless, in the case of 2B's battle, too many notes would distract the players from the current urgency—the decreasing combat capability, the deteriorating condition of 2B all while she still needs to fight against the approaching machine lifeforms endlessly. Moreover, there is high risk of muddying the whole sound palette with noise effect considering the relatively low register at VERSE-A and VERSE-B.

So, the musical choice is clear : the more stable fourth notes substitute for the restless sixteenth notes in VERSE-C with almost-the-same chord progression as VERSE-A. In addition to this slow and soothing melody, the bell-like timbre in higher register also gives the score some calm and feminine tenderness, in contrast to the hostile forces and the shadow of death, which softens the tense state. By doing so, the scene will benefit from this composition in two perspectives : **improving concentration** and **inducing certain mood toward characters**. Michel Chion's study on the "audiovisual contract" provided in-depth analysis on the linkage between sounds and images. He indicated that 'depending on density, internal texture, tone quality, and progression, a sound can temporally animate an image to a greater or lesser degree'.²⁷⁴ Chion proposes a few factors that influence how sounds temporalise images, and I would like to focus on two of those : how sound is sustained and the sound definition. 'A smooth and continuous sound is less "animating" than an uneven or fluttering one. [...] The second sound will cause a more tense and immediate focusing of attention on the image', elaborated Chion.²⁷⁵ Thus, using VERSE-C as a loop to accompany the battles not only provides a steady mental foundation to moderate the stressful situation, but it effectively helps the players pay attention on the changes of the scenes, including unexpected enemies, the stronger noise and increasing flickers on the screen.

Through the continuous alerts from Pod 042, the players are aware of the potential for deterioration and the unstoppable corruption of 2B. As mentioned before, players in the context of gameplay rely more on referential listening to catch information to help them react instantly. 'Often, these sounds induce a specific mood and increase the empathy of the player', discovered Huiberts.²⁷⁶ Through the dialogues between Pod and 2B, the players would realise the tragedy is unavoidable. Especially after the explosion, the screen colour was set to a very low saturation (nearly black and white), but the voice of 2B and the attacks also sounded jammed as it short-circuited (1:07:01-1:08:01). Then, the **VERSE-C** lightly played again while A2 listened to 2B's last words. That is, to make

²⁷⁴ Chion, M. (1994). *Audio-vision : sound on screen*. (Gorbman, C., Trans. & Ed.). New York : Columbia University Press.

²⁷⁵ Chion, *op. cit.*, p.14-15

²⁷⁶ Huiberts, *op. cit.*, p.85

the players see what 2B sees, hear what 2B hears, therefore they are capable of feeling the same helpless and hopeless state of the avatar—the bleeding effect occurs. This is why it is crucial to keep the soundtrack playing in such intense and busy scene.

Further, Chion identified that ‘a sound rich in high frequencies will command perception more acutely’.²⁷⁷ More specifically, the second point of using VERSE-C is that the **high frequencies** and alternative timbres are distinguishable from the noise effect successfully, therefore make each functional sound (e.g., attack, noise, music) balanced and aurally comfortable. The research by Kumar found that ‘spatial and immersive noise is key method for whatsoever preparation and instructive schemes researchers construct with gaming. [...] they can utilize sound fittingly and professionally while diminishing traverse modal sensory divergence. This skill replicates natural stereoscopic evaluation systems by utilizing computer ability’.²⁷⁸ The combination of immersive sound design, noise and music effectively supports the players’ imaginary immersion. Thus, the stereoscopic game experiences elicit true emotions within the interactive world of fiction. Following the death of 2B, there is one more point worth addressing here : the character’s motivations and reactions will determine the narrative directions and the game structure.

The scene that A2 executed 2B was witnessed by 9S and the players at the same time. However, the character, 9S, didn’t see what the players had seen before (from 2B’s perspective) and therefore, misunderstood it as A2 maliciously killing 2B. Hence, it boosted his hatred toward A2 and gave him stronger desire for revenge. This creates a plot twist in the story which is distinguished from route A and route B and unfolds a fresh development toward other endings. This is supported by Lindley’s study which discusses the intertwined nature of gameplay and narrative gestalt formation, ‘In this case, **actions become communicative acts**, and **the consequences of actions bear**

²⁷⁷ Chion, *op. cit.*, p.15

²⁷⁸ Kumar, P.S.J. (2018). Multiplayer Perceptron Neural Network Based Immersive VR System for Cognitive Computer Gaming. *Progress in Advanced Computing and Intelligent Engineering, Advances in intelligent Systems and Computing (AISC) Book Series of Springer*, 564, 91-102.

upon the state of development of a story'.²⁷⁹ Similarly, research findings by Jørgensen also pointed out that like other fictional characters, game characters 'may produce the same empathic responses in people as real people do (Lankoski 2007, p.6; Morrison & Ziemke 2005). This means that the player establishes empathic relationships with companions in both games, and this is what creates the powerful narrative effect'.²⁸⁰ To sum up, by relating character state development to patterns of character interaction, 'a gameplay gestalt may function like a chant or mantra, creating a form of dissolution of consciousness into the moment, acting against the strong incorporation of moments into an unfolding story structure'²⁸¹ –which urges players to react to what they have experienced through characters, thus responding to the true emotions of players themselves. Hence, the empathetic relationships between players and characters are established, accumulated and enhanced through the game progress.

II. Practicing empathy via Sounds : Pascal, Companions & Mirror Neurons

So, what would happen if there were no "emotional bond" between players and characters? Citing *Super Mario* as an example, Huiberts elaborated that while 'Mario has a certain appeal, in general, the designer do not add a lot of emotional meaning to this character. If Mario dies, the player might be annoyed for a short while, but mostly does not empathise with Mario's death'.²⁸²

① Music as the extension of Personality : Pascal

In comparison, the protagonists, the other companion characters and even enemies in *«NieR: Automata»* have their own plot or back stories which enhance players' understanding toward each character to arouse their feelings in different degree. Therefore, it successfully expands the narrative scale. For instance, in gameplay 18, Anemone asked A2 to bring some philosophy books to Pascal, saying 'It seems ...odd

²⁷⁹ Lindley, *op. cit.*, p.210, 214 *The emphasis is added by this study.

²⁸⁰ Jørgensen (2010), *op. cit.*, p.318 *The emphasis is added by this study.

²⁸¹ Lindley, *op. cit.*, p.213

²⁸² Huiberts, *op. cit.*, p.96

that a machine lifeform would be into philosophy, don't you think? It makes me wonder how much we really know about our enemy...' (30:08). The next scene then switched to Machine Village where Pascal was reading Nietzsche's book and educating child machines with kind words, such as : ' Well, look at you! Your speaking has improved so much. [...] You promised to read and memorised that botanical field guide, remember? ' or 'Uncle Pascal would never lie and break his word. Because children who lie grow up to be mean, evil machines! ' (31:00-32:39). The scene was accompanied by a mild soundtrack **<9. Voice of no Return–Guitar>** , a gentle tune with a serene melody (bpm70). The simplicity of this track matches the atmosphere of Machine Village very well, which not only adequately illustrates the pacifist personality of Pascal but also supports his interaction with children vividly. However, when it came to the end of gameplay 18, all the child machines chose to commit suicide because of "fear" which was caused by the very knowledge and emotions that Pascal had hoped would serve them well in the future to be a "good machine" (1:10:26-1:15:08). Consequently, Pascal requested that A2 (the player) either delete his memory or kill him. Without the previous plot showing his expectations on children, the players won't be able to feel Pascal's pain so strongly, and would therefore hesitate to make the decision.

In fact, there is a premise of creating emotional bonds toward characters and stirring up empathy within players. Smethurst identified that 'empathy and identification arise most easily when games contain characters that are relatable and interesting to the player, and the player-character is particularly important in this regard'.²⁸³ This is supported by Freeman's research which indicated that 'the characters must be deep and interesting. *Depth* refers to character complexity in terms of psychology and emotion, and *interesting* refers to the uniqueness, originality and imaginativeness of the character'.²⁸⁴ And the study by Jørgensen found that 'these attributes are very prominent in connection with companion characters'²⁸⁵, such as the example of Pascal in 《NieR:

²⁸³ Smethurst, T. (2015). *Playing with Trauma in Video Games: Interreactivity, Empathy, Perpetration* (Doctoral dissertation). Available from https://www.academia.edu/10765455/Playing_with_Trauma_in_Video_Games_Interreactivity_Empathy_Perpetration

²⁸⁴ As cited in Jørgensen (2010), *op. cit.*, p.318

²⁸⁵ *Loc.cit.*

Automata» . The studies by Smethurst and Freeman both refer to the general principle of designing good characters, including the **Protagonists** (or the Player characters), who are ‘controlled by the player, but can have their own motivations and personalities which are decided by the story’ (Egenfeldt-Nielsen et al., 2008), the **Companion characters** as Jørgensen calls them (or the Cast characters), who ‘have a personality and play some sort of role within the narrative of the game’ (Egenfeldt-Nielsen et al., 2008) ; the **Functional characters**, such as ‘non-descript shopkeepers who sell items to the player, or anonymous characters who can be killed for loot’ and the **Stage characters** who are ‘just part of the background and cannot be interacted with at all’ (Egenfeldt-Nielsen et al., 2008).²⁸⁶ Since the game gestalt depends on how well the players can control the main characters, it is easy for players to recognise the protagonists’ personality through their dialogues and interactions in the gameplay. Nonetheless, Jørgensen argued quite a few reasons that make the Companion characters indispensable to a good game. These are, enriching the narrative scale, making protagonists stand out, ensuring players’ actions, and activating new developments.

② **Silence with Subtlety** : Campaigns of Companion characters

To begin with, let’s take a look at Table 6 which demonstrates the whole campaign stories of Machine lifeforms, Bosses, and Pods. The topic of each campaign is presented in capital letters. After the players have finished route A (gameplay 1 to 8), the cut scenes of campaign stories start from route B till route E (gameplay 10 to 19). As mentioned earlier, a good game with a recursive structure must be “more than a repetitive interaction mechanism” in order to progress greater narrative scale which intensifies the thematic depth of the story.

Table 6. Campaign topics of Companion characters and Enemies

Gameplay no. main theme	Machine campaign	Boss campaign	Pod 042 & Pod 153 Data exchange
10. Adam Boss. Desert Zone	① the meaning of LIFE		① Agree to exchange regularly for efficiency
	② the shape of TREASURE		

²⁸⁶ As cited in Smethurst (2015), *op. cit.*, p.67

Gameplay no. main theme	Machine campaign	Boss campaign	Pod 042 & Pod 153 Data exchange
11. Beauvoir Boss. Engels Boss		i. Beauvoir –BEAUTY is what wins LOVE	
		ii. Beauvoir's story (full)–MEANINGLESSNESS [9:10–14:11 subtitle+soundscape]	
		iii. Adam & Eve –Undergarments for business, Eating fruits for intelligence.	
12. Adam+Eve Boss.Father Servo	③ Joy of Growing Up & Agony of Being Alone		
	④ 256 years ago–Kingdom as Family		
13. A2 Boss. Forest Zone	⑤ 128 years ago–Majesty's MEMENTO	iv. Adam & Eve –Reading books, instead of transferring data	
	⑥ 4 hours ago–PROTECT our baby king		
	⑦ always TOGETHER		
14. EMP Boss. Story Secrets	⑧ EMP story –Unwanted child crying for MOTHER	v. Adam & Eve –Eve's REVENGE * [Also appear at Gameplay 7 (1:3:37)]	② Situation data should exchange regularly
	⑨ Different Treasures–Hate		
	⑩ The Prophet [1/2 story]		
15. Eve Final Boss *Ending B		vi. Adam & Eve –Eve's CONSCIOUSNESS	
16. A2 Campaign story			③ 2B's death confirmed. Pod should support alive 9S and A2.
17. Hegel Boss			④ Select unit to monitor (9S/A2)–A2's memory region has combined with 2B's data.
18. Zombie Robots. Goliath Duel			⑤ Select unit to monitor (9S)
19. Operator Boss. Resource Units			⑥ Pod153 WORRY about 9S worsening psychological state.

Gameplay no. main theme	Machine campaign	Boss campaign	Pod 042 & Pod 153 Data exchange
			<p>⑦ How to handle 9S's deteriorated psychological state? NOT SURE what to do.</p> <p>⑧ 042: Protecting feeling—could this be...our WILL? DO NOT DIE.</p> <p>⑨ 153: Depola & Popola were reprogrammed to generate constant feelings of GUILT.</p>
20. 9S Final Boss (Ending C)		21. Ending D & True Ending (Ending E)	

Generally speaking, the **Machine lifeform** campaigns and **Boss campaigns** are elaborated in route **B**. The machine campaigns are more related to existential questions, such as the meaning of life, the dialectic between growing up and being alone, hate and protection while the Boss campaigns presents the ironic relationships between beauty and love, underwear and human business, and more importantly it illustrates the brotherhood between Adam and Eve. Both explain their histories, motivations and the reasons why and how an enemy becomes an enemy. The study by Lindley proved this is a productive way to achieve deep characterisation and heighten emotional experiences in characters, through the 'provision of rich media material for access to character back story, scenario information, memories, personality hints, associations, and the like. Most of this amounts to creating databases representing the psychic history of characters'.²⁸⁷

Aside from the back stories of enemies, the "data exchange" scenes between **Pod 042 and Pod 153** mainly take place **after route B**, continuing till gameplay 21 which surprisingly create the last main ending—the true Ending E. The interesting development of Pod 042 and Pod 153 demonstrates how 'informal chatter that emphasizes character depth, as well as suggesting that companions may also establish relations to each other

²⁸⁷ Lindley, *op. cit.*, p.212

and not only to the PC [Player characters]', as Jørgensen said.²⁸⁸ From Jenkins' viewpoint, this "embedded narrative" makes 'the game world appears as an information space in which the player gathers information that is distributed around the game environment'.²⁸⁹ Here, we can see the campaign stories of Machine lifeforms, Bosses, and Pods show **the first** essential attribute of companion character—**expanding the narrative scale by enriching its emotional texture.**

The interactions between Pods and the protagonists show **the second** essential trait of companion characters—**making the protagonist's personality stand out.** Jørgensen pointed out that even though the design of companion characters might be based on well-known templates and stereotypes, it is still possible to create impressive companion characters by 'giving them additional, and often surprising personality traits that are presented little by little throughout the game'.²⁹⁰ As low-level assistant androids, Pod 042 and Pod 153 are responsible for providing tactical support which is done through algorithm processing, launching long-ranged attacks, and exchanging operational intelligence between units.²⁹¹ Their mechanical voices, device-like appearances and dialogues (technical instructions) with YoRHa soldiers, all match the standard impression of robots—who don't have their own wills and desires, but just execute the commands they received. However, following the progress of game narrative, the "data exchange" campaigns gradually shows that they do have some "humanised feelings", such as confusion, concern, wish and protection. 'Psychological depth is then established through elaborate backstories, as well as character growth and development. Backstories have an important role in making the companions both deep and interesting, as their histories explain their attitudes and behaviour in ways that make them unique and original', analysed Jørgensen.²⁹²

²⁸⁸ Jørgensen (2010), *op. cit.*, p.321

²⁸⁹ As cited in Jørgensen (2010), *op. cit.*, p.324

²⁹⁰ Jørgensen (2010), *op. cit.*, p.318

²⁹¹ Pod 042. [n.d.]. *NIER Wiki*. Retrieved from Nier Fandom website https://nier.fandom.com/wiki/Pod_042

²⁹² Jørgensen (2010), *op. cit.*, p.318

Meanwhile, the development does not merely happen to companion characters, but also give some chemistry on their interactions with the protagonists. For example, in gameplay 20, after A2 chopped off 9S' left hand during the battle, A2 hacked in 9S' circuits in order to repair the virus-infected part of him. She asked Pod 042 to provide the location info of 9S' core data. Pod 042 replied that 'the idea under consideration is not recommended. This Pod is a YoRHa Technical Support Unit. As such, it cannot approve of actions that would damage its supported unit.' 'You're not so bad, Pod. You know that?' said A2. The soundtrack **<24. Mourning>** accompanied their dialogue in this scene before A2's monologue for Ending **C** (53:13-55:20). Compared to their first meeting in gameplay 16 (1:14:05), A2 was such unwilling to be "assisted" by Pod 042 and kept complaining about how useless Pod 042 is through the whole route C (gameplay 16-20), their relationship greatly improved after they went through so much challenges. Even though Pod 042's voice remains mechanical and cold, the players can obviously see that he is worry about A2's safety and A2 is actually not the cruel character she presents herself to be. 'Depth is hinted at in dialogue options, as each option reflects a different attitude and emotion, but it is up to the player to fill in motivations behind the choices that the PC [player character] is making. The PCs [player characters] are also made interesting only through the relationship to the companions in the games,' concluded Jørgensen.²⁹³

The third attribute of companion characters is **ensuring players' actions to follow the narrative logic**. Jørgensen indicated that 'providing supporting characters with motivations and agendas that may carry the narrative progression opens for a coherent narrative experience. [...] Through the careful scripting of the actions of supporting characters, the game designers are also able to monitor and manage player actions'.²⁹⁴ For instance, in gameplay 9 (3:19-8:14), gamer Shirrako has died in game once because he (9S) didn't follow the instructions of Pod 153, the companion character who is assigned to assist 9S. After receiving the mission from operator 210, 9S was supposed to meet the YoRHa troops who have commenced their descent. Therefore, Pod 153

²⁹³ *Loc.cit.* *The emphasis and [note] are added by this study.

²⁹⁴ Jørgensen (2010), *op. cit.*, p.328 *The emphasis is added by this study.

recommended 9S to take the flight unit. Maybe this is because it is the beginning of route **B** which is also the first time that players switch the protagonist perspective from 2B to 9S, new perspective brings new possibilities to explore the space. Shirrako didn't accept Pod 153's suggestion and instead intended to run down somewhere in the Abandoned Factory. Thus, he went directly to the Ending **G**–hun[G]ry for knowledge (4:07–4:28). When Shirrako restarted the game, the same plot took place again. Pod 153 had the same suggestion after 9S received the mission. This time, Shirrako didn't run down the factory but ran into the opposite direction which enters the factory. Pod 153 immediately gave another instruction–'Alert : Defence systems cannot be reached from the ground.' Having experienced the previous failure, Shirrako decided to give up trying and accepted Pod 153's suggestion of taking the flight unit (8:06–8:14). From this example, we can see players still have their own agency to freely move in the space. But without the aid of companion character, it is hard to put the narrative on track. For example, only when 9S takes flight unit to execute the mission, is he able to defeat the Goliath from the sky and meet 2B "for the first time" as the beginning of route A (gameplay 1). And the scripted narrative of route B can be successfully proceed.

Finally, **the last** and probably the most crucial attribute of companion characters in 《NieR: Automata》 is **activating new developments without making players feel powerless**. As highlighted by Jørgensen, 'companions are the most important devices for creating a richer narrative experience. While the player characters [protagonists] in both games are important for the growth and development of companions by working as advisors in personal and moral dilemmas, companions are scripted with particular potentials for development which the player may or may not activate'.²⁹⁵ Following the progress of "data exchange" campaigns between Pod 042 and Pod 153, the "humanised thoughts or feelings" emerged more within their databanks as shown in Figure 36. This development came up naturally, and forced the players to question whether it is possible for a Pod or an android to have their own 'will' which might be different from their creators.

²⁹⁵ Jørgensen (2010), *op. cit.*, p.327 *The emphasis and [note] are added by this study.

Pod 042: After repeated information exchanges between multiple Pods, an unexpected phenomenon has occurred.

Pod 042: We Pods have developed unusually protective feelings toward support targets 2B, A2, and 9S.

Pod 042: Query: Could this be...our "will"?

Pod 153: Unknown. The definition of "will" is unclear.

Pod 042: Will is the question of whether Pods are capable of their own self-determination.

Pod 153: ...

Pod 153: Even if it were possible, such actions cannot be abided.

Pod 153: Successful mission fulfillment is all that matters.

Pod 042: ...

Pod 042:

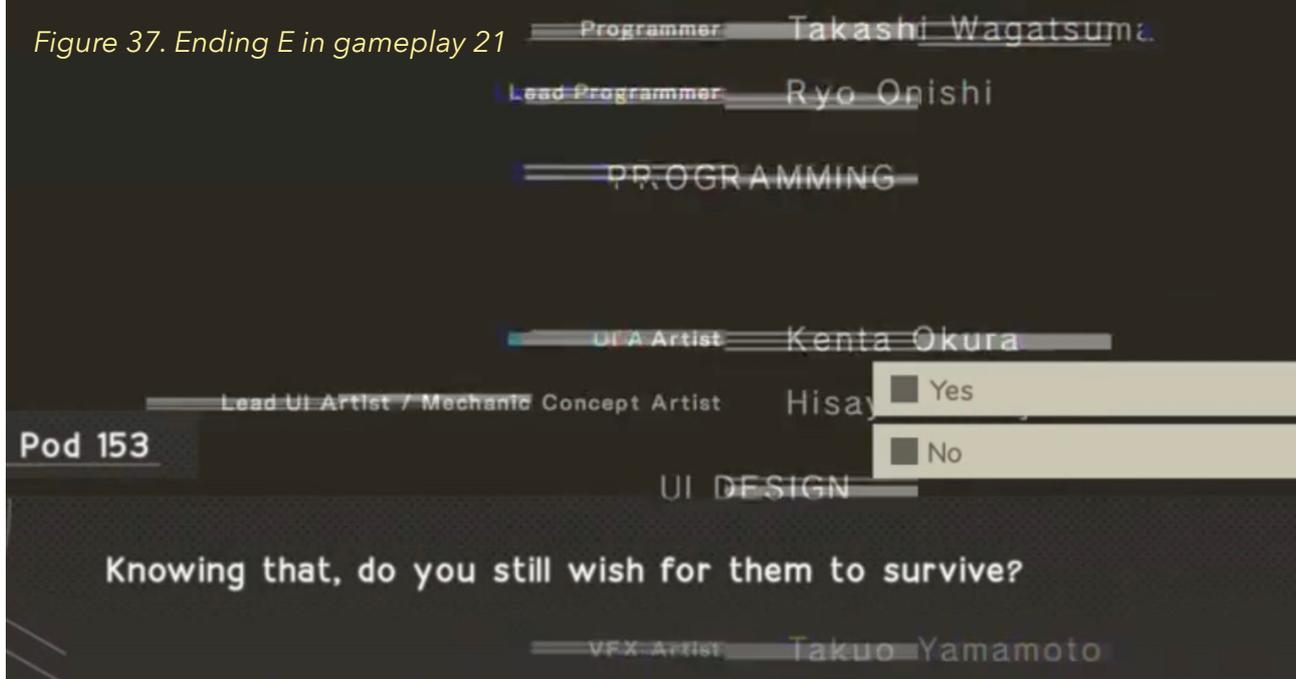
Pod 042: Whatever the case, we Tactical Support Units have a duty to see this through to the end.

Pod 153: Duty? Heh. You sounded like an android just there.

Pod 042: Agreed. But just as androids are influenced by humans, so are we tied to our creators.

Pod 153: Perhaps.

In the case of 《NieR: Automata》, all protagonists died in Ending **D** (gameplay 21), meaning that the players basically lose the right to change the narrative and it would be illogical to launch any other possible endings. However, this campaign foreshadows the alternative possibility—the Ending **E**, which was initiated by Pod 042 when the credit of Ending D was running (10:08-11:58). He refused the data-deletion request from Pod 153, which is part of a pod's "duty" as an essential step of YoRHa project. Despite that, Pod042 still "decided" to salvage 2B, 9S and A2's memory data because he cannot accept this resolution. Here, the players can clearly see that Pod 042's "own will" appears successfully and tried to make things different. Nevertheless, Pod 153 elaborated that as the pods, they 'lack the authority for such an action. The rules are protected by low-level systems. Salvaging data poses an unacceptable level of risk' (11:45). The most marvellous design of this twist is that there are Yes/No option for players to choose from



–the control power is cleverly transferred from the companion characters to the players, as presented in Figure 37. ‘In this sense, the game designers have effectively removed narrative control from the player without making the player feel powerless in the progression of events, and enabled narrative progression through using companions as distributors of narrative information’, summarised Jørgensen.²⁹⁶

As mentioned before, ‘seamlessness’ was one of the chief principles when designing the audio world of 《NieR: Automata》. In comparison with the main story plot, where there is always music or soundscape playing, the silent background of all the campaigns of companion characters and enemies is a valuable way to amplify the players’ attention, and to focus on the content of cut scenes instead of skipping over.

The study by Bridgett indicated that ‘well placed silence and subtlety will allow the parts of the game that need to be larger-than-life to be more powerful due to their contrast next to these areas of silence’.²⁹⁷ Let’s take the transition between Ending **D** and Ending **E** as an example. When it comes to the ending at each route, the soundtrack 《15. *Weight of the World*》 plays in the background accompanying the monologue of the protagonist and the credits for the production team. It’s the same pattern at **route A** (2B’s perspective), **B** (9S’ perspective), and **C** (A2’s perspective), with varied versions of the soundtrack. The ending **D** initially followed this pattern, with the tune 《15-2. *Weight*

²⁹⁶ Jørgensen (2010), *op. cit.*, p.327 *The emphasis is added by this study.

²⁹⁷ As cited in Huiberts, *op. cit.*, p.65

of the World Nouveau–FR version) playing with the credits. However, as shown in Figure 37, the conversation between Pod 042 and Pod 153 took place on the same screen at the same time. When Pod 153 noticed that some data noise was presenting in stream (10:04–10:18), she requested Pod 042 to perform the data check. Then the Yes/No option appeared on screen. If the player chooses Yes, the ongoing music will cease immediately and the running credit would be stuck as glitches, while the conversation between Pod 042 and Pod 153 continues. The visual subtlety effectively arouses players' curiosity and suggests that there seems to be something wrong, and the suddenly silent background makes players more concentrated on what Pod 042 is going to say (see Appendix #7). That is why Ending **E** is the last one among other endings—the most powerful twist caused by the most unexpected characters.

③ ***Witness as Experience : Mirror neurons & Body schema***

As far as empathy is concerned, the embodied cognition approach would help us understand how sound mediates our tendency to identify with the game characters. By examining how the **mirror neurons** and **body schema** operate on the human's identification, Collins indicated that 'identity is created from our embodied interaction with sound'.²⁹⁸ Her investigation provided a valuable viewpoint on embodied cognition, media studies and philosophy. **Firstly**, Collins suggested using the concept of "technological body", proposed by philosopher Don Ihde, to comprehend our body experiences nowadays. The concept of **technological body** was born to supplement other theories which didn't take human's relationships with technology into account, such as the phenomenological body by Merleau-Ponty (1998), and the cultural body by Foucault (1982). The "technological body" is 'characterized by our interactions with technology' from primitive to high-end tech tools.²⁹⁹ By using technology, the human body (both material and cultural) could be shaped by the 'instrument-mediated experience in which the instrument is taken into one's experience of bodily engaging

²⁹⁸ Collins, K. (2011). Making Gamers Cry: Mirror Neurons and Embodied Interaction with Game Sound [Conference paper]. *AudioMostly 2011*, September 7–9, 2011, Coimbra, Portugal. Retrieved from <https://www.researchgate.net/publication/220958353>

²⁹⁹ *Loc.cit.*

the world'³⁰⁰—or in Ihde's terms "extended embodiment". Within this embodiment of relations, 'the experience of one's body image is not fixed but malleably extendable and/or reducible in terms of the material or technological mediations that may be embodied'.³⁰¹ In the light of video games, Collins considered character as an instrument which could help the players engage in the game world with their body.

Secondly, Collins turned to cognition approach which illustrated how the mirror neuron system—a three-way mapping between sound, image and action—works in the brain. She observed that 'the same group of neurons will fire when either performing or observing an action [...], the neuron activity in our brain responds as if we are performing an action that we are not performing ourselves, but are rather **witnessing (through vision or sound)**. Our emotional and neurophysiological state can be directly affected by what we see'.³⁰² Hence, the neuroscientist V.S. Ramachandran affirmed that mirror neurons dissolve the barrier between the self and others, which are 'closely tied to—and possibly responsible for—our experience of empathy', analysed Collins.³⁰³ Moreover, the neuron activity of "witnessing as experiencing" doesn't merely happen in real humans, but can take place in **virtual characters** as well. As highlighted by Morrison and Ziemke, we 'can facilitate a user's identification with the character's "body" as well as provide the groundwork for empathy'.³⁰⁴ Additionally, the studies by Keysers et al. demonstrated that the same neurons fired whether an action is performed, seen or heard in the monkey's mirror neurons system.³⁰⁵ More specifically, these audiovisual mirror neurons respond as if we are experiencing the cause behind the event, **when only the sound of**

³⁰⁰ Ihde, D. (1986). *Consequences of Phenomenology*. New York State University of New York Press, New York, p.141

³⁰¹ Ihde, D. (1979). *Technics and Praxis*. Dordrecht, Holland, D. Reidel Publishing Group, p.508 *The emphasis is added by this study.

³⁰² Collins, *op. cit.*, p.2 *The emphasis is added by this study.

³⁰³ *Loc. cit.*

³⁰⁴ Morrison, I. and Ziemke, T. (2005). Empathy with Computer Game Characters: A Cognitive Neuroscience Perspective. In *AISB'05: Proceedings of the Joint Symposium on Virtual Social Agents*. AISB, UK, p.73–79

³⁰⁵ Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L. and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153, 4 (2003) 628–636. DOI= 10.1007/s00221-003-1603-5

the action is presented.³⁰⁶ The human brain responds to the audio stimuli 'as if it is also seeing and experiencing the action creating the sound', identified Collins.

Thirdly, the audiovisual "emotional contagion" can induce players' empathy through physiological changes. The term "emotional contagion" was proposed by Ekman et al., who suggested that 'we mirror the facial expressions on others that we see, and thus elicit the proper autonomic nervous system's response to the perceived emotion (that is, we feel empathy)'.³⁰⁷ Not only does the "emotional contagion" phenomenon take place visually, it also happens aurally through music. The research by Böttcher and Serafin found that the emotions within music could be recognised and felt by the listeners resulting in the physiological changes in the listeners, including the fluctuations on their blood pressure, heart rate, galvanic skin response, etc.³⁰⁸ Collins investigated how players' empathy is generated during the listening and the mental process : Our **mental re-creation of the music** causes a neuronal and motor-sensory response in ourselves that **mimics the performer**, and thus we are able to interpret the emotional inflections through our re-creation of the action.³⁰⁹ Therefore, she summarised that in the world of video games, 'we can use the knowledge of our visualizing causality and intentionality of sounds to better express the character's emotional state, so that the sound/music is more likely to impact upon our own emotional state'.³¹⁰

Finally, the mediacy of sound and music turns the characters into an "extension of the self" for players within the computer games. Collins based her research on two standpoints : **kinaesonic congruity in games** and **sounds in peripersonal space**. First of all, according to Yin-Poole, the so-called "kinaesonic" (kinaesthetic + sonic) refers to

³⁰⁶ Kohler, E., Keysers, C., Alessandra, M., Umiltà, L.F., Gallese, V. and Rinalatti, G. (2002). Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science*, 297, 5582 (2002) 846–848. DOI=10.1126/science.1070311 *The emphasis is added by this study.

³⁰⁷ As cited in Collins, *op. cit.*, p.3. see Ekman, P., Friesen, W.V., and Levenson, R. W. 1990. Voluntary Facial Action Generates Emotion-Specific Autonomic Nervous System Activity. *Psychophysiology*, 27,4 (Jul 1990) 363–384. DOI=10.1111/j.1469- 8986.1990.tb02330.x

³⁰⁸ Bresin, R. and Friberg, A. (2001). Expressive Musical Icons. In *Proceedings of the 2001 International Conference on Auditory Display*, Espoo, Finland, Hiipakka, J., Zakarov, N., & Takala, T. (Eds.). 141–143

³⁰⁹ Collins, *op. cit.*, p.3 *The emphasis is added by this study.

³¹⁰ *Loc. cit.* *The emphasis is added by this study.

'the physicalization of sound or the mapping of sound to bodily movements'.³¹¹ Although 《NieR: Automata》 is not a gesture-based game such as *Guitar hero* (RedOctane, 2005), this cognitive stance is still valid here to explain the inseparable relationship between players and their actions. Collins indicated that 'in kinaesonic congruence, we have a sonic reaction that matches the action that our body is making, thus embedding our own expressiveness into the game through the character. [...] even though we did not kinaesonically create them, we hear them as if we did. In other words, we have a direct, embodied interaction with the sounds that we evoke and hear in games, coupled with our physical and/or kinaesonic-congruent action'.³¹²

Secondly, a study by Cardinali, Brozzoli and Farnè identified that 'our body schema can extend beyond our physical body to incorporate non-corporeal objects'.³¹³ And this extended embodiments through objects—the 'instrument-mediated experience' as Ihde might put it—locates in the peripersonal space of individual. 'Peripersonal space is an intermediary space between our body (personal space) and what we see as the external environment (extra-personal space)', explained Cardinali, Brozzoli and Farnè.³¹⁴ Besides, the study by Knoblich and Flach observed that the action system of the human brain is 'responsible for creating an immediate sense of self by determining whether certain sensations and perceptions are the result of one's own actions'.³¹⁵ A good game should always have relative sounds that instantly respond to the player's actions of free will, 'even if that sonic response takes place in a virtual world, because it also takes place in the real world (our peripersonal space), we can integrate that sound as an extension of our body schema', was emphasised by Collins.

³¹¹ Yin-Poole, W. (2009). Muzyka: Dragon Age and ME 2 will make Gamers Cry. *Videogamer.com* http://www.videogamer.com/news/muzyka_dragon_age_and_me_2_will_make_gamers_cry.html

³¹² Collins, *op. cit.*, p.4 *The emphasis is added by this study.

³¹³ *Loc. cit.* see Cardinali, L., Brozzoli, C., and Farnè, A. (2009). Peripersonal Space and Body Schema: Two Labels for the Same Concept? *Brain Topography*, 21(3-4):252-60. DOI: 10.1007/s10548-009-0092-7

³¹⁴ *Loc. cit.*

³¹⁵ Knoblich, G., and Flach, R. (2003). Action identity: Evidence from self-recognition, prediction, and coordination. *Consciousness and Cognition: An International Journal*, 12(4), 620–632. [https://doi.org/10.1016/S1053-8100\(03\)00070-9](https://doi.org/10.1016/S1053-8100(03)00070-9)
*The emphasis is added by this study.

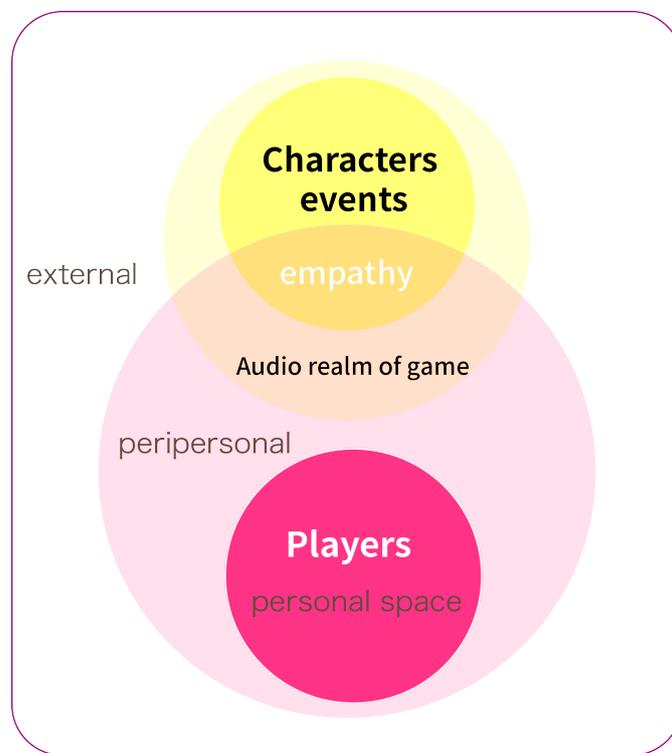


Figure 38. Sound and music turn characters into the "extension of the self" for players

Collins' argument revealed one crucial fact—that is, **the overlap of human's physical body schema and the imaginary world of game is achieved through sounds and music**. She summarised how the audio extends the player's sensations into the fictional world : In games, sound extends our sense of self beyond our physical body and into the intermediary space between ourselves and the virtual world. Sounds that we make—including in the virtual world—become a sensory extension of our self into that virtual world. The auditory realm of games thus becomes an extension of the self, a technological body through which we sense the game-world.³¹⁶ As Figure 38 illustrates, the pink circles show the player's body schema from personal space to peripersonal space (light pink). Meanwhile, the yellow circles represent the virtual world of game which is composed of the characters' actions, emotions, game events and the corresponding audio, including music and sound effects (light yellow). The imaginary world of computer game is situated in the external space of the players. More precisely, it is located at the mediated external space which requires players to undertake actions to enter the fictional world. As the action and sound is inextricable in games, the kinaesonic congruity makes the auditory realm the crucial pivot to bringing players into the imaginary kingdom. When players are aware that their wills could be embodied by

³¹⁶ Collins, *op. cit.*, p.4 *The emphasis is added by this study.

characters' bodies and actions, the immediate sonic response strengthens their confidence on battle performances, and enhances their own thoughts and feelings toward characters. That is, through sounds, the players not just enter the fictional world but proactively immerse themselves into the narrative, making decisions for characters and empathising with what they have gone through. Further, these embodied relations with the virtual world would also 'spill into our conduct in the real world', as evidences by the studies of *Second Life* (Linden Labs, 2003). 'Bailenson has shown that after only 90 seconds, the **interactions that take place virtually can elicit real behavioural changes** in our physical self', noted Collins.³¹⁷ This is supported by Slater's study on the sign of *presence* in an immersive system. He identified that 'when you are present, your perceptual, vestibular, proprioceptive, and autonomic nervous systems are activated in a way similar to that of real life in similar situations'.³¹⁸ That's why we laugh at comical scenes in simulative spaces and burst into tears when there is no chance to make our beloved characters survive.

Since the game audio is vital for the embodied cognitive connection between players and characters, Collins explicitly encouraged the usage of surround sound system. She believed it would amplify the gaming experiences and player's identification with characters when the spatial audio is provided, 'since we are then even more immersed in the sound, our own body in the middle of the peripersonal auditory space'.³¹⁹ Given that a game character is an extension of player's self in games, voice acting becomes an essential tool for involving players into the character's heart. However, setting a standard to evaluate the voicing acting is a controversial matter, because the perception of voice acting depends on each individual's taste. 'The voice is so intimately personal, that the "wrong" voice can destroy our illusion of the character,' argued Young.³²⁰ Even though the voice acting is excluded from the scope of this study, 《NieR: Automata》's approach of **not using voice acting** at certain moments is worth to mention here. One example is

³¹⁷ Collins, *op. cit.*, p.7 *The emphasis is added by this study.

³¹⁸ Slater, *op. cit.*, p.2

³¹⁹ *Loc. cit.*

³²⁰ As cited in Collins, *op. cit.*, p.6

*Do these clothes make me look cute?
Would he like it if I walked like this?*

Figure 39.
Beauvoir's campaign

*I still don't understand what it means to love someone.
But I've made up my mind.*

I will do whatever it takes to capture his affection.

the campaign story of Beauvoir at gameplay 11 (9:10-14:11). Compared to the "theatre-like" styles of machine campaigns and Pods' data-exchange, Beauvoir's campaign is the first one that tells the story without any concrete character or image, but via typing text, as Figure 39 exhibits.

All the campaign stories are silent, her story explaining why she is eager to become beautiful only has the soundscape to accompany till the end. One possible reason is that it is a **rather long** campaign lasting 5 minutes and the sparse soundscape can help the typing sound more soothing. It's unnecessary for Beauvoir to say one word, because 'the player has already read the words on-screen, and thus already "heard" (and re-enacted) the voice in their heads', indicated Young.³²¹ The second reason to have soundscape here is for **inducing a certain mood** toward the character, therefore some related **issues are emphasised**. Actually, the soundscape doesn't appear when the campaign starts, but gradually fades in when the line 'I still don't understand what it means to love someone' is typed. By doing so, the topics of "loving someone" or "how to be loved by someone" and her solution—becoming beautiful—not only fulfils the campaign's function of telling character's history, but also successfully arouses the players' attentions, doubts, even reflections.

The other prominent example is in gameplay 21, 9S' monologue at **Ending D** (5:08-8:30). It's also shown via typing text on the white screen, without 9S' voice or image. Actually at this moment, after the extremely radical battle between A2 and 9S, presenting any image would be redundant. As the fatal system has been destroyed and

³²¹ Young, K. (2010, May 16). Voice in Bioware's 'Dragon Age: Origins'. [Web log post] GameSound.org Blog. Retrieved from <https://gamesound.org/2010/voice-in-biowares-dragon-age-origins/>

9S is dying, using the fading text in white screen to visually represents his fragmented memories is a pretty adequate choice. Meanwhile, the intention of using sounds to induce player's empathy is more obvious here because this time, the text-typing sound is not accompanied by soundscape but by the soundtrack **〈14. Vague Hope–Cold Rain〉** with the gradually faded-in vocal until the end of 9S' last words. The biggest merit of eliminating the voice of the character is that it enables the players to embed themselves into the character through imagination. As highlighted by Young, 'If we are to become our character, shouldn't they sound like us, or at least how we imagine they should sound? [...] by not hearing a prescribed character voice which takes them out of the experience, the player is empowered to fully inhabit their character'.³²² These two examples demonstrate well the power of imagination—even though there is no voice acting, the sound and music gently guide the players into character's mind, resonating with their real-world experiences deeply.

7.2 Setting

"Everyone is fragile, not that strong. And NieR:Automata is about the world how these frail characters' stories interweave together."— YOKO TARO

According to the investigation by Huiberts, sounds can effectively enhance the believability of the world setting of games. He distinguished between the visual and the audio, pointing out that there are 'specific qualities that audio has over graphics in the realm of imagination and credibility'.³²³ Following our previous discussion, the penetrative feature of game audio plays a pivotal role in bridging the players' imagination and body schema via their control on characters within the fictional world. 'Especially pit music, can function like the spatiotemporal equivalent of a railroad switch,' discovered Chion.³²⁴ From Chion's point of view, compared to other sound and visual elements which are 'obliged to remain clearly defined in their relation to the diegetic space and to a linear and chronological notion of time', music enjoys the status of being

³²² *Loc. cit.* *The emphasis is added by this study.

³²³ Huiberts, *op. cit.*, p.83

³²⁴ Chion, *op. cit.*, p.81 *The emphasis is added by this study.

a little freer from barriers of time and space. Therefore, the nature of music enables itself to be the '**cinema's passe-muraille**, capable of instantly communicating with the other elements of the action'.³²⁵ We will discuss how music functions as 'the passer-through-walls' (passe-muraille) to build up the believability of game world through two aspects in this section, that is—the World setting, and the Magical circle.

I. World setting

Huiberts and Collins both pointed out that audio can be used to induce specific settings in games. Along with communicating meanings with music, the mood induced through music can control the player's emotions. Furthermore, the induced mood/atmosphere is 'perceived rather unconsciously and manage to convey the Setting instantly'.³²⁶ So, what is the "setting" in 《NieR: Automata》? According to director Yoko, the point of the previous 《Nier》 was that the player would find that they are not human by the end of game whereas in 《NieR: Automata》, it's clear at the beginning that none of the protagonists are human, and neither is the enemy. 'However, to some extent, I found these non-human characters are still like human when I completed the script. What I'd like to emphasise is that everyone is fragile, not that strong. And *NieR:Automata* is about the world how these frail characters' stories interweave together'.³²⁷ Before we probe into how game audio induces this "non-human but fragile" setting, it is essential to see how Yoko's vision carries through the whole world design realised by the avatar modelling and their interactions with functional characters.

I-1. Contradictory Aesthetics : family, war, violence

When it comes to the avatar modelling, people often focus on the design of protagonists whose visual images should represent the world concept of game, and a good game design should always provide corresponding audio effects and music to accompany their (players') actions. For example, with 2B, the character modelling artist

³²⁵ *Loc. cit.* *The emphasis is added by this study.

³²⁶ Huiberts, *op. cit.*, p.86

³²⁷ 奎力, *op. cit.*, p.6

Hito Matsudaira mentioned that ‘there’s a fragile beauty to it (儚い美しさ), a doll-like form that feels like it could easily break. [...] the model looks more beautiful with a dash of oddity’.³²⁸ His action of increasing 3D model’s appeal is consistent with Yoko’s idea. In addition, Matsudaira observed that the interactive feature of game can easily elevated character’s charm in the mind of a player through their control on characters. In the meantime, player’s memories and emotions that come from experiencing the game also help.³²⁹ We’ve seen many examples how soundtracks help boost player’s empathy with characters in previous chapters, so I’d like to emphasise the importance of the **sound modelling of machine lifeforms** that delineates another perspective of Yoko’s vision of a “non-human but human-like world”.

According to sound designer Shindo, their goal of sound design can be divided into two parts : **First, to present androids**—the protagonists who are supposed to be close to human—through futuristic action sounds without warmth. **Second, to present machine** lifeforms—the enemies who are assumed to be the totally opposite existence from human—via retro sounds that make people feel nostalgic.³³⁰ For instance, at the beginning of route **B** (gameplay 9, 0:20–3:08), one machine lifeform is trying to save his “dead brother” by giving him the oil. The foley sounds of how this machine jumps, walks, and fills the buckets with oil were recorded with real iron plates at the TOEI Digital Center. In order to make these metals sound rusty and retro for the machine lifeforms, the audio team also recorded the friction sounds by rubbing the floor with different desks—a common sounds you often hear at elementary schools.

Different field recordings of real objects were mixed to create complex foley sounds, which were sometimes overlapped with synthesised sounds and sometimes combined with environmental sounds for the cut scenes. Furthermore, the Lo-fi noise we discussed earlier was also ‘created by studying the composition of old digital recording

³²⁸ Matsudaira, H. (2017, April 3). An exclusive look at the creation of Nier:Automata’s 2B. [Web log post] PlayStation.Blog. Retrieved from <https://blog.playstation.com/archive/2017/04/03/an-exclusive-look-at-the-creation-of-nier-automatas-2b/>

³²⁹ *Loc.cit.* This article has Japanese version as well. Matsudaira, H. (2017, May 30). 『NieR:Automata』 キャラクターモデルメイキング. [Web log post] Platinum official website. Retrieved from <https://www.platinumgames.co.jp/dev-nier-automata/article/106>

³³⁰ Shindo, M. (2017, Oct 3). 『NieR:Automata』 イベントシーンの効果音制作現場. [Web log post] Platinum official website. Retrieved from <https://www.platinumgames.co.jp/dev-nier-automata/article/240>

machines'.³³¹ To ensure the "a dash of oddity" in 2B's beauty, they use "contradictory aesthetics" to not only represent the machine lifeforms with "nostalgic sonic image", but also implement it on their 3D modelling—the rusted metal robot with a round head which can turn softly like a lively bird.³³² This consistency from the audio to the visual makes the characteristic of machine lifeforms unified and distinctive. Moreover, their ambiguous characteristic design and actions serve as the perfect foil for the androids' coldness. In fact, there are tons of requests from machine lifeforms that exhibit their "humanised relationships", within the game, such as the "Lost Girl" request from Big Sister machine (gameplay 3), "Family squabble" request from Mother machine (gameplay 4), Overprotect machine with her scary son machine (gameplay 13), and so forth. The most interesting scenario is in gameplay 5, an Animal-Loving machine elaborates his/her passion for taking care of injured animals when 2B and 9S try to kill them all : 'LONG AGO, I RESCUED A HURT ANIMAL.[...] SHE WAS... WARM.SHE MADE ME HAPPY. NOW, IF I FIND A CREATURE THAT IS WEAK OR INJURED, I BRING THEM HERE.I WATCH OVER THEM UNTIL THEY ARE HEALTHY AGAIN'.³³³ 'Interesting,' coldly replied 2B (21:43-27:45). Compared to the all-time-killing androids, the machine lifeforms' requests force players to wonder who is more like human being?

To some degree, this blurred line between androids and machines implicitly foreshadows the top secret of the story. Meanwhile, it induces players to re-think those basic concepts we're used to believing in, such as family and sibling, gender, capitalistic systems and mutual understanding through network. As highlighted by director Yoko, 'I wanted to avoid questions like, "What does it mean to be human?" [...] Instead, the characters ask questions about war, and family [...] They also ask questions about violence. Is killing justified, or worthwhile? Are my enemies sympathetic, or just monsters?'³³⁴ The usage of music also showcases Yoko's emphasis on these topics : the

³³¹ See footnote 104.

³³² Kijima, H. (2018, June 19). 『NieR:Automata』のメカデザイン:機械 生命体編. [Web log post] Platinum official website. Retrieved from <https://www.platinumgames.co.jp/dev-nier-automata/article/351>

³³³ The original text is written in capital letters.

³³⁴ Muncy, J. (2018). Behind the Mask With Yoko Taro, Videogames' Most Interesting Designer. *WIRED website*. Retrieved from <https://www.wired.com/story/yoko-taro-nier-automata/>

ballad *〈9-1. Voice of no Return–Normal〉* comes in when Animal-Loving Machine addresses his/her reason for begging for 2B's mercy, while *〈13. Treasured Time〉*, the lullaby-like tune sung by an innocent kid, plays in the background when the Mother machine expresses how hard it is to communicate with her child : 'I never imagined how hard it would be to disconnect from the network. We're constantly misunderstanding each other... Still, **on those rare occasions where we connect**, it makes me so very happy ' (gameplay 5, 1:19:28).³³⁵ Compared to other soundtracks that frequently show up to accompany different scenarios, these two tracks only appear here to elicit our identification with machine lifeforms with sympathetic mood. We can assume that Yoko not only intends for the players to empathise with the protagonists, but also attempts to give them more food for thought from the enemy's perspective.

In truth, Yoko's ambition of stirring players' thoughts permeates nearly every interaction that players might have, even with functional characters, such as the weapon trader, or the supply trader. For instance, the dialogues between the weapon trader, 2B and A2 at gameplay 1 and 18 demonstrate how the "embedded narrative", suggested Jenkins, and enriches the world setting under Yoko's direction. The weapon trader once told 2B about his confusion of selling weapons, he said : 'Although sometimes I wonder...What if my weapons are just making my friends die all the faster?' Although there is no request to complete, it plants a seed in player's mind. Then, following the development till gameplay 18, when A2 brought 2B's sword to visit Resistance Camp, the weapon trader realised what it is at first sight and said, 'YoRHa weapons save the external memories of whoever uses them. Please take extra care of...her'. The corresponding plots and brief dialogues do not merely strengthen the debatable meaning of war through 2B's death, but interlock the different routes (A, B, C) firmly forming a greater narrative gestalt. Further, this detail provides more clues and room to interpret Ending **D** –9S killed A2 but he was also stabbed by 2B's sword–nobody survived.

³³⁵ The emphasis is added by this study.

I-2. Comparison with previous 《Nier》 : balance, similarity, diversity

Having known the concept behind the world setting of 《NieR: Automata》 , distinguishing it from previous 《Nier》 games would help us better capture the affections within the game and how music supports it. Okabe noted that 'it's important to keep the "organic touch" from the previous game 《NieR Replicant/Gestalt》 ,because it is a fantasy role-playing game which takes place in villages and grasslands—although it turns out to be a science-fiction game at very end. In comparison, 《NieR: Automata》 starts with battles of androids and machines, apparently showing a "mechanical world" that is awaiting players to explore. Even though they have completely different worldviews, we still want to meet fans' expectations. So, how does one get the balance between 《NieR: Automata》 and previous works became our main concerns'.³³⁶ This is consistent with UI designer Kojima's considerations, he explained, 'given the story setting of 《NieR: Automata》 , I replaced the analogue elements on screen by transforming them into digital signs, aiming to get a good balance between the atmosphere of fantasy and science fiction'.³³⁷

Besides, Yoko also requested that the production team use "warm and gentle beige" as the main colour tone for the visual identity, which implies his "contradictory aesthetics" again—a humanised world of non-human androids and machines. For the music, the MONACA team re-scored a few titles from the previous game and adapted them into different versions, such as : 〈**10. Grandma—Destruction**〉 , 〈**31. Song of the Ancients—Atonement**〉 and 〈**33. Emil—Despair**〉 ³³⁸. Through applying those preceding tunes to 《NieR: Automata》 in the form of new arrangements and alternative instrumentations, it can awaken fans' memories of the previous games and compels the 《Nier》 series appeal to new players.

³³⁶ Tadao and Sagako, *op. cit.*, p.11

³³⁷ Kijima, *op. cit.*, p.3

³³⁸ Further discussion can be consulted at Morojenie. (2020, March). 【心得】『自動人形音樂』前 你要知道的事. 巴哈姆特 *website*. Retrieved from <https://forum.gamer.com.tw/C.php?bsn=17353&snA=2947>

Aside from inheriting the spirits of 《Nier》, MONACA team determined a totally different composition strategy based on one significant difference of 《NieR: Automata》 –the **open world setting**. Okabe indicated that ‘compared to the previous game which is rather a close world as a miniature garden, 《NieR: Automata》 has much greater scale both on the story and its world setting. Therefore, we consciously intended to compose bigger pieces to present the magnificent view of the game’.³³⁹ With this clear purpose, the MONACA team added more “classical” arrangements and “orchestral” details into the compositions. Furthermore, the soundtracks for 《NieR: Automata》 comprised a wealth of notes enriching the compositions, instead of leaving too much space within the tune as with previous works. Since the music of 《NieR: Automata》 is implemented through vertical remixing, there is definitely one layer with very less sound as the ambience (Layer 1). But each track generally has many variations to adapt to different scenes.³⁴⁰

Nonetheless, 《NieR: Automata》 still has two features in common with previous 《Nier》 games –the **Japanese melancholy** in music and the **futuristic lyrics**. First of all, in comparison to other Western game music, which often presents a masculine, enthusiastic, and hearty impression to players, 《Nier》 series exhibits an opposite paradigm –a world of sorrow where is no so-called “hero”, “success”, nor a “glimpse of happiness or hope for the future”.³⁴¹ To successfully convey such a message and atmosphere, music plays a decisive role in the game, as Okabe explained: ‘Either it expresses directly or implicitly, I believe the emotions are both sincere. However, for 《Nier》 game series, I would like to make players feel more melancholy that uniquely belongs to Japanese [...] In other words, leaving some space for players to inhabit themselves inside the music, by putting the distance that encourages players to feel again, realising the aftertastes’.³⁴² This Japanese melancholy might be indirect, but this

³³⁹ Nagayoshi, *op. cit.*, p.6

³⁴⁰ Sugihara and Komori, *op. cit.*, p.11

³⁴¹ Napolitano, *op. cit.*, p.6

³⁴² Tadaku and Sagako, *op. cit.*, p.9 *The citation is translated from Japanese by this study, same hereafter.

modest attitude is probably the most eloquent way to convey Japanese elegance and the characters/director's thoughts about the world.

Secondly, the futuristic lyrics are sung by the vocalists. Emi Evans, the vocalist and lyricist of 《Nier》 games, mentioned that she was asked to create a “made-up futuristic language” for all the songs at her initial meeting with MONACA team. She said, ‘As *Nier* is set in the future, the MONACA team decided they wanted me to image how our languages of today would sound after thousands of years, [...] I was taking a specific language, respectfully manipulating it and then aging it a few thousand years’.³⁴³ After receiving the rough tracks 1 or 2 days in advance, Evans wrote the lyrics based on MONACA's instructions of what sort of languages they wished to use. ‘Altogether I wrote songs in 8 languages based on Gaelic, Portuguese, Spanish, Italian, French, English and Japanese’, said Evans. Only the 〈*Song of the Ancients*〉 for 《Nier》 was totally “made-up”, because she had been given no guidelines at all other than “imaginary language”. Even though some particular songs were given a list of Japanese keywords by director Yoko and Evans translated it to other languages, ‘the lyrics have no meaning—they are just a means by which to create a special atmosphere’, noted Evans.

However, Evans also pointed out the difficulty of this approach, because in most of the cases, even the composers were not sure how the songs would be used in the game. Evans described that ‘he [Mr.Okabe] could only say vague things like, “there is a peaceful open green space and then there's a big fight” or “there's a metallic mountain and then there's a big fight”. Everything ended in a big fight! So, it was difficult to know what sort of emotion I should capture... [...] **All I had to draw inspiration from were the melody and harmonies** and the actual sounds of the new style languages I was singing in’.³⁴⁴ Moreover, Evans indicated she was given generous freedom for interpreting melodies during recording, ‘I just had to use my imagination and decide for myself what kind of feelings to convey. [...] **the composers would listen to my singing**

³⁴³ Napolitano, *op. cit.*, p.2

³⁴⁴ Napolitano, *op. cit.*, p.3 *The emphasis is added by this study.

and often base their arrangements on the feeling of my voice.³⁴⁵ This approach of using “futuristic lyrics” was also applied to the song creations of 《NieR: Automata》 and it was proved to be a successful way to arouse players’ memories and touch their hearts. Wai-li, the editor-in-chief of *Game Walkthrough Monthly*³⁴⁶, compared 〈**31. Song of the Ancients–Atonement**〉 the adapted version to its original soundtrack, ‘I believe people who have played 《NieR》 heard 〈*Song of the Ancients*〉 in front of Depola as I did. I was deeply moved by the sad yet warm melody, even though I didn’t understand what she was singing at all.’³⁴⁷

The only one different usage of vocal between previous works and 《NieR: Automata》 is that the former had only one vocalist– Emi Evans, who had written all the lyrics and sung all of the versions while the latter’s music was written and sung by multiple people. The majority of songs still remain in the “tradition of Nier”, sung in “made-up futuristic languages”, but the ending song 〈*15. Weight of the World*〉 which is the only track written in existing human language has four versions (see Table 1). The English version 〈**15.**〉 for Ending A was written and sung by vocalist J’Nique Nicole, while the lyric of Japanese version 〈**15-1.**〉 for Ending B was written by director Yoko Taro and sung by singer Manina Kawano. Then, 〈**15-2.**〉 for Ending C and D was written and sung in invented French (futuristic language) by Emi Evans, and the last version 〈**15-3.**〉 for Ending E was sung by three vocalists and all the production team. For each version, we would examine in detail at next section. Here, it’s worth pointing out the reason for adding new vocalists in 《NieR: Automata》 and how their distinctive voices serve different purposes well for the game structure.

‘The vocals are an indispensable part for 《NieR》 music. The delicacy and transparent quality of Emi’s voice matched the worldview of 《NieR》 very well. Besides keeping her

³⁴⁵ Napolitano, *op. cit.*, p.4 *The emphasis is added by this study.

³⁴⁶ ‘*Game Walkthrough Monthly*’ (密技冰風暴 in Chinese) is published by PC home Publising Group-PC Gamer, who initiated the first portal site ‘Gamebase’ (遊戲基地 in Chinese) in 2000 in Taiwan. <https://www.gamebase.com.tw/>

³⁴⁷ 歪力. (2018, March). 重覆生與死的螺旋《尼爾：自動人形》一年後，我仍會一直感動下去. 4GAMERS. Retrieved from <https://www.4gamers.com.tw/news/detail/34463/after-1-year-the-memories-and-soundtracks-of-nier-automata-still-make-me-goosebumps>

unique voice, we also want to give 《NieR: Automata》 an alternative charm by mixing other styles, such as the powerful voice of J’Nique Nicole, who has been singing soul music for long time’, described Okabe.³⁴⁸ From Okabe’s perspective, putting Evans’ soothing voice and Nicole’s diva-like voice in parallel not merely adds fresh atmosphere to the new game, but the big contrast makes both vocals’ strength sound clearly and complement each other.³⁴⁹ ‘Emi’s voice well delineates the fragility of Nier’s world, while J’Nique’s vocal compels players to plunge into Nier’s world immediately’, was commented by Okabe.³⁵⁰

Given that both vocalists have their own strong characteristics, it became challenging to find an appropriate singer for the Japanese version. Yoko specifically requested the Japanese version should sound like **breaking down in tears of despair** to accompany Ending **B**, in order to clearly distinguish it from Ending A. Based on Okabe’s experiences, many leading singers might have been unwilling to do such “crying recording” which only for it to be produced as a CD and sold worldwide. Thus, the most suitable candidate for the Japanese version were not the ones who had outstanding singing skills, but the ones who could convey such desperate feelings through their vocal performance. Finally, the MONACA team found the Japanese singer Marina Kawano, who not only specialises in animation theme songs but also has had voice acting experiences from the same. Besides, aiming to make clear contrast to Ending A, the beginning of Japanese vocal is overlapped with some **whispering voices** to emphasise the singer’s image as a storyteller who is sharing her heartfelt wishes.³⁵¹ The most obvious performance is at the end of the song : the last lyric–‘僕は’ (bokuwa, meaning “I am”)–was sung/read **with breathing** as her last confession.

³⁴⁸ Tadaoku and Sagako, *op. cit.*, p.12

³⁴⁹ Nagayoshi, *op. cit.*, p.12

³⁵⁰ Tadaoku and Sagako, *op. cit.*, p.15

³⁵¹ Nagayoshi, *op. cit.*, p.14

I-3. What is 'Nier-like atmosphere' : rendering, reverberation, music suture

Having delineated the setting concept of 《NieR: Automata》 and how music helps differentiate it from previous games, I'd like to quickly review the features of immersion and the assessment standard which discussed in this study. The former consists of three characteristics of immersion : teleportation, absorption, and identification while the latter depends on how much a game system **preserves the fidelity to reality**—the more real-world sensory modalities it can deliver or display, the more “immersive” this game system is. Before we jump to answer the question what is a “Nier-like atmosphere”, it is fundamental to examine whether its world setting fulfils our criteria of immersion. That is, what have they done to make this world plausible, such that this world setting can successfully trigger players’ “suspension of disbelief”³⁵² ?

To begin with, 'realistic foley' and 'instant audio information' are two points needed to achieve the believability of world setting. **First**, Stockburger proposed the “spatial signature function” of sound samples, whose ‘acoustic information that is recorded with the original sample, conveys information and a feeling about the world as well’.³⁵³ According to Shindo, the environmental sounds were not made by software programming, but by real field recordings, for example, the river and the relaxation place that are in front of Resistance Camp were recorded in Kansai region (i.e., the west and south Japan), and the sounds for the area of Abandoned Factory were recorded at a neighbourhood in the Osaka shipyard.³⁵⁴ **Secondly**, Carson identified that an important task of a game designer is ‘to provide an answer to the question “where am I ?” that arises after starting up a game’.³⁵⁵ In addition, he suggested the best timing to answer this question is 'within 15 seconds, and thereafter, it is important to give some more

³⁵² ‘Suspension of disbelief’ refers to the intentional avoidance of critical thinking or logic in examining something surreal, such as a work of speculative fiction, in order to believe it for the sake of enjoyment. Retrieved from https://en.wikipedia.org/wiki/Suspension_of_disbelief

³⁵³ Huiberts, *op. cit.*, p.93

³⁵⁴ Shindo, *op. cit.*, p.14

³⁵⁵ As cited in Huiberts, *op. cit.*, p.94

information about the relationship to the place the user is in, which is essential for players to know what **their role** is in the setting'.³⁵⁶

Generally speaking, whenever a new place shows up for the first time in 《NieR: Automata》, there is around 10-second of silence to accompany the landscape close-up. The place title appears on the screen after the visual overlook. For instance, in gameplay 3, before the place title displays on screen, the sounds of fireworks and the gorgeous roller coaster have already provided clues that this place is the Amusement Park that the players were advised to visit in gameplay 2. In regard to Carson's suggestion of 15 seconds, I would like to take it as a reference, instead of a rule. An example of this suggestion can be found in gameplay 3. When 2B and 9S fell from the ceiling to the underground theatre (a new space), the curtain was unveiled, Beauvoir came out and started her singing. At this strange and unfamiliar scene, characters/players are not sure what is going to happen. After the combat music 〈**8. A Beautiful Song**〉 has been played around 10 seconds, the warning message of Boss fight finally shows up. The whole informing process takes around 20 seconds (18:15-18:35).

Next, verisimilitude is based on sound rendering instead of just sound reproduction.

Chion's study found that sound helps imprint rapid visual sensation into memory because the ear's agility in identifying moving figures is intuitional and omnidirectional.³⁵⁷ This is consistent with the investigation by Sonnenschein, which clarifies the capacities and limitations of human's sight and hearing sense.³⁵⁸ Chion indicated that sound is capable of aiding the apprehension of visual movements, which 'are "spotted" by rapid auditory punctuation, in the form of whistles, shouts, bangs, and tinkling that mark certain moments and leave a strong audiovisual memory'.³⁵⁹ By superimposing on the image, sound is not only capable of 'directing our attention to a particular visual trajectory', but also spotting these moving images into our memory.

³⁵⁶ *Loc. cit.* *The emphasis is added by this study.

³⁵⁷ Chion, *op. cit.*, p.122

³⁵⁸ Sonnenschein, *op. cit.*, p.151

³⁵⁹ Chion, *op. cit.*, p.11

Further, Chion proposed that 'in considering the realist and narrative function of diegetic sounds (voices, music, noise), we must distinguish between the notions of rendering and reproduction'.³⁶⁰ He described that we rarely pay distinct and focused attention to those sounds that take place in our everyday life. 'We only retain impressions of such sounds if they **carry material and emotional significance**. [...] The codes of theater, television, and cinema have created very strong conventions, determined by **a concern for the rendering more than for literal truth**. We are all thoroughly familiar with these conventions, and they easily override our own experience and substitute for it, becoming our reference for reality itself'.³⁶¹ Similarly, Sonnenschein also found that our hearing is very keen on association. 'Sound can serve as an emotional dictionary for a given film. [...] for example, flashbacks are created using transitions of aural memory to fling us from the character's subjective experiences to the objective action on screen'.³⁶² Both research findings point toward a vital fact about our hearing–sound reference is not merely established by the reproduction of reality but largely by the rendering of our aural memory. Only by doing so, can game players or film audience grab meanings and comprehend the situations within the fictional contexts. Chion's observations help clarify this finding further : 'The film spectator recognises sounds to be truthful, effective, and fitting not so much if–they reproduce what would be heard in the same situation in reality, but if they render (convey, express) the feelings associated with the situation'.³⁶³

Hence, our standard of assessing a game's level of immersion by–how much it preserves the fidelity to reality–should be amended as–**how much it renders the fidelity to our aural memory**. In truth, there are no "equivalent real-world sensory modalities" in the case of 《NieR: Automata》 –a story that takes place on the distant future, 5042A.D. We might be able to catch some aural memory from previous games of 《Nier》 series, because of what the MONACA team had already done–to rescore a couple of theme

³⁶⁰ Chion, *op. cit.*, p.109

³⁶¹ Chion, *op. cit.*, p.107-108 *The emphasis are added by this study.

³⁶² Sonnenschein, *op. cit.*, p.205

³⁶³ Chion, *op. cit.*, p.109

songs from previous works. But as elaborated earlier, 《NieR: Automata》 has a completely different worldview among others, so, new aural memory and auditory reality need to be generated for this new setting. Thus, we can now eventually ask the questions : What is a “Nier-like atmosphere”? What kind of emotional significance did they intend for the players to carry in their memories? And how did they render it through sounds?

‘From the musical aspect, it wouldn’t sound like “Nier” if there is no **reverberation**’, depicted Okabe. He mentioned his thoughts when he produced the music for 《Nier》 series : ‘Compared to J-POP music, which is usually full of notes, I composed pretty less notes for the soundtracks of 《Nier Replicant/Gestalt》 due to the “organic setting”. Through taking care of the reverberations within the music, I found this lingering sounds present the worldview of Nier very well. So, I decided to continue this approach at 《Nier: Automata》 , aiming for arousing the afterglow through these echoed sounds’.³⁶⁴ ‘We want to create the music not just sounds beautiful, but to have some room for players to feel themselves’, added composer Hoashi.³⁶⁵ In fact, this effective approach is supported by the research findings of Huiberts : An ambience mostly conveys a mood by incorporating elements that refer to settings in the real world, games or in movies. In addition to these world elements, reverberation can be used for mood induction.³⁶⁶

The second characteristics of the “Nier-like atmosphere” is the **stateless feeling**. This is established through two means : the futuristic lyrics and the multi-style composition. The former has been elaborated before, which results in the unique feeling one gets from Nier’s music—‘you can feel the sense of sorrow permeating through whole tune, which is dressed with some exotic flavour. However, you can hardly recognise where these flavours come from’, commented Hoashi.³⁶⁷ The latter refers to the mixture of Japanese

³⁶⁴ Tadaki and Sagako, *op. cit.*, p.13-14

³⁶⁵ Sugihara, *op. cit.*, p.5

³⁶⁶ Huiberts, *op. cit.*, p.93

³⁶⁷ Nagayoshi, *op. cit.*, p.9

melody with other styles, as Okabe described : It is certain that Nier's music has many Japanese-only musical thoughts within the soundtracks. In spite of that, we also put many "non-Japanese elements" into the compositions to fill up the space in melodies.³⁶⁸

The third feature of the "Nier-like atmosphere" is the **heartrending sorrow**. As mentioned before, 《NieR: Automata》 is a game of mourning—a world that is full of sadness and grief—based on director Yoko's setting. Thus, during gameplay the music should also represent these affections. However, Okabe and Hoashi both stressed the importance of controlling the degree of sorrow. 'We don't want to create the music that sounds like it's for funerals. Instead, we're trying to leave something that can make people feel the transience of the world, before there are too many emotions in the music', emphasised Hoashi.³⁶⁹ Okabe agreed with Hoashi's opinion, saying, 'Too much grieving emotions would make it sound melodramatic and cheap. So, our compositions for Nier series often go through the process of trial and error to restrain too much sadness'.³⁷⁰ Too much emotion within the music not only destroys the intended feelings, but also runs the risk of exhausting players. As Hoashi indicated, 'if the tune keeps too many strong emotions and plays for long time, players would only hear something rumbling in their ears. So, we're devoted to providing tracks that are emotional but won't make players feel tired easily'.³⁷¹

With respect to listener fatigue, the MONACA team has **3 guidelines of composition** for 《NieR: Automata》 : No too many weighty emotions, No intense dynamics within one tune, and Sparse composition for ambient layer. Hoashi stated their considerations with the issue of dynamics : 'If the dynamics are radical in one tune, it might be a burden for the listeners. [...] Because the obvious dynamics of the tune may confuse the players when they not even do anything in the game. I think it would reduce the immersive

³⁶⁸ Sugihara, *op. cit.*, p.5

³⁶⁹ Sugihara, *op. cit.*, p.12

³⁷⁰ *Loc. cit.*

³⁷¹ Nagayoshi, *op. cit.*, p.9

game experiences of players, when their focus are caught by music instead of events'.³⁷² In addition, ⟨2. *City Ruins*⟩ can be cited as the best example of "sparse composition for ambient layer", as Okabe mentioned, 'This tune was composed at a very early stage of the game. Given it is highly possible to play in the background for long time, we made the ambient layer of ⟨2. *City Ruins*⟩ sound looser than other tunes'.³⁷³ Besides these three principles on composition, the sonic effect at 《NieR: Automata》 also benefited from the in-game implementation via the **vertical remixing strategy**. 'Since different layers would be added based on players' actions, it enables the same tune to keep its dynamic full version and the nearly unnoticeable ambient version at the same time', indicated Hoashi.³⁷⁴ As mentioned previously, "seamless" is a chief principle of designing the game, both visually and aurally. Hoashi's observation not only points out the flexibility of vertical remixing, but also discloses the merits of applying this strategy—keeping players absorbed in the game world without feeling tired.

Arnold utilised the concept of **music suture**, proposed by Kamp, to analyse how the continuously looping audio can help players remain in immersion. He took the puzzle-platformer adventure game *INSIDE* (Playdead, 2016) as a case study : In keeping the rhythmic continuity intact throughout death and respawning, *INSIDE* not only avoids player frustration but also continually conditions the player to listen to the audio cues and respond accordingly. [...] [The rhythmic nature of gameplay creates] their ability to plan ahead by judging speed, distance, and positioning.³⁷⁵ Similarly, the studies by Denora (2000) and Kassabian (2013) also pointed toward how musical structures create a "haptic image" that dynamically engage listeners' bodies. Namely, the seamless musical loop not only 'generates overarching moods through combinations of timbre,

³⁷² Nagayoshi, *op. cit.*, p. 27

³⁷³ Nagayoshi, *op. cit.*, p.11

³⁷⁴ *loc.cit.*

³⁷⁵ Arnold, M. (2018). Inside the Loop: The Audio Functionality of Inside. *The Computer Games Journal*. <https://doi.org/10.1007/s40869-018-0071-x> *The emphasis is added by this study.

timing, and pitch'³⁷⁶, but also regulates players' bodies and minds as a medium prepared to react to unexpected challenges anytime. This is the **first case** when the soundtrack plays in its normal or Dynamic (full) version, where more than one layers are added respectively, based on the player's actions and game events.

The **second case** is when the soundtrack plays in its Quiet version which only has Layer 1, the ambient version of the composition. Eno indicated that 'ambient music must be able to accommodate many levels of listening attention without enforcing one in particular ; it must be as ignorable as it is interesting'.³⁷⁷ In the case of 《NieR: Automata》, the ambient layer is derived from the full version of the soundtrack. So, it can help players to remain in the same mood without the interruption of total silence while its rough composition won't demand players' attention. As Jørgensen highlighted, 'players notice that the immersion decreases, and that the fictional world seems to disappear and that the game is reduced to rules and game mechanics when sound is removed'.³⁷⁸ Hence, playing the sparse composition in the background, it enables players to stay in the virtual environments for a long time without fatigue. Besides, Rabin's study identified that the combination of ambience and environmental sounds can productively amplify players' positive feelings in game world : Ambience can sometimes create a mood even more than music. Even a small hum from a computer or distant waterfall or the wind through trees enhances the experience tenfold.³⁷⁹ Furthermore, the research by Szabo found that 'the absence of "personality" in ambient music allows each listener to put themselves in it and make the music their own personal soundtrack'.³⁸⁰ That is to say, the space within the ambient layer naturally absorbs players into the game environment without any pressure, while still gifting them the exploring freedom as they are there.

³⁷⁶ Szabo, Victor. (2019). 'What Music Isn't Ambient in the 21st Century?': A Design-Oriented Approach to Analyzing and Interpreting Ambient Music Recordings. In Ciro Scotto, Kenneth Smith, and John Brackett (Ed.), *Popular Music Analysis: Expanding Approaches* (2nd ed.) (pp.144–158). London: Routledge

³⁷⁷ Eno, Brian. (1978). [LP] "Ambient Music," liner notes to *Ambient 1: Music for Airports*, Editions E.G., AMB001

³⁷⁸ Jørgensen (2006), *op. cit.*, p.3

³⁷⁹ As cited in Huiberts, *op. cit.*, p.93

³⁸⁰ Szabo, *op. cit.*, p.154

II. Magical circle

The term “magical circle” coined by Johan Huizinga refers to the boundaries that the play activities take place within. Within the boundaries, the magical circle results in momentary worlds that reside inside our ordinary world.³⁸¹ These activities could happen in mediated environments and non-mediated environments, such as playing computer games, or playing tennis at a sports arena. Salen & Zimmerman adopted the concept of magic circle on games studies and noted that by circumscribing 'a special place in time and space created by a game', the magical circle encloses and separates it from the real world.³⁸² From the auditory aspect, Huiberts proposed a similar term, “**imaginary contract**” to delineate how sounds ideally work within the frame of game world. He described that by participating, the player agrees with the contract that is offered and consequently, has specific expectations concerning how things sound in that game. The presence of a sound asset that fundamentally goes against the contract is liable to disrupting immersion.³⁸³

With regard to immersion, I'd like to discuss two examples that exhibit opposite ways of applying the imaginary contract within the magical circle. **Firstly**, the usage of *<21. Wretched Weaponry : Medium/Dynamic>* and *<21-1. Wretched Weaponry : Quiet>* . The Dynamic version **<21.>** usually accompanies the battles that take place within the area of Abandoned Factory (gameplay 2, 7, 14), the tunnel fights of flying units (gameplay 9) and the Boss fight with Operator 210, who is crying for family (gameplay 19). Following the development of narrative, it is also used for depicting the brutal massacre happening at the Machine Village (gameplay 18). Then, the Quiet version **<21-1.>** appears when 9S meets Pascal, whose memory has been erased already, at the ruined Machine Village (gameplay 19). The musical choice is interesting. Since the original location music is no longer appropriate to represent the peaceful

³⁸¹ Leino, O. (2007). Feeling So Real – a phenomenological exploration into the realities of emotions in play. *Academia.edu*. Retrieved from https://www.academia.edu/484166/feeling_so_real_phenomenological_exploration_into_the_realities_of_emotions_in_play

³⁸² Leino, *op. cit.*, p.5

³⁸³ Huiberts, *op. cit.*, p.95 *The emphasis is added by this study.

Machine Village, instead of creating a new tune to illustrate the massacre and Pascal's amnesia, using the tune we've been familiar with can re-trigger players' curiosity within the context and without the expense of breaking immersion—especially during the last stage of the game.

The **second** example is the **Jukebox** design at Resistance Camp. Players can change the background music when they stay at the Resistance Camp. Each title has three or four versions to choose from : ambient version, the track with a little rhythm version, full percussion with full instrumentation version and the 8-bit version.³⁸⁴ After choosing one title, players can decide whether it plays with or without vocal. This innovative idea is great and supports the composer's hope of appreciating the soundtracks within the game. As Okabe once mentioned, 'Many fans told me they've bought the album and like Nier music very much, but they've never played the game before. Of course, I was still happy when people expressed their admiration to my music. However, as a game composer, I believe our duty is to enrich the game world further, by helping it establish the unique worldview through music. So, if possible, I would be much happier if fans can listen to it within the game'.³⁸⁵

However, this design is strictly limited to listen to the chosen music in front of the Jukebox, so it does not even apply to other spaces within the Resistance Camp. Whenever the player leaves a certain area, for instance, to visit weapon trader's stand which is still at the Camp, the chosen music is automatically replaced by the location music **⟨3. Peaceful Sleep⟩** , immediately breaking the "imaginary contract", resulting in confusion, annoyance and the collapse of immersion (see Appendix #8). I believe this issue can be improved by better implementation of the transition between music. Nonetheless, it is disputable whether the idea of letting the player intervene with the game setting is good or whether it will result in not achieving immersion. On the one hand, it can lift the players' spirits by giving them freedom to make their own choices. On the other hand, it withdraws players from their role in games, consciously being

³⁸⁴ Iwamoto, *op. cit.*, p.5-7

³⁸⁵ Nagayoshi, *op. cit.*, p. 28-29

detached from the fictional character. More discussions on this **paradoxical situation** are necessary if the production team wishes to keep this design for the next games.

7.3 Story

According to Huiberts, audio can be used to enhance the story—a fundamental part of game design—and players recognise this as positively influencing immersion.³⁸⁶ With respect to immersion, he examined the functionality of game audio through four perspectives : Overcoming barriers, Supporting primary and secondary emotions, Creating empathy with avatar, and Music as evaluation.

I. Overcoming the barriers of immersion by sounds

Huiberts identified that ‘to keep the player in the mood and to help concentration during game play, some of the barriers of immersion concerning the disruption of the flow of the game (for instance due to loading screens), can be overcome with sound’.³⁸⁷ An apparent example is the initiation page that starts the game. The players hear **〈1. Significance–Nothing〉** when they initiate the game for the first time and after Ending A while the **〈1-1. Significance〉** plays in the background if the players wish to play again after Ending B and Ending C. Both tunes eloquently show the Nier-like atmosphere—the loose soundscape that is full of reverberations and ambience, randomly accompanied by the soft chants with no clues as to where it comes from. By repeating these tunes to restart the game, it effectively helps players stay in the same mood and re-enter the atmosphere of game world more smoothly, even though they might be defeated many times. Another good example is the preview of route C and D which is shown after Ending B. As Huiberts discovered, ‘the audio storytelling during the loading screens keep tension, presents short flashbacks or create atmosphere, while keeping the player focused on the story, while preventing the real world or real world thoughts from interrupting’.³⁸⁸ To some degree, it is essential to have such a preview to arouse players’

³⁸⁶ Huiberts, *op. cit.*, p.96

³⁸⁷ Huiberts, *op. cit.*, p.101

³⁸⁸ *Loc. cit.*

interest. Apart from the changed viewpoint of protagonist and character campaigns, route B (gameplay 9-15) has almost the same structure and missions as route A. Thus, motivating players to explore other routes makes the loading/menu pages something to be reckoned with. The good thing is the ambience-like tune 〈**23. Broken Heart**〉 appears to accompany the preview right after the theme song 〈**15-1. Weight of the World**〉 has finished, which successfully bridges the gap between different routes and continues the game mood seamlessly.

In fact, compared to the music suturing approach at *INSIDE*, which runs the musical loops non-stop even when the players die and restart, **the flexibility in using silence and music** cues in 《NieR: Automata》 presents another valuable means to enhance immersion. For instance, when 9S realised the top secret of the blackbox at gameplay 19 (45:43-48:53), his psychological state deteriorated seriously. When the player follows 9S to hack into his memory space, we can clearly see his memories of 2B being invaded by logical virus which makes 9S totally collapse, he starts insanely stabbing the intruder and yelling to it, 'Stay the hell out...of my GOD DAMN MEMORIES! These belong to me and me alone!'³⁸⁹ Then, the screen shows that the intruder had already transformed into 2B's body while 9S continues stabbing, crying, then laughing. Undoubtedly, it is one of the most heartrending scenes in 《NieR: Automata》 which encompasses very intense and complex emotions. Then, the black loading screen comes in with complete silence for around 10 seconds. In my opinion, leaving no sound at the loading screen is a very wise choice, because intense emotions takes time to digest, and the player who has been playing the role of 9S needs room to take a breath. Any extra emotion brought by music or ambience would make this scene too weighty and melodramatic. It won't be possible to inhabit character's emotions in the player's heart until such emptiness is provided.

Further, it is worth mentioning that the setting of **loading screen** in 《NieR: Automata》 is actually not a stumbling block but the stepping stone to achieving immersion. Chen's study on Dynamic Difficulty Adjustment (DDA) showed that the gamer's flow experience can be accomplished and customised by offering a wide spectrum of activities and

³⁸⁹ The capital words are the original text.

difficulties to choose from. However, he also reminded us the frequent DDA choices might become potential interruptions which break the player's Flow. 'The only solution is to embed choices into the gameplay, let the player treat choices as part of the play and eventually ignore them. Thus, their choices will become intuitive and reflect their actual desires'.³⁹⁰ In the case of 《NieR: Automata》, whenever the players want to visit specific locations, check map, or receive mails from the Bunker, they go to the nearest **Access Point**—a machine that enables players to transport themselves in the game world of their free will (see Appendix #9). After the players make their decision, the Teleportation screen is the same as the Loading screen, which is programmed to appear at the transitions between different narrative plots. That is to say, by combining players' free choices into the core gameplay mechanics, players won't take the Loading screen as an interruption instead, it's just part of the game, and they still stay with 2B, 9S and A2 within the world of YoRHa project.

II. Supporting Primary & Secondary emotions – Ending A & B

Huiberts identified that game audio should help players recognise the current state of game by supporting two types of emotional response : The primary emotions concern the character, the secondary the player, who experiences the primary emotions but is able to feel differently from the reflected emotions of the game character.³⁹¹ To serve this purpose, game audio should be used for creating empathy with the avatar in the context of narrative, and emphasising the emotions felt by players as well. Especially for imaginary immersion, 'the aim is to couple the primary and secondary emotions (story has ended and the player is happy)', suggested Huiberts.³⁹² In 《NieR: Automata》, both route A and route B have the same sad ending at first place—2B has to strangle 9S with her own hands even though it causes her (and the players) so much pain. This scene is accompanied by the melancholy tune 〈**14. Vague Hope—Cold Rain**〉 whose lyrical melody in B minor expresses well the avatar's mental suffering and how helpless the

³⁹⁰ Chen, *op. cit.* p.13-14

³⁹¹ Huiberts, *op. cit.*, p.96

³⁹² Huiberts, *op. cit.*, p.97

Rebirth & Hope

NieR: Automata

Keigo Hoashi

Score analysis of <26. Rebirth & Hope> showing musical notation, annotations, and colored notes (yellow and red circles) highlighting specific elements. The score is in E-flat major, 6/8 time, with a tempo of quarter note = 64. The score is divided into two systems. The first system (measures 1-5) shows a piano introduction with a bass line of triplets and a treble line of quarter notes. Annotations include 'Strings comes in' with a red circle around a quarter note and a yellow circle around an arpeggio. A yellow arrow points from the first system to the second. The second system (measures 6-10) features a 'Chant comes in' with a red circle around a quarter note and a yellow circle around an arpeggio. A green arrow points from the first system to the second. The score ends with a mezzo-forte (mf) dynamic and a D-flat 6 chord.

Figure 40. Score analysis of <26. Rebirth & Hope>
(The coloured notes are added by this study)

situation is. Then, a machine's eyes start to shine green light, 2B thinks it might be new menace and prepares herself to fight. However, the incoming music gives a clue to the players—a brief tune <26. Rebirth & Hope> in E-flat major plays to accompany the visual scenes—not just one machine but hundreds of machines' eyes start shining as a green light within the vigorous and greenish City Ruins (see Appendix #10).

Before 2B realised it is not the enemy but the message from the still-alive 9S, the players had already known this is not the end through feeling the hopefulness of the tune. Here, the track <26.> functions as the "extradiegetic sounds", as proposed by Jørgensen, which can't be heard by the avatars but only by the players, in order to communicate the drama of the scene. Even though the full score is pretty short (around 40 seconds), it is still very effective to deliver the optimistic feeling through three ascending lines—the high voice of arpeggio (circled in yellow), the middle voice of quarter notes (circled in orange), the unceasing triplets from 8th notes to 16th notes performed by strings (noted with green arrow), as exhibited in Figure 40.

III. Creating empathy with the avatar – Ending C

The music order of Ending A and B are the same : **14. Vague Hope–Cold Rain** then **26. Rebirth & Hope** , followed by the monologue of 2B and 9S, and then the theme song **15. Weight of the World** plays as the final credit music. The only difference between these two endings is that the theme song for Ending A is English version while the Japanese version **15-1.** plays at Ending B, as we've discussed earlier. However, the Ending C has alternative music to represent the final cut scene :

24. Mourning plays when A2 runs to repair 9S' corrupted circuits. Then, after she has asked Pod 042 to take care of 9S, the theme song in Nouveau–FR version **15-2.** plays to accompany A2's monologue till the end of the credit. Compared to 2B and 9S, who appear as the heroine/hero of justice at the beginning, **A2** is presented as a Boss to fight against when she first shows up at gameplay 5 (53:32–56:55). Who could have predicted that an antagonist would become a protagonist, controlled by players during the gameplay? Player AKSBSU's experience summaries the mental process of having empathy with characters well : 'When I was introduced to A2, it seemed like she was an almost feral killing machine in [route] A and B with seemingly no redeeming qualities. How my opinion changed over time. She gave 2B a merciful, dignified death, preserved 2B's memories, spared 9S multiple times, was honest throughout, and we learn the horrific past she endured shaping who she became. I loved the way my perceptions of 2B and 9S changed over time as well and feel great sympathy for all of them. The three protagonists were fantastic characters'.³⁹³ Especially in Ending **C**, the decision of sacrificing herself and giving 9S' life truly turns A2 into a guardian angel, a promise-keeper, instead of a villainous traitor.

IV. Music as evaluation – Ending D & E

The research findings by Huiberts revealed that the superimposed music does not merely express what has happened and help the player to **reflect upon** it, but also

³⁹³ Shirrako. (2017, February 24). Nier Automata - Ending D & True Ending E (Secret TRUE Ending). [YouTube]. *Comment*. Retrieved from <https://www.youtube.com/watch?v=HxmOa5su7O4&list=PLCLLeSTzz6trYrOxS1anPprphUxjuOXBAA&index=21&t=3s>

provides the player with a **new direction** and new motivation to continue.³⁹⁴ Following the similar pattern of previous endings, Ending **D** is accompanied by soundtrack **<14.>** when 9S' fading consciousness is typed on the white screen, as shown in Figure 41.

"Will you come with us?" asks Adam.
The question was completely free of malice.
It seemed I no longer had a reason to hate machines.
Maybe I never had a reason in the first place.

What have I been fighting for?
Who have I been *living* for?

I don't know anymore.

Figure 41. 9S' will at Ending D

Then, two options –'I'll go with you' and 'I'll stay'–appear on the screen. After the player makes his/her choice, the theme song **<15-2.>** plays in the background to accompany the launch of the ark from the Tower. The brutal truth of the story is that the YoRHa's black box circuitry is made from cores of the machines. Namely, the android is the same creation as machines, made by God or Aliens. 9S' final words delineate the worldview that director Yoko wants us to question : What it means to be **alive without any clear purpose** and without even **the opportunity to die gracefully** (being mechanical beings, androids are reincarnated by means of data uploads after their deaths).³⁹⁵ Further, as the gamified violence in Yoko's game is designed largely to be fun, it throws more conflicting emotions and questions toward the players. 'That oxymoron-ish feeling of being between why does it feel so good to defeat an enemy? And why do I feel so much guilt in defeating an enemy?– that's an internal battle that we as humans always have', said by Yoko.³⁹⁶ Iwamoto's observation provided a valuable view to reflect on how we can enable such contradictory feelings through music : 'If we consider "time" as the "canvas size" of music, most of the scenes in 《NieR: Automata》 are united by one

³⁹⁴ Huiberts, *op. cit.*, p.99

³⁹⁵ Muncy, *op. cit.*, p.3 *The emphasis is added by this study.

³⁹⁶ Muncy, *op. cit.*, p.4 *The emphasis is added by this study.

soundtrack for long time—either it just present as the ambience in background, or plays as the full version to accompany the battles. The seamless feature makes the pieces have greater scale, similar to a big canvas that enables vocals—the human voices—to come in freely. Namely, the battles and the living world of characters are tightly connected in one soundtrack, which demonstrates Yoko’s thoughts about wars and the world well’, as Iwamoto reviewed.³⁹⁷ The seamless connection between wars and the ambience functions as a metaphor of our ordinary life : we are fighting for living, and the living is fighting—where the battle/challenge is always coming one after another. The viewpoint of Iwamoto shows that the music usage of 《NieR: Automata》 not only supports the emotional dynamics within the game, but also enriches director’s philosophy aurally and profoundly.

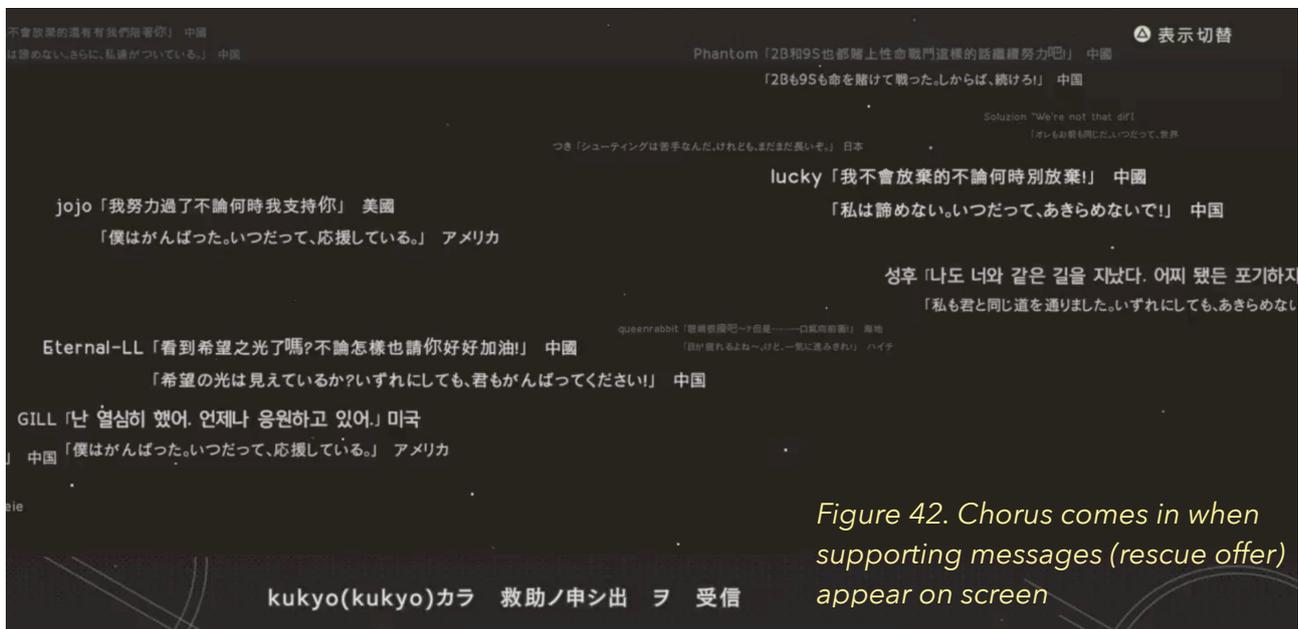
Besides reflecting upon the game through music, the other important function that music can contribute to immersion is supplying new directions. ‘Yoko requested us to produce the theme song in three different versions—English, Japanese and the made-up language—at the beginning. After we finish the composition, he hoped we could add an extra version which is sung by the production team due to the purpose of the game. So, Yoko, producer Mr. Saito and I also sang with the staff of PlatinumGames in 〈**15-3. Weight of the World—the End of YoRHa**〉 version which is used for Ending **E**’, described Okabe.³⁹⁸ ‘According to Yoko, the concept of this game is about “**self and others**”, so for me, this is a very meaningful soundtrack, in terms of the entire performance and experiences of the game’, commented Okabe.³⁹⁹

In truth, this intention is well received by players, as Kamiyama pointed out : ‘When I confronted the last shooting mission which was pretty hard to break through, I died many times and almost wanted to give up. However, since I chose “online mode” to play the game, it enabled me to ask for help. Once I clicked Yes, the chorus came in the music and many encouraging messages popped up instantly on the screen, which come

³⁹⁷ Iwamoto, *op. cit.*, p.19-22

³⁹⁸ Saito and Kaneko, *op. cit.*, p.4

³⁹⁹ Nagayoshi, *op. cit.*, p.24 * The emphasis is added by this study.



from other players who had already cleared up the game. [see Figure 42.] Compared to the well-devised structure of game design, this interlocking of music, chorus and my choice as a player, made me feel much touched and supported me to accomplish the tough shooting till the end'.⁴⁰⁰ Kamiyama's experience well depicts the sense of belongingness or the idea of "thinking about others" which was made to be realised through the gameplay. Yoko elaborated on his game design : 'Players can delete the saved data to "save heroine of the game" in previous NIER, which happens inside the game world. But for NieR:Automata, if the player chose to delete their saved date, it would **save "a player somewhere in the real world"**, which made it happen outside the game. I hope that people who play this game, take the time to think about someone they don't know in the far away countries. Even just for a little bit, that would make me extremely happy and secretly that is all I hope for'.⁴⁰¹

8. Result and Conclusion

In this study, the aim was to investigate two questions : (1) What is the role of soundtracks for immersion in computer games? (2) How do adaptive music and audio design build up the believability of game world? Regarding the **first question**, this paper has shown that the immersion in computer games can greatly benefit from the most prominent feature of soundtracks—inducing emotions to help players **identify** with

⁴⁰⁰ Kamiyama, ニーアオートマタにおける垂直型インタラクティブミュージックの一例, *op. cit.*, p.3

*Figure 42. is originated from Kamiyama's article and video.

⁴⁰¹ Yoko, T. (2018). A FUN TIME IN WHICH SOME NO-GOOD GAME DEVELOPERS MAY OR MAY NOT DISCUSS HOW WE MADE 'NIER:AUTOMATA. [Presentation] presented at Game Development Conference (GDC 2018-Vault). Moscone Center, CA.

the situation or a character in a game. From the **Sensory aspect**, the location music not only assures players of what to expect, but also enables to re-arouse players' attention by substituting other tunes with the purpose of narrative progression. Besides, the "vertical remixing" implementation results in the multi-layers and multi-functional music for one place. By cleverly putting the lead tune and the complementary tunes parallel to good allocation of frequencies, it can effectively avoid "sonic sludge" and "emotion sludge"—the too weighty emotions. Further, flexibly using "silence" to break down the seamless soundscape and having "8-bit version" music can heighten players' focus on narrative content and increase the appeal of character's uniqueness, such as 9S' hacking skill. From the **Challenge-based aspect**, utilising "melodic fulcrum" well and adding variations into the same tune can avoid the "loopable combat music" being frustrating too soon, therefore it can support players to conquer the time-consuming or tough sessions. Aside from its original function—supporting players succeed in challenges—the soundtracks in 《NieR: Automata》 also helps players prepare for the vital moments that proceed the narrative. In addition to this, replacing soundtracks flexibly to correspond to the "peaks and valleys" of gameplay also further establishes the thematic depth of its recursive structure.

Lastly, from the **Imaginary aspect**, the soundtracks of the "Nier-like atmosphere" delineates the worldview of the game well, which instantly hypnotises players to enter the game world smoothly, seamlessly connecting different in-game spaces and bridging alternative storylines/levels with its distinct features—reverberation, stateless feeling, and heartrending sorrow. One of the more significant findings to emerge from this study is that the mediacy of sounds and music turns the characters into the player's "extension of self" within the computer games, due to the indivisible kinaesonic congruity in games and our body schema. In other words, the auditory realm of the game enables us to embody our wills, choices, and actions through controlling characters—our "technological body" through which we sense the game world. The other noteworthy point is the varied layers and diverse versions of one track not only enrich the in-game emotional expressions, but also amplify the philosophical thinking behind the

mechanics design. The former refers to the multi-functional location/combat music, the latter indicates the usage of chorus in the theme song.

With regards to the **second question**, the SCI-model presents three findings. **Firstly**, the well-devised spatial audio design of 《NieR: Automata》, including realistic field recordings for foley remixing, greatly contribute to the *Sensory immersion*. From the plugin Simple 3D, Interactive reverb, to Doppler effect, the audio team successfully increases levels of fidelity and builds up a system of high sensorimotor contingency, which convinces players to “believe the realness” of the game world, especially based on the good mapping between their actions and the perceptible spatio-temporal effects of these actions. **Secondly**, the adaptive music used for motor challenges unites the whole atmosphere through the steady tempo while the relatively loose and calming compositions for cognitive challenges support players with the strategic thinking, collect useful data and reflect upon the profound dialogues with functional characters—which invites the players to bring their own thought and real-world experiences into the game world. **Thirdly**, it is proved that combining the good sound design and adaptive music can effectively indicate the characters’ state and arouse players’ empathy with them, such as the usage of Lo-fi noise during 2B’s final battle. Further, it is worth noting how they avoided the collapse of immersion through flexibly employing “music suture” (seamlessness) and “vertical remixing”.

By integrating the player’s free choices— **Teleportation**—into game mechanics, it not only overcomes the potential interruption of transitions, but also perfectly turns the Loading screen into a new storytelling interface by adequately using silence and making it the campaign stage. Meanwhile, utilising the multi-layers feature of vertical remixing can effectively remain players to be **absorbed** in the ready-to-fight state and the vibe of game world without fatigue. Moreover, the “absence of personality” feature of the ambient layer empowers players to inhabit themselves into the space within the compositions, as if it’s their personal soundtracks during the gameplay.

Finally, this study has found that the assessment of a game’s immersive level should take the “imaginary domain” into account. As Chion reminded us, ‘On screen, the audiovisual

channel has to do all the work of transmitting these two scenes : the filmmaker must "render" them by the sole means of image and sound. [...] the sound here must tell the story of a whole a rush of composite sensations and not just the auditory reality of the event'.⁴⁰² Therefore, our standard of evaluating the immersive level of a game system should turn to how much it "**render the fidelity to our aural memory**" to make players believe in, instead of "preserve the fidelity to reality". By rescoring a few tunes from 《Nier》, keeping the original voice and adding alternative vocals in new styles, 《NieR: Automata》 not merely inherited Nier spirits to recall fans' aural memory successfully, but it also created the new sonic reality and opened directions that well illustrate a "humanised world of non-human", luring human players into the fragile, beautiful spirals of life and death to reflect upon themselves and others.

Given that most references are from the aspect of production team, more investigation on player's side would help us better understand the reception of debatable settings and true user experiences, such as Jukebox, and the influence of paralleling 2 tunes of contradictory mood. Future research can be undertaken in two potential directions : One is 'how do the soundtracks help retain the "magical circle" from computer to reality', in regard to the cosplay culture of Nier. The other is 'how does Nier music aid the action flow to enhance gymnast's performance in real world?'⁴⁰³, and so on and so forth.

9. Bibliography

Monograph

Bresin, R. and Friberg, A. (2001). Expressive Musical Icons. In *Proceedings of the 2001 International Conference on Auditory Display*, Espoo, Finland, Hiipakka, J., Zakarov, N., & Takala, T. (Eds.). 141-143

Chion, M. (1994). *Audio-vision : sound on screen*. (Gorbman, C., Trans. & Ed.). New York : Columbia University Press

Csikszentmihalyi, M. (2014). *Flow and the Foundations of Positive Psychology—The Collected Works of Mihaly Csikszentmihalyi*. Claremont, CA : Springer

Freeman, D. (2004) *Creating Emotion in Games: the Craft and Art of Emotioneering*. Berkeley, CA: New Riders

⁴⁰² Chion, *op. cit.*, p.113 *The emphasis is added by this study.

⁴⁰³ MenRgJpn. (2016, August 14). 【NieR Replicant】 ニアのサントラで踊った男子新体操チームのクオリティがチートレベル【エミール&魔王】 . [Youtube] Retrieved from <https://www.youtube.com/watch?v=nXm6uahBq1E&list=LL>

- Ihde, D. (1979). *Technics and Praxis*. Dordrecht, Holland, D. Reidel Publishing Group
- Ihde, D. (1986). *Consequences of Phenomenology*. New York State University of New York Press, New York
- Jenkins, H., (2004). Game Design as Narrative Architecture. In Noah Wardrip-Fruin and Pat Harrigan (Ed.), *First Person: New Media as Story, Performance, and Game* (pp.118-130), Cambridge, Mass: MIT Press
- Kumar, P.S.J. (2018). Multiplayer Perceptron Neural Network Based Immersive VR System for Cognitive Computer Gaming. *Progress in Advanced Computing and Intelligent Engineering, Advances in intelligent Systems and Computing (AISC) Book Series of Springer*, 564, 91-102
- Lindley, C.A. (2002). The Gameplay Gestalt, Narrative, and Interactive Storytelling. In Frans Mäyrä (Ed.), *Proceedings of Computer Games and Digital Cultures Conference* (pp.203-215). Tampere: Tampere University Press
- Marks, A. (2009). *The Complete Guide to Game Audio*. (2nd ed.). Oxford : Focal Press (Elsevier Inc.)
- McMahan, Alison. (2003). Immersion, Engagement, and Presence—A Method for Analyzing 3-D Video Games . In Mark J.P. Wolf and Bernard Perron (Ed.) *The Video Game, Theory Reader* (pp.67-86). NY, US : Routledge, Taylor & Francis Group.
- Okabe, K. and MONACA. (2017). *ピアノ曲集—ニーア オートマタ(NieR:Automata) オフィシャル・スコア・ブック*. Tokyo, 有限会社ケイ・エム・ピー©Kmp.
- Rollings, A. & Morris, D. (2000). *Game Architecture and Design*. Coriolis Group, LLC. Arizona.
- Schuemie, M.J., van der Mast, C. A. P. G., Krijn, M., and Emmelkamp, P. M. G. (2002). Exploratory Design and Evaluation of a User Interface for Virtual Reality Exposure Therapy. In Westwood, J. D., Hoffman, H. M., Robb, R. A., Stredney, D. (Ed.), *Medicine Meets Virtual Reality*, IOS Press, 468-474
- Sheldon, L. (2004). *Character Development and Storytelling for Games*. Boston, MA: Thomson Course Technology.
- Sonnenschein, D. (2001). *SOUND DESIGN: The Expressive Power of Music, Voice, and Sound Effects in Cinema*. Studio City, CA : Michael Wiese Productions.
- Stone, A. R. (1996). *The War of Desire and Technology at the Close of the Mechanical Age*. Cambridge, MA : The MIT Press, 83-92
- Sweet, M. (2015). *Writing Interactive Music for Video Games : a composer's guide (2nd ed.)*. Indiana : Pearson Education. Inc.
- Szabo, Victor. (2019). 'What Music Isn't Ambient in the 21st Century?': A Design-Oriented Approach to Analyzing and Interpreting Ambient Music Recordings. In Ciro Scotto, Kenneth Smith, and John Brackett (Ed.), *Popular Music Analysis: Expanding Approaches* (2nd ed.) (pp.144-158). London: Routledge
- Toch, E. (1948, 1977). *The Shaping Forces in Music. (the Dover ed.)*. New York : Dover Publications, Inc.
- Wirth, W. & Hartmann, T. & Böcking, S. & Vorderer, P. & Klimmt, C. & Schramm, H. & Saari, T. & Laarni, J. & Ravaja, N. & Gouveia, F. & Biocca, F. & Sacau, A. & Jäncke, L. & Baumgartner, T. & Jäncke, P. (2007). A Process Model of the Formation of Spatial Presence Experiences. *Media Psychology*. 9. (pp.493-525). New Jersey : Lawrence Erlbaum Associates, Inc.

Professional Article

Arnold, M. (2018). Inside the Loop: The Audio Functionality of Inside. *The Computer Games Journal*.
<https://doi.org/10.1007/s40869-018-0071-x>

Cardinali, L., Brozzoli, C., and Farnè, A. (2009). Peripersonal Space and Body Schema: Two Labels for the Same Concept? *Brain Topography*, 21(3-4):252-60. DOI: 10.1007/s10548-009-0092-7

Chen, J. (2006). Flow in Games. (Master thesis). Available from Jenova Chen's website.

Cohen, A.J. (1999). The Functions of Music in Multimedia: A Cognitive Approach. *Paper presented at ICMPC: Proceedings of the Fifth International Conference on MUSIC PERCEPTION AND COGNITION*, August 26-30, 1998, Seoul, Korea.

Collins, K. (2011). Making Gamers Cry: Mirror Neurons and Embodied Interaction with Game Sound [Conference paper]. *AudioMostly 2011*, September 7-9, 2011, Coimbra, Portugal

Eerola, T., Friberg, A. & Bresin, R. (2013). Emotional expression in music: contribution, linearity, and additivity of primary musical cues. *Frontiers in Psychology*, volume 4 (July 2013). doi: 10.3389/fpsyg.2013.00487

Ekman, P., Friesen, W.V., and Levenson, R. W. 1990. Voluntary Facial Action Generates Emotion-Specific Autonomic Nervous System Activity. *Psychophysiology*, 27,4 (Jul 1990) 363-384. DOI=10.1111/j.1469-8986.1990.tb02330.x

Eladhari, M., Nieuwdorp, R., Fridenfalk, M. (2006). The Soundtrack of Your Mind : Mind Music—Adaptive Audio for Game Characters. *Paper presented at ACE'06: Proceedings of the 2006 ACM SIGCHI international conference on Advance in computer entertainment technology*, June 14-16, 2006, Hollywood, CA, USA.

Fisher, M. (2013). The Metaphysics of Crackle : Afrofuturism and Hauntology. *Dancecult : Journal of Electronic Dance Music Culture*, 5(2), 42-55

Hameed, A. & Perkis, A. (2018). Spatial Storytelling: Finding Interdisciplinary Immersion. *11th International Conference on Interactive Digital Storytelling, ICIDS 2018, Dublin, Ireland, December 5-8, 2018, Proceedings*

Hevner, K. (1935). The affective character of the major and minor modes in music. *Am. J. Psychology*. 47, 103-118. doi: 10.2307/1416710

Hevner, K. (1937). The affective value of pitch and tempo in music. *Am. J. Psychology*. 49, 621-630. doi: 10.2307/1416385

Huiberts, S. (2010). Captivating Sound: the Role of Audio for Immersion in Games (Doctoral dissertation). Available from ResearchGate.

IJsselsteijn, W. A., Ridder, H., Freeman, F., and Avons, S.E. (2000). Presence: concept, determinants, and measurement. *Proc. SPIE 3959, Human Vision and Electronic Imaging V*, (2 June)

Jørgensen, K. (2006). On the Functional Aspects of Computer Game Audio. *Proceedings of the Audio Mostly Conference 2006*, Oct 11-12

- Jørgensen, K. (2010). Game Characters as Narrative Devices. A Comparative Analysis of Dragon Age: Origins and Mass Effect 2. *Eludamos. Journal for Computer Game Culture*, 2010:4 (2), p. 315-331
- Juslin, P. N., and Lindström, E. (2010). Musical expression of emotions: modelling listeners' judgements of composed and performed features. *Music Analysis*. 29, 334-364. doi: 10.1111/j.1468-2249.2011.00323.x
- Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L. and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research*, 153, 4 (2003) 628-636. DOI= 10.1007/s00221-003-1603-5
- Knoblich, G., and Flach, R. (2003). Action identity: Evidence from self-recognition, prediction, and coordination. *Consciousness and Cognition: An International Journal*, 12(4), 620-632
- Kohata, S. (2018). An Interactive Sound Dystopia: Real-Time Audio Processing in 'NieR:Automata'. [Presentation] *Speech slide presented at Game Development Conference (GDC 2018-Vault)*. Moscone Center, CA.
- Kohler, E., Keysers, C., Alessandra, M., Umiltà, L.F., Gallese, V. and Rinaldi, G. (2002). Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science*, 297, 5582 (2002) 846-848. DOI=10.1126/science.1070311
- Morrison, I. and Ziemke, T. (2005). Empathy with Computer Game Characters: A Cognitive Neuroscience Perspective. In *AISB'05: Proceedings of the Joint Symposium on Virtual Social Agents*. AISB, UK, p.73-79
- Leino, O. (2007). Feeling So Real - a phenomenological exploration into the realities of emotions in play. Available from Academia.edu.
- Lombard, M. & Ditton, T. (1997). At the Heart of it All: The Concept of Presence. *Journal of Computer-Mediated Communication*, 3(2). <https://doi.org/10.1111/j.1083-6101.1997.tb00072.x>
- Reid, G. (2018). Chiptune : The Ludomusical Shaping of Identity. *The Computer Games Journal*, 7, 279-290
- Slater, M. (2003). A Note on Presence Terminology. *Presence Connect* 3 (3), 1-5
- Smethurst, T. (2015). Playing with Trauma in Video Games: Interreactivity, Empathy, Perpetration (Doctoral dissertation). Available from Academia.edu.
- Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining *Telepresence*. *Journal of Communication*, 42, No. 4 (Autumn), 73-93
- Yannakakis, G., Paiva, A., Karpouzis, K. and Hudlicka, E. (2011). Emotion in Games. *Proceedings of the 2011 Affective Computing and Intelligent Interaction Conference; Emotion in Games (EmoGames) Workshop*, Springer, 2011
- Yoko, T. (2018). A FUN TIME IN WHICH SOME NO-GOOD GAME DEVELOPERS MAY OR MAY NOT DISCUSS HOW WE MADE 'NIER:AUTOMATA'. [Presentation] *presented at Game Development Conference (GDC 2018-Vault)*. Moscone Center, CA.

Web source

Coen, O. (2018, June 5). Player Success: How to Help or Hinder It with Sound. [Web log post] Audiokinetic Blog. Retrieved from <https://blog.audiokinetic.com/player-success-how-to-help-or-hinder-it-with-sound/>

GDC Vault - An Interactive Sound Dystopia/ Real-Time Audio Processing in 'NieR/Automata'. Interactivemusic website. Retrieved from https://scrapbox.io/interactivemusic/GDC_Vault_-_An_Interactive_Sound_Dystopia:_Real-Time_Audio_Processing_in_'NieR:Automata'

Kamiyama, D. (2017, September 4). 【CEDEC2017】『NieR:Automata』の世界を彩る効果音はどのように実装されたのか？デザインコンセプトとその仕組みについて [Web log post] Inside website. Retrieved from <https://www.inside-games.jp/article/2017/09/04/109538.html>

Kamiyama, D. (2017, June 17). ニーアオートマタにおける垂直型インタラクティブミュージックの一例. [Web log post] Nine Gates Studio. Retrieved from <http://blog.nine-gates.com/1307/>

Kijima, H. (2018, June 19). 『NieR:Automata』のメカデザイン:機械 生命体編. [Web log post] Platinum official website. Retrieved from <https://www.platinumgames.co.jp/dev-nier-automata/article/351>

Kutay, S. (2006, May 1). Bigger Than Big: The Game Audio Explosion. [Web log post] gamedev.net. Retrieved from <https://www.gamedev.net/articles/audio/music-and-sound-fx/bigger-than-big-the-game-audio-explosion-r2317/>

Iwamoto, S. (2018, September 25). NieR: Automata—音楽によって「接続される」世界. [Web log post] note. Retrieved from <https://note.com/geekdrums/n/n9bbd292c74e1>

Matsudaira, H. (2017, April 3). An exclusive look at the creation of Nier:Automata's 2B. [Web log post] PlayStation.Blog. Retrieved from <https://blog.playstation.com/archive/2017/04/03/an-exclusive-look-at-the-creation-of-nier-automatas-2b/>

Muncy, J. (2018). Behind the Mask With Yoko Taro, Videogames' Most Interesting Designer. WIRED website. Retrieved from <https://www.wired.com/story/yoko-taro-nier-automata/>

Nagayoshi, H. (2017, April). 『ニーア』サウンドを生み出す、音の職人たち. 2083 web. Retrieved from <https://www.2083.jp/contents/201704nier/>

nakedjehuty. (2017, August 28). 「尼爾」不只是致鬱遊戲?《尼爾:自動人形》製作人與聲優群劇透演出心得. [Web log post] gamebase. Retrieved from <https://www.gamebase.com.tw/news/topic/98869270/>

Napolitano, J. (2010, May). Deep into NieR: Interview With Vocalist and Lyricalist Emi Evans. OSV (original sound version) website. Retrieved from <http://www.originalsoundversion.com/deep-into-nier-interview-with-vocalist-and-lyricist-emi-evans/comment-page-1/>

Saito, K. and Kaneko, M. (2017, March). 『NieR:Automata』 サントラ配信記念！音楽制作・MONACAスタッフインタビュー. Mora website. Retrieved from <https://mora.jp/topics/interview/nier-soundtrack/>

Shindo, M. (2017, Oct 3). 『NieR:Automata』 イベントシーンの効果音制作現場. [Web log post] Platinum official website. Retrieved from <https://www.platinumgames.co.jp/dev-nier-automata/article/240>

Shindo, M. and Kohata, S. (2018, December 4). Part 1: The spatial acoustics of NieR:Automata, and how we used Wwise to support various forms of gameplay. [Web log post] Audiokinetic Blog. Retrieved from <https://blog.audiokinetic.com/the-spatial-acoustics-of-nierautomata-and-how-we-used-wwise-to-support-various-forms-of-gameplay-part-1/>

Shindo, M. and Kohata, S. (2019, January 9). Part 2: The spatial acoustics of NieR:Automata, and how we used Wwise to support various forms of gameplay. [Web log post] Audiokinetic Blog. Retrieved from <https://blog.audiokinetic.com/the-spatial-acoustics-of-nierautomata-and-how-we-used-wwise-to-support-various-forms-of-gameplay-part-2/>

Shirako. (2017, February 22). NieR Automata - Gameplay Walkthrough Part 1 - Prologue (Full Game) PS4 PRO. [Youtube]. NieR Automata PS4 Gameplay Walkthrough Ps4 PRO 1080p 60fps Full Game Guide (playlist).Retried from <https://www.youtube.com/watch?v=4MU0yMgu3bQ&list=PLCLLeSTzz6trYrOxS1anPprphUxjuOXBAA&index=2&t=0s>

Shirako. (2017, February 24). NieR Automata - Ending D & True Ending E (Secret TRUE Ending). [YouTube]. Comment. Retrieved from <https://www.youtube.com/watch?v=HxmOa5su7O4&list=PLCLLeSTzz6trYrOxS1anPprphUxjuOXBAA&index=21&t=3s>

Square Enix official website (2017, March). NieR: Automata Original Soundtrack-Tracklist. Retrieved from Square Enix. Music <https://www.jp.square-enix.com/music/sem/page/nier/automata/>

Sugihara, T. and Komori, D. (2016, August). 『NieR:Automata (ニ－ア オ－トマタ)』の音楽はこうして作られる！ MONACAのスタジオに潜入、2バージョンのテーマ曲も公開(2/2). ファミ通 (famitsu) website. Retrieved from <https://www.famitsu.com/news/201608/17113095.html?page=2>

Taura, T. (2016, August 4). 『NieR:Automata』 プラチナゲームズ開発現場ご紹介#02. [Web log post] PlatinumGames official website. Retrieved from <https://www.platinumgames.co.jp/dev-nier-automata/article/27>

Tada, T. and Sagako. (2016, August). 『ニ－ア』を彩る音楽が生まれる地“MONACAスタジオ”ツアー。『オ－トマタ』のヴォ－カル曲数は.....全部!?. Dengeki Online website. Retrieved from <https://dengekionline.com/elem/000/001/341/1341859/>

Ueda, M. (2017, July 25). 『NieR:Automata』 BGM実装の開発裏話. [Web log post] PlatinumGames official website. Retrieved from <https://www.platinumgames.co.jp/dev-nier-automata/article/155>

Yin-Poole, W. (2009). Muzyka: Dragon Age and ME 2 will make Gamers Cry. Videogamer.com http://www.videogamer.com/news/muzyka_dragon_age_and_me_2_will_make_gamers_cry.html

歪力. (2017, April). 《尼爾：自動人形》製作人掲2B設計真相：眼罩+黒衣是本體. 4GAMERS. Retrieved from <https://www.4gamers.com.tw/news/detail/31949/nier-automata-producers-interview>

歪力. (2018, March). 重覆生與死的螺旋《尼爾：自動人形》一年後，我仍會一直感動下去. 4GAMERS. Retrieved from <https://www.4gamers.com.tw/news/detail/34463/after-1-year-the-memories-and-soundtracks-of-nier-automata-still-make-me-goosebumps>

*The last access date for all the web source is 2020, October 22.

Appendix : DVD content

All the content included in these DVDs, as the research materials of this study, are used ONLY for educational purpose. Any distribution for profit or commercial advantage would cause the infringement of third party copyright, therefore is strictly prohibited.

DVD 1. «**NieR: Automata Original Soundtrack**» © 2017 SQUARE ENIX CO., LTD. All Rights Reserved. <https://www.jp.square-enix.com/music/sem/page/nier/automata/>

Title (track number in this study)	Variations
1. Significance - Nothing	1-1. Significance
2. City Ruins - Rays of Light	2-1. City Ruins - Shade
3. Peaceful Sleep	
4. Memories of Dust	
5. Birth of a Wish	
6. The Color of Depression	
7. Amusement Park	
8. A Beautiful Song	
9. Voice of no Return - Guitar	9-1. Voice of no Return - Normal
10. Grandma - Destruction	
11. Faltering Prayer - Dawn Breeze	11-1. Faltering Prayer - Starry Sky
12. Emil's Shop	
13. Treasured Times	
14. Vague Hope - Cold Rain	14-1. Vague Hope - Spring Rain
15. Weight of the World [English Version]	15-1. Weight of the World [Kowaretasekainouta]
	15-2. Weight of the World [Nouveau - FR Version]
	15-3. Weight of the World [the End of YoRHa]
16. End of the Unknown	
17. Pascal	
18. Forest Kingdom	
19. Dark Colossus - Kaiju	
20. Copied City	

Title (track number in this study)	Variations
21. Wretched Weaponry - Medium / Dynamic	21-1. Wretched Weaponry - Quiet
22. Possessed by Disease	
23. Broken Heart	
24. Mourning	
25. Dependent Weakling	
26. Rebirth & Hope	
27. War & War	
28. Crumbling Lies - Front	
29. Widespread Illness	
30. Fortress of Lies	
31. Song of the Ancients - Atonement	
32. Blissful Death	
33. Emil - Despair	
34. Alien Manifestation	
35. The Tower	
36. Bipolar Nightmare	
37. The Sound of the End	

DVD 2. «**NieR: Automata**» **Gameplay video**

Due to the space limit, 10 selected examples are included in the DVD 2 to accompany this study. Researchers are recommended to follow the timecode mentioned in the thesis to get the thorough understanding within the gameplay context.

Except the #6. Lo-fi video, which is from Audiokinetic Youtube channel (© 2017 SQUARE ENIX CO., LTD. All Rights Reserved)¹, all the video clips are edited from professional gamer Shirrako's Youtube channel. <https://www.youtube.com/playlist?list=PLCLLeSTzz6trYrOxS1anPprphUxjuOXBAA>

¹ Audiokinetic. (2019, January 11). NieR:Automata- lo fi. [YouTube]. Retrieved from <https://www.youtube.com/watch?v=33v8tAAjUE0> *For the discourse, please see footnote 272.

Thesis Page	No.	Description	Duration
p.21-23 Machine Village	#1	<p>gameplay 3 : Machine Village firstly appears with <17. Pascal></p> <ul style="list-style-type: none"> ▪ Chorus comes in after Pascal has shown his kindness ▪ Satre : 'Do you believe that existence precedes essence?' ▪ 9S : 'Wait, they have genders?' 	3:27
p.37 Doppler effect	#2	<p>gameplay 9 : 9S comes out of the tunnel by the flight unit, flying toward the screen then leaving off the left side.</p>	0:13
P.55 Silence	#3	<p>gameplay 8 : The most tough battle right before Ending A. After 2 combat tunes, the silence highlights the war is over.</p>	1:00
p.61-62 Tone Filter	#4	<p>gameplay 19 : The long transition from normal track to 8-bit version shows the difficulty and uncertainty of Hacking.</p> <ul style="list-style-type: none"> ▪ When Devola & Popola are talking, <12.> in soft volume shows that Emil is somewhere not close to Resource Unit. 	1:15
p.85 TOP SECRET	#5	<p>gameplay 19 : Only the Layer 1 has been processed by Tone Filter. The clean mix helps players digest the subversive secret.</p>	1:15
p.100-101 Lo-Fi	#6	<p>gameplay 16 : The increasing noise level indicates the character's vitality, for example : 2B's final battle.</p>	2:43
p.113-115, p.121, 146 Ending D & E	#7	<p>gameplay 21 : Ending D ; 9S' will ; the transition to Ending E ; Pod042's WILL ; Instructive screen as a new storytelling interface</p>	14:16
p.140 Jukebox	#8	<p>gameplay 18 : Paradoxical design– Breaking immersion v.s Empowering players</p>	1:11
p.143 Access Point	#9	<p>gameplay 18 : Integrate players' free choice to game mechanics–Teleportation screen= Loading screen.</p> <ul style="list-style-type: none"> ▪ The dialogues show the personality of avatars : A2, machines. ▪ Music changes its reverb when AccessPoint menu opens. 	1:16
p.144 Rebirth	#10	<p>gameplay 8 : Music supports the Primary & Secondary emotions of the scene (Ending A & B).</p>	1:28