

Emotionales Sounddesign in Horror-Videospielen am Beispiel von

SILENT HILL 2 Remake

Bachelorarbeit im Studiengang Audiovisuelle Medien

vorgelegt von Lionie Berger

Matrikel-Nr.: 42092

Am 14. Mai 2025 an der Hochschule der Medien Stuttgart

zur Erlangung des akademischen Grades
eines Bachelor of Engineering

Erstprüfer: Prof. Oliver Curdt

Zweitprüfer: Prof. Dr. Oliver Wiesener

Ehrenwörtliche Erklärung

Hiermit versichere ich, Lionie Berger, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: „Emotionales Sounddesign in Horror-Videospielen am Beispiel von Silent Hill 2: Remake“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Ebenso sind alle Stellen, die mit Hilfe eines KI-basierten Schreibwerkzeugs erstellt oder überarbeitet wurden, kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§ 24 Abs. 2 Bachelor-SPO, § 23 Abs. 2 Master-SPO (Vollzeit)) einer unrichtigen oder Mai 2023 Prüfungsverwaltung 6 unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.¹

Lionie Berger



Stuttgart, 14.05.2025

Kurzfassung

Der Klang des Horrors verkörpert allein durch seine Zielsetzung, vorrangig negative Emotionen hervorzurufen, ein psychologisch einzigartiges Mediengebiet. Überträgt man diese Wirkung in die interaktive Welt der Videospiele, steigt die Komplexität des Gestaltungsprozesses ebenso wie die Anforderungen der Kunden an die auditive Ebene der Spielerfahrung. Diese Arbeit ergründet die Zusammenhänge zwischen Ton und Emotion im Kontext von Horror-Videospielen sowie deren Relevanz für das Spielerlebnis. Hierfür werden etablierte und potenzielle Methoden des Sounddesigns einschließlich ihrer technischen und emotionalen Besonderheiten analysiert. Aus den Ergebnissen geht hervor, dass Sounddesign im Horror-game-Genre einen strategischen Ansatz verfolgt, um intensive emotionale Reaktionen in einem kontrollierten Ausmaß zu induzieren. Die vertiefende Analyse des Videospiels *Silent Hill 2: Remake* belegt diese Erkenntnis und zeigt auf, mit welchen Mitteln das Sounddesign des Werkes ein emotional immersives, beängstigendes und symbolisch reichhaltiges Spielerlebnis erschafft.

Abstract

With its aim to primarily evoke negative emotions, the sound of horror embodies a psychologically unique area of media. When transferred to the interactive sphere of videogames, the complexity of the sounddesign-process increases, as do the customer's demands on the auditory level of the gaming experience. This thesis explores the correlations between sound and emotion in the context of horror videogames, along with their relevance for the user experience. For this purpose, established and potential auditory methods as well as their technical and emotional characteristics are analysed. The results suggest, that sound in the horror game genre follows a strategic approach to induce intense emotional reactions in a controlled manner. The analysis of the video game *Silent Hill 2: Remake* proves this insight and shows the means by which the sounddesign of the game creates an emotionally immersive, frightening and symbolically rich gaming experience.

Inhaltsverzeichnis

Abbildungsverzeichnis

Abkürzungsverzeichnis

Glossar

Anmerkungen

1. Einleitung.....	1
2. Emotionspsychologische Grundlagen im Kontext von Sound	3
2.1 Geräusche: Wahrnehmung und Reaktion.....	3
2.1.1 Elemente auditiver Wahrnehmung	3
2.1.2 Emotionale Wahrnehmung und Interpretation	4
2.1.3 Emotionale und physiologische Reaktionen	6
2.1.4 Evolutionsbedingte Wirkung von Geräuschen.....	9
2.2 Musik: Wahrnehmung und Reaktion	9
2.3 Wahrnehmung von Kontinuität und Stille	12
2.4 Emotionalität durch Videospiele.....	13
3. Grundlagen des Sounddesigns in Videospiele.....	14
3.1 Strukturen und Besonderheiten von Videospiele	14
3.2 Besonderheiten des Videospiel-Audios.....	15
3.3 Diegetisches und nicht-diegetisches Audio.....	16
3.4 Bestandteile des Sounddesigns in Videospiele	17
3.5 Ziele von Sounddesign in Videospiele	20
3.5.1 Infomation	20
3.5.2 Immersion	22
3.5.3 Unterhaltung.....	23
4. Audio in Horror-Videospiele	25
4.1 Timing und Lautstärke.....	26
4.2 3D-Audio	27
4.3 Vorwarnung	29
4.4 Transdiegetisches Audio	31
4.5 Simulation veränderter Bewusstseinszustände	33

4.6 Audio Uncanny Valley	37
4.7 Monster und der Sound der Einsamkeit.....	39
4.8 Musik	42
4.8.1 Musik als Warnsystem und Leitmotiv	43
4.8.2 Darstellung von Charakteremotionen und Umgebungen	44
4.8.3 Manipulation: Musik und Stille.....	44
5. Analyse von Silent Hill 2: Remake.....	47
5.1 Handlung	47
5.2 Sounddesign.....	49
5.2.1 Ambience (Video 1).....	49
5.2.2 Warnsystem (Video 2).....	52
5.2.3 Monster (Video 3)	54
5.2.4 Auditive Halluzinationen und Jumpscares (Video 4).....	57
5.2.5 Sonstige Soundeffekte (Video 5).....	60
5.2.6 Musik (Video 6)	63
5.2.7 Ende (Video 7).....	65
6. Fazit	68
Literaturverzeichnis.....	70
Videos.....	76
Videospiele.....	76

Abbildungsverzeichnis

Abbildung 1: Schematische Darstellung von Affektkonzepten.....	6
Abbildung 2: Kurve des Uncanny Valley	37
Abbildung 3: Gegenüberstellung der Audio-Signale bei drei Angriffen des Mandarin	56

Abkürzungsverzeichnis

HUD: Heads-Up Display

ILD: Interaural Level Difference/ Interaurale Pegeldifferenz

ITD: Interaural Time Difference/ Interaurale Laufzeitdifferenz

NPC: Non Player Charakter

SFX: Sound Effects

UI: User Interface

VR: Virtual Reality

Glossar

Angst: Diffuse, emotionale Reaktion, welche sich durch ein unbestimmtes Unbehagen ohne eine konkrete Bedrohung äußert

Emotion: Gesamtheit des Empfindungsprozesses, bestehend aus Gefühlen, körperlichen Reaktionen und Denkprozessen

Furcht: Emotionale Reaktion auf eine konkrete Bedrohung

Gefühl: Bestandteil der Emotion, welcher eine Empfindung zu einem bestimmten Zeitpunkt beschreibt

Natürlichkeit: Authentische, naturähnliche auditive Empfindung durch wenig oder keine künstliche Bearbeitung

Rauheit: Unregelmäßige, unruhige auditive Empfindung durch schnelle Frequenz- oder Amplitudenmodulationen bei der Überlagerung nahe beieinanderliegender Tonsignale

Schärfe: Stechende auditive Empfindung durch stark betonte hohe Frequenzen oder Sibilanten eines Signals

Anmerkungen

Die Verwendung englischer Fachbegriffe dient der besseren Lesbarkeit, Verständlichkeit sowie der Vermeidung wörtlicher Wiederholungen. Die Definition von Begriffen, welche im Literaturteil nicht vertieft erläutert werden, wie etwa „Furcht“ und „Angst“, erfolgt im Glossar. Zur Wortfindung einzelner Fachbegriffe sowie zur sprachlich optimierten Umformulierung vereinzelter Sätze wurde ChatGPT unterstützend verwendet.

1. Einleitung

“The oldest and strongest emotion of mankind is fear, and the oldest and strongest kind of fear is fear of the unknown” (H. P. Lovecraft, 1927, S. 12)

Das Gefühl der Angst, welches Lovecraft als die stärkste Emotion des Menschen beschreibt, verkörpert eine primordiale und faszinierende Komponente der menschlichen Psyche. Von Literatur über Filme bis hin zu Videospielen entwickelte sich in Form des Horror-Genres eine große Bandbreite an Medien und Unterkategorien, die es dem Rezipienten ermöglichen, einen Zustand der Angst und Spannung in einem von Gefahren isolierten Umfeld zu erleben. Insbesondere Horror-Videospiele bieten durch ihre interaktiven, nichtlinearen Eigenschaften eine Vertiefung der affektiven Wahrnehmung ihrer Rezipienten. In von Gefahr und Dunkelheit geprägten fiktiven Welten sich selbst überlassen, ist der Spieler besonders auf seine Ohren angewiesen, um zu überleben. Bezuglich der Aufgabe des Horrors, den Spieler durch eine gefährliche Umgebung zu leiten und simultan ein unheimliches Erlebnis hervorzurufen, stellt der Ton in diesen Videospielen ein unverzichtbares, enorm effektives Gestaltungsmittel dar.

Die Erweiterung audiotechnischer Möglichkeiten und die damit einhergehende Verbesserung der räumlichen, dramaturgischen und psychologischen Gestaltung, stellen einen essenziellen Faktor für die steigende Immersion und Beliebtheit angsteflößender Spiele dar. Auch wenn sich das Genre Stand 2024 weit hinter den Kategorien Shooter und Action im globalen Videospiel-Marktanteil verzeichnet (Kemp, 2024, Folie 146), ist ein deutlicher Anstieg der Nachfrage innerhalb der Kategorie Horror erkennbar. Allein auf der Videospiele-Plattform Steam kommt es von 2022 bis 2024 zu einer Verdopplung der veröffentlichten Horror-Spiele (SteamDB, o. D.).

Trotz des, insbesondere in den letzten Jahren, rapiden Wachstums, finden sich wenige wissenschaftliche Forschungen über die fachspezifische Nutzung psychologischen Sounddesigns und dessen Auswirkung auf die Spielerwahrnehmung sowie -reaktion. Ein vermutlicher Grund für diesbezügliche Mängel liegt in der grundsätzlichen Komplexität einer Übertragung der Psychoakustik auf Medien, welche durch die Konkretisierung auf eine spezifische Art Videospiele erst recht verstärkt wird. Da mit dem Fortschritt technologischer Möglichkeiten gleichzeitig der Nutzeranspruch an das Spielerlebnis wächst, kann die Erforschung und Anwendung akustischer Methoden auf der emotionalen Ebene als dringend notwendig betrachtet werden.

Das zentrale Ziel dieser Arbeit besteht daher in der Feststellung und Analyse der Methoden des emotionalen Sounddesigns in Horror-Spielen sowie der damit verbundenen Emotionen, welche beim Spieler induziert werden können. Etablierte Methoden werden dabei mithilfe von Anwendungsbeispielen verschiedener Videospiele erläutert. Außerdem wird der in Zukunft mögliche Einsatz bisher theoretischer psychologischer sowie technischer Konzepte beleuchtet.

Die dargelegten Ermittlungen widmen sich der Frage, durch welche Aspekte der grundlegende Zusammenhang zwischen Sound und Emotion gekennzeichnet ist, und untersuchen, inwiefern sich diese Korrelation auf dem Gebiet der Horror-Games äußert. Hierfür müssen zudem die Funktionen und Formen des Audios in Videospiele, sowie die Bestandteile, aus denen sich die akustische Nutzerwahrnehmung zusammensetzt, klargestellt werden.

Zu Beginn erfolgt die Erläuterung der emotionspsychologischen Ebene von Sound, welche über physikalische sowie subjektive Faktoren der Perzeption und Reaktion in Verbindung mit akustischen Reizen aufklären soll. Im Fokus des zweiten Kapitels stehen die Besonderheiten, Bestandteile und Ziele von Sounddesign in Videospiele. Kapitel drei kombiniert die vorangegangenen Themen und setzt sie in den Kontext des Horror-Genres. In Zuge dessen werden verschiedene auditive Herangehensweisen zur Erzeugung eines unheimlichen Spielerlebnisses beleuchtet. Um die ermittelten Komponenten des Horror-Sounddesigns anschaulich zu beleuchten, zieht diese Thesis schlussendlich zur beispielhaften Analyse das Werk *Silent Hill 2: Remake* (Konami, 2024) heran. Die Reihe der Silent Hill-Videospiele kann sich aufgrund ihrer beständigen und zahlreichen Fans seit 2001 als das Epitom der Kategorie des Survival-Horrors bezeichnen. In einem dystopischen Setting sorgt die visuelle sowie auditive Untermalung der psychologisch tiefgreifenden Handlung für eine Faszination, die viele Nutzer das Spiel als atmosphärisches, verstörendes Meisterwerk betiteln lässt (IMDb, o.D.). Mit der Neuauflage des Spiels wird eine allgemeine Optimierung der audiovisuellen Charakteristiken, welche dem Spiel seine Bekanntheit verleihen, beabsichtigt.

2. Emotionspsychologische Grundlagen im Kontext von Sound

Jeder Mensch entwickelt aufgrund unterschiedlicher emotionaler und physiologischer Voraussetzungen eine individuelle Interpretation akustischer Reize. Um auditive Wahrnehmungs-Konzepte auf das Genre der Horror-Games zu übertragen, bedarf es zunächst einer Erläuterung der wichtigsten akustischen Aspekte, welche das subjektive Hörerlebnis beeinflussen. Abschnitt 2.1 legt den Fokus neben Geräuschen auf Sound im Allgemeinen und bezieht sich daher stellenweise auf musikalische Stimuli.

2.1 Geräusche: Wahrnehmung und Reaktion

2.1.1 Elemente auditiver Wahrnehmung

Schall entsteht bei Anregung eines Mediums durch eine Schallquelle, wobei die Moleküle des Mediums in Schwingung gebracht werden. Die entstehenden Druck- und Dichteschwankungen resultieren in einer Ausbreitung des Schalls. In der Luft tritt dies in Form von Longitudinalwellen auf (Raffaseder, 2010, S. 86). Die über die Luft übertragenen Druckschwankungen gelangen zum Ohr und wirken auf das Trommelfell ein. Durch die Verarbeitung in Mittel- und Innenohr werden mechanische Schwingungen schließlich in elektrische Nervenimpulse umgewandelt, die an das Gehirn weitergeleitet werden. So kommt es schlussendlich zur Wahrnehmung von Geräuschen und Musik. Erwähnenswert für die Verarbeitung im Gehirn ist, dass die Signale unter anderem auch im limbischen System des Hirns verarbeitet werden, welches für Hormonhaushalt, Körperfunktionen und Emotionen zuständig ist (Raffaseder, 2010, S. 107–108). Das menschliche Gehör ist vor allem für Kommunikation, Lokalisierung und die Verarbeitung wichtiger akustischer Informationen, wie etwa Warnsignale, essenziell. Dabei besteht ein logarithmischer Zusammenhang zwischen der objektiven Intensität eines akustischen Stimulus und der subjektiven Empfindung dieses Reizes (Görne, 2017, S. 55). Die menschliche Lautstärkeempfindung ist frequenzabhängig, wobei der Grad der Frequenzabhängigkeit durch den Pegel beeinflusst wird. Generell werden tiefe Frequenzen vom Gehör bei gleichem Schalldruckpegel leiser als hohe Frequenzen wahrgenommen (Görne, 2017, S. 57). Ein gesundes Ohr nimmt Sound innerhalb einer Frequenzspanne von 20 bis 20 000 Hz wahr. Dabei reagiert es besonders empfänglich für akustische Stimuli im Bereich von 500 bis 4000 Hz (Dobie R. A. & Van Hemel S., 2004). Insbesondere bezüglich der Sprachverständlichkeit spielt der Frequenzbereich von 2 bis 5 kHz eine große Rolle, da sich in diesem Spektrum vorwiegend stimmlose Konsonanten befinden (Görne, 2017, S. 37).

Eine elementare Voraussetzung, um Informationen eines Sounds aufnehmen und in eine emotionale Reaktion wandeln zu können, ist eine angemessene Informationsdichte. Der sogenannte Shannon'sche Informationsgehalt eines Signals wird durch dessen zeitliche Änderungsrate bestimmt. Während sich ein deterministisches Signal wie der Sinuston als äußerst gering informationshaltig darstellt, ist bei einem stochastischen Signal wie weißem Rauschen genau das Gegenteil der Fall. Um im Hirn möglichst großes Interesse an der Aufnahme akustischer Reize zu wecken, stellen sich Signale wie Sprache und Musik als ideal heraus. Da sie eine mittlere Informationsdichte besitzen, sind sie für das menschliche Ohr gut deutbar und zugleich interessant, ungeachtet ihrer tatsächlichen Informationen. Um ein angenehmes Hörereignis sicherzustellen, darf das Gehirn also weder gelangweilt noch überfordert werden (Görne, 2017, S. 26–27). Ein Schallereignis enthält darüber hinaus Informationen bezüglich seiner Quelle, Anregung und Art des Raumes, in dem es stattfindet. Zum Beispiel kann das Zufallen einer Tür über deren Material und Dicke Aufschluss geben, ein Schrittgeräusch vermittelt dem Hörer Informationen über die Beschaffenheit der Schuhe, des Untergrunds sowie den körperlichen oder emotionalen Zustand der Person. Zur Interpretation dieser akustischen Hinweise greift der menschliche Verstand auf bekannte Ursachen für die Schallbildung zurück. Synthetische, komplexe Signale lassen sich dagegen nicht ohne Weiteres einem Entstehungsprozess zuordnen, was oft in einer Verunsicherung über die eigene Wahrnehmung resultiert (Raffaseder, 2010, S. 46–48).

Wenn ein visueller Reiz unerwünscht ist, fällt es dem Rezipienten leicht, seinen Sichtwinkel zu wechseln oder die Augen zu schließen und die Störung damit komplett zu beseitigen. Geht es ums Gehör, gestaltet sich dieser Vorgang komplizierter – es kann nicht einfach weggehört werden. Während das menschliche Sichtfeld auf einen bestimmten Bereich begrenzt ist, kann das Ohr akustische Informationen auf der Frontal-, Sagittal-, sowie Transversalebene außerhalb dieses Bereichs wahrnehmen. Dies hat zur Folge, dass die Anzahl der potenziell wahrnehmbaren Klänge und Geräusche über die nötige Informationsmenge hinausgeht, wodurch eine vollständige Aufnahme und Verarbeitung sämtlicher auditiver Reize unmöglich ist. Die Wahrnehmung von Sound im Alltag kann nicht verhindert werden und spaltet sich lediglich in unbewusste sowie bewusste Wahrnehmung auf (Raffaseder, 2010, S. 19).

2.1.2 Emotionale Wahrnehmung und Interpretation

Die Wahrnehmung von akustischen Ereignissen basiert zum einen auf der Art und den Kennwerten eines Geräusches als auch auf den Umständen, unter denen der Stimulus empfangen wird. Die räumliche Wahrnehmung akustischer Reize erfolgt durch Lokalisierung der Schallquelle sowohl auf der horizontalen als auch der vertikalen Ebene um den Empfänger herum sowie durch die Einschätzung der Distanz entlang der Tiefenachse. Schallereignisse mit höheren Frequenzen und

einer kurzen Attackzeit können präziser lokalisiert werden als tiefere Frequenzen mit einer langen Attackzeit. Die Lokalisation von Schallquellen im Raum funktioniert am besten in den ersten Millisekunden des Attack-Transienten, während die Lokalisation von langanhaltenden Geräuschen erschwert ist. Darüber hinaus beeinträchtigen Reflexionen zunehmend die Einordnung der Position einer Soundquelle. Die Wahrnehmung von Distanz ist insbesondere bei vertrauten Geräuschen erleichtert, da diese intuitiv mit einer bestimmten Lautstärke und einer daran gekoppelten Entfernung assoziiert werden. Bei unbekannten Schallquellen hingegen basiert die Distanzabschätzung primär auf spektralen Hinweisen, wie etwa der Abnahme hoher Frequenzanteile mit zunehmender Entfernung. Zudem gibt das Verhältnis von Direktschall zu Reflexionen Aufschluss über die Raumgröße und ermöglicht somit eine Einschätzung der Distanz (Farnell, 2010, 79–82).

Darüber hinaus kann die Länge eines Geräusches das subjektive Lautstärkeempfinden manipulieren. Bei kurzen Schallereignissen führt die begrenzte Stimulation dazu, dass der akustische Reiz einen geringeren Pegel zu haben scheint. Durch temporale Integration kann die wahrgenommene Lautstärke von Geräuschen mit einer Dauer unter 200 ms erhöht werden. Die Methode der temporalen Integration beschreibt dabei die Verlängerung eines Sounds um ein paar Millisekunden bei unveränderter Lautstärke. Dieser Effekt wird insbesondere im Sounddesign genutzt, um kurze Geräusche wie beispielsweise Schüsse lauter erscheinen zu lassen. Wie bereits dargelegt, erhält sich das Interesse an einem Geräusch durch dessen Informationsrate, woraus resultierend kontinuierliche Signale wenig Effizienz in der Aufrechterhaltung des Interesses aufweisen. Außerdem erscheinen sie mit dem Zeitverlauf deutlich leiser als zu Beginn der Reizaussetzung. Bereits eine kurze Unterbrechung des Signals kann jedoch einen Riss in der Pegeladaption verursachen und somit die ursprünglich empfundene Lautstärke sofort wiederherstellen (Farnell, 2010, S. 82–83).

Zum anderen hängt unsere individuelle akustische Wahrnehmung von persönlichen Erfahrungen sowie Erwartungen ab (Görne, 2017, S. 23). Der externe Sinnesreiz wird erst durch die Verarbeitung im Gehirn und die Übertragung dieser Informationen auf die dreidimensionale Welt zugänglich und erfahrbar gemacht. Görne (2017) beschreibt daher den einwirkenden Schall als „keine Klänge, sondern Kategorien für klangerzeugende Dinge oder Ereignisse“ (S. 23). Spezifische Musikstücke oder Geräusche besitzen das Potenzial, beim Hörer eine Aktivierung bereits gefühlter Emotionen als auch Fantasiebilder zu bewirken. Dabei bestimmen individuelle Erlebnisse und Assoziationen, welches Bild vor dem inneren Auge des Zuhörers entsteht. Durch Sound herbeigeführte, visuelle Szenarien und Emotionen können somit von Person zu Person stark variieren (Weiß, 2015, S. 24–25). Die im Unterbewusstsein des Empfängers bereits existierenden Gefühlsregungen werden beim Hören lediglich ausgelöst, wodurch sich erklären lässt, weshalb derselbe Sound in einer Person

Freude und in der anderen Ablehnung verursacht. Da komplexe auditive Informationen in ihren Bestandteilen nicht zeitgleich aufgenommen und verarbeitet werden können, besteht dieser emotionale Effekt auch nach mehrmaligem Hören und führt dazu, dass der Verstand jedes Mal neue Details in einem Werk oder Geräusch entdeckt (Görne, 2017, S. 46). Des Weiteren stellt sich die Wirkung akustischer Reize als kontextabhängig heraus. Während ein Schreibmaschinengeräusch früher für technischen Fortschritt stand, wird es heute mit Nostalgie verknüpft. Die Symbolik eines Sounds kann für Menschen unterschiedlicher Epochen, Kulturen und Gemütszustände variieren und führt somit oft zu konträren Empfindungen gegenüber einem Reiz (Raffaseder, 2010, S. 49–50).

2.1.3 Emotionale und physiologische Reaktionen

Welches Gefühl ein bestimmter Sound induziert, basiert auf einer Reihe komplexer Vorgänge und Einflussfaktoren. Eine übersichtliche Einteilung der Prozesse in der Emotionsverarbeitung durch auditive Reize gibt Eerolas hybrides Modell (2018, S. 5, siehe Abbildung 1). Dieses stellt verschiedene Affekt-Level und deren Mechanismen, die der Emotionserkennung und Induktion durch Musik zugrunde liegen, dar. Das Modell baut auf bereits bestehende Emotions-Modelle auf und kann über Musik hinaus auch auf andere auditive Signale übertragen werden.

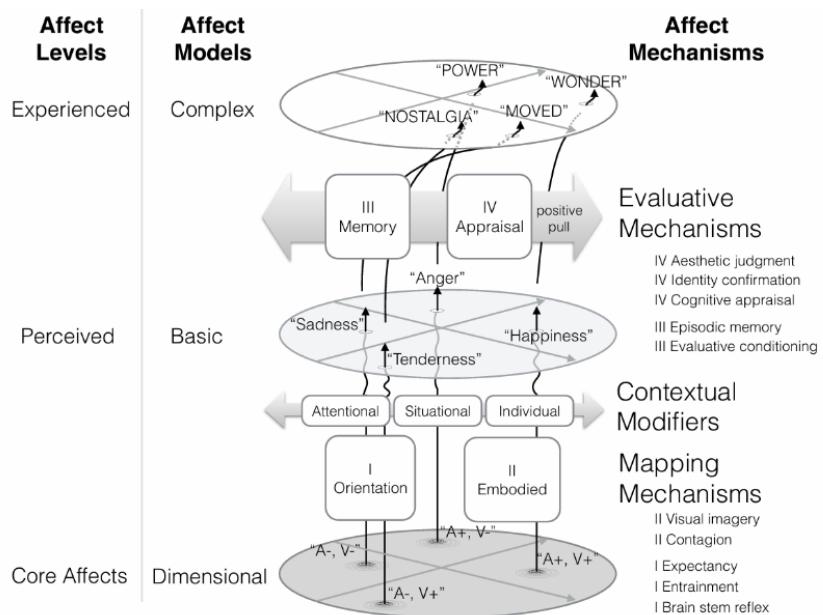


Abbildung 1: Schematische Darstellung von Affektkonzepten
Quelle: Eerola, 2018, S. 543

Das niedrigste Level, die Kernaffekte, beschreibt die Valenz-Erregungs-Dynamik, welche die erste Primitive bei der Einstufung psychologischer Ereignisse ist. Diese Ebene äußert sich, abhängig von der Ortung auf der Skala, durch bestimmte spontane Gefühle. Auf dieser Stufe sind dafür das autonome und endokrine System verantwortlich. In der zweiten Affektebene erfolgt durch

evalutive Mechanismen wie Konditionierung, Identität und Ästhetik-Gespür die Wahrnehmung und Erkennung von Emotionen. Schließlich führt dieser Prozess zu den tatsächlich erfahrenen Emotionen in der komplexen Affektebene (Reybrouck & Eerola, 2017, S. 7–8).

Insbesondere die Skala der ersten Ebene bildet in der Emotionsforschung eine Grundlage zur Kategorisierung induzierter Gefühle. Um emotionale Reaktionen von Rezipienten in ihren Bestandteilen beschreiben zu können, bedarf es einer Unterscheidung der Begriffe aus dieser Darstellung: Arousal und Valence. Dabei ist Arousal bzw. emotionale Erregung als die Intensität einer Gefühlsregung definiert. Valenz verweist auf den Grad, mit dem ein Gefühl positiv oder negativ ist. Aus dem Arousal-Valence-Modell ergibt sich somit beispielsweise für das Gefühl der Angst eine hohe Erregung bei niedriger Valenz und die Kombination niedrige Erregung mit hoher Valenz für Gefühle wie Zufriedenheit oder Ruhe (Neff, o. D.). Dieses emotionspsychologische Konzept kann ebenfalls auf emotionale Reaktionen auf akustische Stimuli übertragen werden. Außerdem besteht eine generelle Assoziation bestimmter akustischer Merkmale mit Positionierungen im Valenz-Erregungs-System: Klangeigenschaften wie Rhythmik, Komplexität, und Kontrast lassen sich dem Begriff der Erregung zuschreiben, Süße, Schärfe, Wärme, Kälte, Dunkelheit und Helligkeit befinden sich gegenüber voneinander auf dem Valenz-Spektrum (Görne, 2017, S. 201).

Wie Asutay et al. (2012, S. 23) veranschaulichen, werden Geräusche von Applaus, Lachen sowie einer Achterbahn mit einem höheren Valenzwert bewertet als der Sound eines Schreis, Alarms oder von Bienen. Dabei zeichnet sich die Unterscheidung zwischen positiven und negativen Reaktionen durch eine deutlich höhere Dynamik aus als die ermittelten Erregungswerte gleicher Geräusch-Samples. Daraus folgt, dass für jedes Geräusch ein gewisser Grad an Erregung gegeben ist, unabhängig von dessen emotionaler Bewertung (Asutay et al., 2012, S. 23, Tabelle 1).

In Bezug auf die Effektivität einer Emotionsinduktion durch Sound spielen die Parameter Lautstärke und Schärfe eine maßgebliche Rolle. Laut Västfjäll (2012), stehe der Grad an emotionaler Erregung vor allem in Verbindung mit der wahrgenommenen Schärfe. Valenz wird primär durch Lautstärke, allerdings auch mittels Rauheit und Natürlichkeit beeinflusst. Es ist ein Anstieg der Erregung bei steigender Schärfe und folglich eine verringerte Wachsamkeit für stumpfe Geräusche zu erkennen. Die Steigerung des Geräuschpegels bewirkt eine Absenkung in der Valenz, was für eine Relation zwischen erhöhtem Pegel und negativen Emotionen spricht (S. 607–608).

Ein weiterer Aspekt, welcher laut den Autoren einen signifikanten Einfluss auf die emotionale Reaktion auf Sound habe, sei die digitale Verfremdung von Geräuschen bzw. die Vorbereitung auf diese. Die Verfremdung der Hörbeispiele wurde mithilfe von moderaten spektralen und zeitlichen Verzerrungen realisiert. Die Autoren verfolgten in ihrem Vorgehen eine erschwerte Identifizierung des Geräusches bei der Erhaltung dessen physikalischer Eigenschaften.

Innerhalb der Klang-Kategorie angenehmer Sounds werden die originalen Stimuli im Vergleich zu ihren prozessierten Gegenstücken als wesentlich positiver wahrgenommen. Allerdings bewirken die als unangenehm kategorisierten Stimuli durch Verfremdung einen deutlichen Abfall an emotionaler Erregung.

Sogenannte „primed Stimuli“ beschreiben das Wahrnehmen des Samples zusammen mit einer auditiven oder informativen Vorbereitung auf das Hörereignis. Sie können beispielsweise in der Kategorie angenehmer Sounds zu einer Senkung der positiven Wahrnehmung im Vergleich zum unvorbereiteten Reizerlebnis führen. Bei unangenehmen Geräuschen fällt die Erregungsreaktion für vorbereitete Stimuli höher aus als für verfremdete (Asutay et al., 2012, S. 24, Abbildung 1). In Bezug auf negative Erregungsreaktionen kann darüber hinaus anhand auditiver Verfremdung ein Gefühl von Lästigkeit bzw. Ärger induziert werden, abhängig davon, wie angenehm oder unangenehm das originale Signal eingestuft ist. Durch die Nutzung von Verzerrung konnte die Studie in der Kategorie unangenehmer Sounds eine signifikante Minderung des Ärger-Empfindens im Vergleich zum Original feststellen (Asutay et al., 2012, S. 24–25, Abbildung 2). Ebenso sinkt die subjektiv empfundene Lautstärke für unangenehme Samples bedeutend aufgrund von Verfremdung und Vorbereitung. Für die Klang-Kategorien „neutral“ und „angenehm“ konnte die Manipulation der Stimuli keinen drastischen Effekt erzielen (Asutay et al., 2012, S. 25, Abbildung 3).

Der Mensch kann parallel zu den aufgeführten gefühlsmäßigen Reaktionen körperliche Regungen als Antwort auf einen akustischen Reiz hervorbringen. Hierbei besteht weitgehend eine Korrelation zwischen der emotionalen und der physiologisch erfassten Reaktion. Der persönlichen Einstellung gegenüber einem Geräusch oder Musikstück entsprechend, können sich physiologische Funktionen wie Atmung, Puls und Muskelspannung anpassen. Ein als unangenehm empfundener Song resultiert in einer beschleunigten und unregelmäßigen Atmung. Die Vorstellung eines Sounds, welcher Teil eines erinnerten Ereignisses ist, erzeugt Atemmuster, die denen beim aktiven Hören des Stimulus sehr ähnlich sind. Darüber hinaus kann es durch rhythmische Komponenten zu einer teils sehr genauen Anpassung der Pulsfrequenz an die Tempodynamik und Lautstärke kommen (Harrer & Harrer, 2008, S. 243–244). Kapitel 2.3 geht näher auf diese Erkenntnisse ein.

Vor allem durch Stress hervorgerufene mentale und physiologische Reaktionen können aufgrund bestimmter Geräusche verursacht werden. Das individuelle Ausmaß der Wirkung ist abhängig von Art, Ort, Zeitpunkt, akustischen Merkmalen des Sounds sowie dem Informationsgehalt für den Empfänger und dessen Einstellung zur Geräuschquelle bzw. deren Verursacher (Bayerisches Landesamt für Umwelt, o.D.). Jedoch kann aus evolutionärer Sicht von gewissen Überschneidungen in der Empfindung spezifischer Geräusche gesprochen werden. Diese Reaktionen, welche zum Teil

auf menschliche Urinstinkte zurückzuführen sind, sollen anhand des folgenden Abschnittes erläutert werden.

2.1.4 Evolutionsbedingte Wirkung von Geräuschen

Die evolutionsbedingte Verstärkung einer Reaktion auf bestimmte, vor allem gefährliche Geräusche, kann an Babys in einer Studie von Erlich et al. (2013) beobachtet werden. In einem Versuch wurden 56 Babys im Alter von neun Monaten 18 Geräusch-Samples mit den Kategorien angenehm, modern Angst-relevant und evolutionär Angst-relevant vorgespielt (Erlich et al., 2013, S. 896). Die Auswertung ergab eine gesteigerte Verlangsamung der Herzfrequenz für evolutionär Angst-relevante Geräusche gegenüber den modern Angst-relevanten und angenehmen Samples. Dabei kam es kategorienübergreifend zu maximaler Verlangsamung für die Geräusche wütender Mann, wütende Frau, Schlange, Sirenen, brechendes Glas und Bürogeräusche. Das Ausmaß der Lidschlag-Reaktion, als auch eine gesteigerte Orientierung des Babys in Richtung des Lautsprechers nahmen für die Kategorie der evolutionär Angst-relevanten Geräusche am stärksten zu (Erlich et al., 2013, S. 898–901).

Weitere, vor allem in der Tierwelt vorkommende, Signale wie der Schrei eines Schimpansen oder Murmeltiers, zeichnen sich häufig durch einen hohen Noise-Gehalt und ein Crescendo aus. Je höher die Dringlichkeit, mit einem Schrei vor Gefahr zu warnen, umso stärker bildet sich der Rauschanteil im Signal ab. Besonders das Genre des Horrors macht sich diese nichtlinearen Geräusche wie Schreie zunutze, um beim Zuhörer einen Zustand von Angst zu erzeugen (Blumstein, 2012).

2.2 Musik: Wahrnehmung und Reaktion

Musik unterscheidet sich von Geräuschen in den Kernpunkten Tonhöhe, Lautstärke und Klangfarbe (Raffaseder, 2010, S. 35). Roederer (1999) ist der Auffassung, anders als Geräusche, wirke Musik als ein Zusammenspiel von unterschiedlicher Struktur, Rhythmus und Tönen einer bestimmten Tonart und weise somit im Grunde keine evolutionär relevanten Informationen auf (S. 17). Jedoch entstanden im Zuge der Entwicklung des menschlichen neuronalen Netzwerkes komplexe Systeme zur Verarbeitung und Erkennung akustischer Signale, wobei vor allem der stimmliche, musikalische Anteil maßgeblich für den Austausch persönlicher und gemeinschaftlicher Gefühle war. Prozesse wie Analyse, Speicherung, Abruf und Vergleich klanglicher Informationen führten zu der Fähigkeit, komplexe akustische Muster wahrzunehmen und sie auf ihren symbolischen Wert zu reduzieren. Neben ihrer evolutionsbedingten Bedeutung wird die Wahrnehmung von Musik außerdem durch kulturell und persönlich antrainierte Vorgänge aktiviert. Der gezielte Einsatz von Musik als

Ausdrucksmittel ermöglicht aufgrund erlernter Assoziationen eine Interpretation von Umweltzuständen, die bedingt durch die Eigenschaften und Zusammenhänge der Klänge variiert. Trotz der Fähigkeit, evolutionär relevante Informationen aus musikalischen Signalen zu deuten, bezieht sich der Sinn von Musik hauptsächlich auf den emotionalen Zustand, den sie hervorruft. Selbst wenn die gegenwärtige Situation keine kognitiven Prozesse zur Bearbeitung eines musikalischen Reizes erfordert, finden diese, vermutlich mit dem Ziel der Unterhaltung, trotzdem statt (Roederer, 1999, S. 17–18; Hesse, 2003, S. 155). Wie bereits erwähnt kann dasselbe Musikstück bei verschiedenen Menschen komplett unterschiedliche Gefühle und Körperreaktionen auslösen. Abhängig sind diese Empfindungen zum einen vom Musikgenre und dessen Familiarität für den Hörer sowie den damit verbundenen Charakteristiken. Ausschlaggebend sind unter anderem Tempo, Lautstärke, Klangfarbe und der Text mit seiner Bedeutung. Zum anderen folgt aus der kulturellen, sozialen und persönlichen Individualität eines jeden Menschen, eine ebenso diverse Musikempfindung, welche nur schwer vorhersehbar ist. Abhängig ist diese von sozialen- und Bildungshintergründen als auch dem physischen und psychischen Zustand der jeweiligen Person. Musik kann sowohl auf Künstler- als auch auf Hörer-Seite eine Metapher der Gefühlswelt erschaffen und ein Gemeinschaftsgefühl bewirken bzw. verstärken. Die Intention und Auffassung dieser Aspekte stehen dabei nicht unbedingt in einem positiven Kontext (Fladt, 2012, S. 24–26; Hesse, 2003, S. 159). Über die persönliche Bedeutung hinaus bestehen kulturelle Assoziationsfelder, in denen neben kulturell bedingten Unterschieden in der Wahrnehmung, international gemeinsame Sound-Codes bestehen. Beispiele für diese erlernten Codes sind die Komponenten Disharmonie und Bass im Film, die als Gefahr interpretiert werden, Violinen, deren Klang für intensive Gefühle wie Liebe und Trauer steht, oder der Einsatz eines Belltrees, um auf Zauber und Geheimnis hinzuweisen (Weiss, 2015, S. 49).

Um das emotionale Empfinden in Bezug auf Musik genauer zu definieren, ist zunächst eine Unterscheidung der wahrgenommenen und gefühlten Emotion nötig. Die wahrgenommene Emotion bezieht sich auf die Perzeption emotionaler Expressionen, also das Erfassen des Gefühlausdrucks, den das Werk vermittelt. Der Hörer ist nicht selbst von den Emotionen betroffen und erlebt lediglich einen Prozess von Wahrnehmung und Reflexion. Gefühlte Emotionen sind das, was wir üblicherweise unter deren Begriff verstehen – die emotionale Reaktion auf den akustischen Reiz. Da es in den meisten Fällen nicht möglich ist, eine ausschließlich perzeptuelle Emotionshaltung zu Musik zu haben, sollten wahrgenommene und gefühlte Emotionen eher als Bestandteile einer Skala betrachtet werden (Gabrielsson, 2001, S. 128). Die induzierte Emotion kann dabei konträr zur wahrgenommenen Emotion sein. Ein fröhliches Stück, das der Hörer mit einem negativen Erlebnis verbindet, kann ein Gefühl von Traurigkeit auslösen. Traurig wahrgenommene Musik, die den Rezipienten tief berührt, führt wiederum zu positiv gefühlten Emotionen (Gabrielsson, 2001, S. 134–

135). Insbesondere auf negative Emotionen bezogen, übertrifft die wahrgenommene Stärke von Gefühlen wie Angst, Sorge, Trauer und Wut die der gefühlten. Dies lässt sich dadurch erklären, dass vor allem negative Verhaltenssymbole in der Musik nicht so einfach auf die Emotion des Hörers übertragen werden können, da dieser sich in einem vor Gefahren und Verlust isolierten Umfeld befindet (Zentner et al., 2008, S. 500–501).

Welche musikalischen Aspekte sind nun aber entscheidend für die Art der induzierten Emotion? Das Verhältnis verschiedener Klänge zueinander kann konsonant oder dissonant, vereinfacht gesagt angenehm oder unangenehm sein. In ihrer physikalischen Definition bedeutet Konsonanz bzw. Harmonie zum einen das Bestehen eines ganzzahligen Frequenzverhältnisses von Intervallen. Je komplexer diese Zahlenverhältnisse der Frequenzen zueinander sind, umso mehr Schwebungen entstehen. Daher nimmt der Mensch Verhältnisse wie eine Oktave (2:1) und Quinte (3:2) als besonders angenehm wahr. Zum anderen sind die Klangeigenschaft der Töne sowie die Übereinstimmung der Teiltöne, aus denen sie bestehen, relevant für den harmonischen Eindruck eines Signals. Es ist anzumerken, dass die Wahrnehmung von Konsonanz und Dissonanz musikalisch weder falsch noch richtig sein kann, sie ist lediglich kulturell bedingt gewohnt oder ungewohnt. Dieser Absatz bezieht sich auf das europäisch-westliche Tonsystem (Görne, 2017, S. 69–76). Der wechselseitige Einsatz von Konsonanz und Dissonanz ist vor allem für die dramaturgische Gestaltung im Sounddesign von Relevanz (Raffaseder, 2010, S. 117).

Neben Harmonie und Dissonanz tragen Attribute wie Tempo, Tonart, Tonhöhe, Lautstärke sowie deren allgemeine Dynamik zur wahrgenommenen Stimmung von Musik bei. In den Untersuchungen von Krumhansl (1997) zeichneten sich die Musikstücke, die in ihrer wahrgenommenen Emotion als traurig intendiert waren, durch Charakteristika wie etwa ein langsames Tempo, Moll-Harmonien sowie Konsistenz in Tonhöhe und Lautstärke aus. Letzteres fand sich ebenso in den fröhlichen Beispielen wieder, welche außerdem ein erhöhtes Tempo, tanzähnliche Rhythmen und Dur-Tonarten aufzeigten. Die mit Furcht assoziierten Beispiele beinhalteten dagegen ein hohes Tempo, Dissonanz, schnelle Lautstärkeänderungen sowie starke Tonhöhenkontraste. Die beabsichtigten Emotionen spiegelten sich, mit der Ausnahme stellenweiser Schwankungen, weitestgehend in der angegebenen Empfindung der Testpersonen wider. Es ist jedoch anzumerken, dass die durch Musik erzeugten Emotionserlebnisse nicht eindeutig mit denen nicht-musikalischer Emotionen vergleichbar sind. Die Autorin bestätigt zudem das Auftreten konsistenter physiologischer Veränderungen in Übereinstimmung mit der Art der musikalischen Emotionen. (S. 348–350).

Hesse (2003) identifiziert diesbezüglich die Parameter Intensität, Zeitablauf, Tonhöhenstruktur und Klangcharakter als entscheidende musikalische Komponenten, welche besonders körperliche

Reaktionen auf Musik bewirken. Die synchronisierte Regulation von Muskulatur und Atmung in Reaktion auf musikalische Stimuli führt zu einer aktivierenden Wirkung bei Merkmalen wie Dissonanz, häufigen Tempowechseln sowie hoher Lautstärke und Tonhöhendynamik. Im Gegensatz dazu werden gleichmäßige, ruhige musikalische Eigenschaften wie Konsonanz, konstantes Tempo, eine geringe Lautstärke und Tonhöhendynamik mit der Ruhefrequenz des Herzschlags assoziiert, was eine tendenziell beruhigende Wirkung zur Folge hat (S. 156–157).

2.3 Wahrnehmung von Kontinuität und Stille

Die menschliche Wahrnehmung kann sich ebenso auf physisch nicht vorhandene oder unvollständige akustische Reize beziehen. Das Gesetz der gestaltgerechten Fortsetzung besagt, dass Sounds als kontinuierlich empfunden werden, selbst wenn sie durch andere akustische Stimuli unterbrochen oder überdeckt werden. Wird beispielsweise ein gesprochener Text durch ein kurzes Geräusch übertönt oder in seinem Fluss zeitweise komplett unterbrochen, sodass einzelne Laute nicht hörbar sind, so wird dennoch der vollständige Text wahrgenommen. Der Mensch bildet eine räumliche Vorstellung, in der das störende Signal vor dem fortlaufenden steht, wodurch er sich der Vollständigkeit des überlagerten Signals bewusst ist. Somit entsteht durch die Rekonstruktion fehlender Informationen die Illusion von Kontinuität trotz Unterbrechung. Ein weiterer, verwandter Effekt ist die scheinbare Fortsetzung. Dieses Phänomen kann anhand einer stetig wiederholten musikalischen Sequenz beobachtet werden, welche nach einem unerwarteten Abbruch oder einer Pegelreduktion in der mentalen Wahrnehmung vervollständigt wird, selbst wenn ihr physischer Reiz nicht mehr hörbar ist. Im Sounddesign stellt die scheinbare Fortsetzung ein wirksames gestalterisches Mittel dar, um eine von den tatsächlich vorhandenen Stimuli abweichende Perzeption im Bewusstsein der Rezipienten zu induzieren (Görne, 2017, S. 98).

Goh et al. (2023) bestätigen in ihrer Studie die mögliche akustische Wahrnehmung von Stille. Diese kann demnach als auditives Event wahrgenommen werden, insbesondere wenn sie in Form von Signal-Lücken mit einer Dauer von ein bis fünf Sekunden auftritt, da das menschliche Gehirn aufgrund alltäglicher Pausen dieser Länge, wie etwa in Gesprächen und Musik, in diesem Bereich eine erhöhte Sensitivität aufweist. Die Versuchspersonen nahmen für unterschiedliche Experimente nicht primär die Dauer der Stille-Intervalle wahr, sondern konstruierten objektähnliche Repräsentationen für diese Lücken (S. 5). Bei der Unterbrechung einer Alltags-Ambience durch Stille-Phasen mit insgesamt genau derselben Dauer, wurde die einteilige Stille als zeitlich länger wahrgenommen als zwei aufeinanderfolgende Stille-Sequenzen mit einer kurzen Pause (Goh et al., 2023, S. 2–3). Diese Illusion entsteht durch die Segmentierung der Stille-Zeiträume und einer daraus

resultierenden Perzeption der Stille als Event. Die Wahrnehmung von Stille kann demnach als inhaltslose Eventrepräsentation mit non akustischer, temporaler Information beschrieben werden (Goh et al., 2023, S. 5).

2.4 Emotionalität durch Videospiele

Das Medium des Videospieles unterscheidet sich von anderen Medienarten im Wesentlichen durch seine Interaktivität und die damit einhergehende Aktion-Folge-Dynamik – der Spieler beobachtet das Geschehen nicht nur, er befindet sich, im übertragenen Sinne, direkt darin. Die Emotionen des Spielers helfen ihm, für seinen Erfolg und den Verlauf des Spiels, wichtige Entscheidungen zu treffen. Das Erreichen oder Scheitern bestimmter Ziele kann Freude, Stolz oder Wut auslösen (Isbister, 2023). Videospiele des Horror-Genres sind hauptsächlich mit Gefühlen wie Anspannung, Furcht, Angst und Trauer verbunden. Wieso also setzen sich die Spieler dieser Art Games absichtlich aus?

Nachgebildete Angstszenarien wie Horror-Filme oder -Games ermöglichen dem Nutzer eine verspielte Variante der emotionalen und physiologischen Stimulation, sozusagen eine angenehme Angsterfahrung. Für solche Situationen kann ein individueller Sweet Spot an Vergnügen entstehen, wenn das Erlebte nicht zu beängstigend, aber auch nicht zu langweilig ist (Clasen, 2023, S. 38–39). Emotional dunkle Spiele übertragen Situationen, die in der echten Welt unerwünscht wären, auf einen medialen Raum. So kann der Rezipient im Spiel mit der Angst wichtige Dinge über sich selbst und die Gefahren der Welt erfahren. Horror-Spiele konfrontieren den Menschen mit seinen eigenen Ängsten, Unsicherheiten und Schuldgefühlen, welche oft anhand eines Spielcharakters verdeutlicht werden (Epps, 2023).

3. Grundlagen des Sounddesigns in Videospielen

3.1 Strukturen und Besonderheiten von Videospielen

Nach Hunicke et al. (2004) lassen sich Videospiele in drei grundlegende Komponenten unterteilen, die im sogenannten MDA-Framework zusammengefasst werden: Mechanics, Dynamics und Aesthetics. Der Begriff Mechanics bezieht sich auf Aktionen, Verhalten und Kontroll-Mechanismen, beispielsweise das Regelwerk oder Assets wie Waffen und Munition. Dynamics beschreiben die Wechselwirkungen zwischen diesen Mechaniken und den Spielerinteraktionen sowie deren Auswirkungen auf den Spielverlauf, wodurch individuelle Spielerfahrungen entstehen. Schließlich umfassen Aesthetics den intendierten sowie hervorgerufenen Grad an Unterhaltung, welcher je nach Genre und Gamedesign verschiedene Formen wie etwa Empfindung, Fantasie oder Challenge annehmen kann (Hunicke et al., 2004, S. 3–4). Daraus folgt, dass je nach Zweck, den die Implementierung von Sounddesign im Spiel bedient, Audio in allen MDA-Komponenten integriert sein kann.

Videospiele zeichnen sich im Vergleich zu anderen Medien insbesondere durch ihre Nichtlinearität aus. Dies bedeutet, dass der Handlungsverlauf nicht chronologisch vorhersehbar ist und von den Entscheidungen des Spielers beeinflusst wird. Der Rezipient greift durch aktive Handlungen in den Verlauf des Spiels ein, wodurch zahlreiche neue Entscheidungsmöglichkeiten und Reaktionen des Umfelds entstehen, die vom Videospiel dargestellt und somit auch in ihrer Vertonung berücksichtigt werden müssen. Nichtlinearität kann dabei weitreichende Konsequenzen haben, indem sie narrative Veränderungen hervorruft und über den Erfolg sowie Misserfolg des Spielers als auch über die Beziehung zwischen Charakteren entscheidet. Der Grad der Ausprägung nichtlinearer Elemente variiert abhängig vom jeweiligen Videospiel (Leenders, 2012, S. 39–40). Ein weiteres zentrales Merkmal des Mediums der Videospiele ist seine interaktive Struktur. Die Spielumgebung setzt sich aus Objekten zusammen, die sowohl sichtbare als auch unsichtbare Attribute aufweisen. Diese umfassen beispielsweise Name, äußere Erscheinung, Verhalten, Inhalt sowie Alter. Auf Grundlage dieser Eigenschaften und anderer Mechanics interagieren sowohl Umgebungsobjekte als auch Spielerobjekte miteinander. Dabei repräsentiert das Spielerobjekt sowohl die visuelle Perspektive als auch die Handlungen des Spielers (Farnell, 2010, S. 315–316).

3.2 Besonderheiten des Videospiel-Audios

Das Sounddesign in linearen Medien wie Filmen bildet die visuell dargestellten Reize zeitgleich nach und bleibt dabei in jedem Durchlauf für alle Rezipienten sowohl chronologisch als auch gestalterisch identisch. Die Erschaffung eines Soundsystems in Videospielen gestaltet sich in dieser Hinsicht deutlich komplexer. Da auditive Events durch die Spieleraktionen unvorhersehbar, wiederholt und kontinuierlich ausgelöst werden können, wird der Audio-Output eines Spiels sozusagen in Echtzeit durch die Aktionen des Spielers gemischt. Um spontane und vielfältige Reaktionen zu erzielen, ist daher die Generierung mehrerer Sample-Versionen sowie variierender Intervalle der Sound-Elemente innerhalb eines Loops nötig. Anders als bei Kinozuschauern ist Spielern die Möglichkeit gegeben, ihr Sichtfeld nach eigenem Ermessen zu ändern und sich frei in der virtuellen Welt zu bewegen. Im Vergleich zum Film folgt daraus eine weitaus höhere Relevanz für die Wahrnehmung von Sound aus dem hinteren Feld des Hörers. Durch Drehung der Charakterkamera sowie seine Bewegungsfreiheit ist der Spieler befähigt und darin bekräftigt, die Geräuschquelle ausfindig zu machen und gegebenenfalls mit ihr zu interagieren (Suter et al., 2021, S. 92–93).

Auf der räumlichen Wahrnehmungsebene wird Game-Audio in folgende dimensionale Arten segmentiert: nicht lokalisierte 2D-Sounds, deren Hörereignis unabhängig von der Spielerposition ist, sowie lokalisierte 3D-Sounds, welche aus einer bestimmten Richtung emittieren und deren akustische Wahrnehmung an die Spielerposition gekoppelt ist (Sinclair, 2020, S. 51). Ein Übergang zwischen diesen Dimensionsebenen ist in Form von Distanz-Crossfades möglich (siehe Abschnitt 3.5.1). Die Implementierung von 3D-Audio in Videospielen kann grundlegend objektbasiert oder kanalbasiert realisiert werden. Objektbasiertes Audio bezeichnet Soundquellen, die mithilfe eines Decoders anhand von Metadaten im dreidimensionalen Raum als Objekt platziert werden und somit auf unterschiedliche Wiedergabesysteme übertragen werden können. Im Gegensatz zu kanalbasierten Systemen wie Stereo oder 5.1 zeichnen sich objektbasierte Technologien wie Dolby Atmos und DTS:X daher durch eine erweiterte Ton-Dimension in der Vertikalen sowie eine höhere Anpassungsfähigkeit an unterschiedliche Wiedergabesysteme aus. Da Audio in Videospielen vorwiegend über Kopfhörer rezipiert wird, empfiehlt sich vor allem die Implementierung von 3D-Audio mittels binauraler Aufnahmen oder Renderings. Durch die Berücksichtigung interauraler Zeit- und Pegeldifferenzen sowie der Head-Related-Transfer-Function im Rahmen des binauralen Audios entsteht ein immersives Hörerlebnis sowie eine realitätsnahe Lokalisierung im dreidimensionalen Raum. Diese kann jedoch aufgrund anatomischer Unterschiede an die individuelle Empfindung nur angenähert werden (Sinclair, 2020, S. 58–60).

Neben der Varietät in mehrfach auslösbarer und geloopten Soundeffekten besteht mit prozeduralem Audio eine wichtige Diversitäts-Komponente in Games. Prozedurales Audio

beschreibt die synthetische Erzeugung von Audio während der Laufzeit des Spiels ohne eine vorher gerenderte Soundquelle. Physikalische Merkmale und das Verhalten einzelner Instanzen werden von der prozeduralen Audio-Engine verarbeitet, um darauf basierend, sowie mithilfe auditiver Variablen, Soundeffekte oder Musik in Echtzeit zu generieren. Auf diese Weise entsteht eine akustische Umgebung, die in Wechselwirkung mit den Mechaniken des Spiels steht und sich folglich in spezifischen Verhaltensmustern äußert. So kann beispielsweise das Vorbeifahren an anderen Fahrzeugen durch einen situationsabhängigen, dynamisch generierten Rausch-Sound realitätsnah abgebildet werden. Ebenso lässt sich die Vokalisation prozedural generierter Kreaturen an deren visuelle Charakteristika anpassen. Prozedurales Audio bietet keineswegs einen vollständigen Ersatz für sämtliche gestalterische Prozesse, jedoch eine solide Grundlage für die Implementierung von adaptivem, dynamischem Audio in Videospielen (Farnell, 2010, S. 321; Mraz, 2021).

3.3 Diegetisches und nicht-diegetisches Audio

Zunächst ist die Einteilung aller Audio-Objekte innerhalb eines Videospiels in zwei Kategorien notwendig. Ein auditives Ereignis wird als diegetisch definiert, wenn seine Schallquelle innerhalb der Spielwelt verortet ist. Dabei können sich diegetische Elemente im sichtbaren Bereich als auch offscreen befinden. Nicht-diegetische, auch sogenannte extradiegetische Sounds hingegen werden nicht durch eine Quelle in der virtuellen Welt emittiert, sondern sind vielmehr als metaphorische Darstellung des Unterbewusstseins des Spielers und dessen Charakters zu betrachten. Dabei handelt es sich häufig um Klänge, die von der Figur selbst nicht wahrgenommen werden, jedoch die Stimmung eines Charakters oder einer Situation vermitteln, um bestimmte Gefühlszustände bei den Spielenden zu induzieren. In Gegenüberstellung zum Medium Film verfügen die Rezipienten von Videospielen über die Fähigkeit, mit direkten Handlungen auf interne sowie externe Sound-Ereignisse zu reagieren. Insbesondere ist das Wirkungsausmaß der extradiegetischen Ebene hervorzuheben, da diese in Form auditiver Informationen Einfluss auf die Spielerhandlung und das diegetische Geschehen ausübt, ohne unmittelbar in der fiktiven Welt platziert oder den darin befindlichen Charakteren bekannt zu sein. Kommt es zu einem Übergang zwischen diegetischem und nicht-diegetischem Audio, bezeichnet man dieses als transdiegetisch (Bridgett, 2021, S. 77–82; Jørgensen, 2007, S. 106–107).

3.4 Bestandteile des Sounddesigns in Videospielen

Um die Rolle von Sound im Kontext von Videospielen adäquat analysieren zu können, ist eine Erläuterung der grundlegenden Audiodesign-Bausteine innerhalb des Mediums erforderlich. Bridgett (2021) kategorisiert die zentralen Elemente des Sounddesigns in vier Gruppen, welche die wichtigsten Bestandteile und deren Relation zueinander darstellen. Dabei werden die Gruppen Sound, Musik und Stimme durch gestalterische Prozesse in der Gruppe Mix über die Dauer eines Mediums verbunden und aufeinander abgestimmt (S. 10–11).

Die Kategorie **Sound** umfasst eine Vielzahl akustischer Elemente, die zur auditiven Gestaltung eines interaktiven Mediums beitragen. Dazu zählen unter anderem Soundeffekte (SFX), Foleys, Ambience sowie speziell im Bereich der Videospiele: Waffen-, User Interface- sowie Heads-Up Display-Sounds. Der Begriff der Soundeffekte umfasst ein breites Spektrum an Audio-Events und dient als Sammelbezeichnung für sämtliche im Spiel vorkommende Sounds, die weder Musik noch Sprache oder einer spezifischen Sound-Art zugeordnet werden können. Unter der allgemeinen Definition der SFX versteht man einzelne oder mehrschichtige Klänge und Geräusche, die für spezifische Spielkomponenten erstellt und angepasst werden. Unter den Begriff der SFX fallen beispielsweise Objekt-Geräusche wie Waffensounds als auch Aktions-Geräusche wie etwa Kollisionen. Da die Kategorisierung der Soundeffekte mit den Entwicklern eines Spiels variiert, erfolgt nun eine Erläuterung der zentralen SFX-Kategorien und ihrer wichtigsten Merkmale (Bridgett, 2021, S. 12–13).

Als grundlegender Baustein fördert die ambiente Audio-Ebene insbesondere die atmosphärische Stimmungsbildung eines Spiels. Eine Ambience ist meist in mehreren Schichten aufgebaut und beinhaltet unterschiedliche Komponenten wie z. B. gelooptes Hintergrundrauschen und Klänge sowie variierend auftretende Umgebungsgeräusche (Sinclair, 2020, S. 174). Sie kann in statischer Form als zonenbasierte Klangkulisse realisiert werden, beispielsweise als 2D- oder Quad-Ambience (vierkanalig). Darüber hinaus kann sie in dynamischer Form auftreten, indem sie sich in Abhängigkeit von der Spieler- oder Kameraposition verändert. Dies wird beispielsweise durch Ambisonic-Formate ermöglicht. Ambience setzt sich aus mono- oder multikanaligen Audio-Komponenten zusammen, die innerhalb einer virtuellen Umgebung platziert werden und sich, wenn erforderlich, auf randomisierten oder vorgegebenen Wegen durch den Raum bewegen können (Bridgett, 2021, S. 11–12). Ambience trägt wesentlich zur Wahrnehmung einer lebendigen und glaubhaften Spiel-Welt bei, indem sie auch in fiktionalen oder stilisierten Umgebungen akustische Erwartungen erfüllt, die bei einer bestimmten Visualisierung entstehen. Durch die konsistente und subtile, hintergrundige Präsenz von Umgebungsgeräuschen wird ein

unterbewusstes Zufriedenheitsgefühl beim Rezipienten erzeugt. Ebenfalls kann Musik als klangliche Gestaltung einer Ambience dienen, wobei hier jedoch auf die Verwendung kontrapunktischer Melodien und perkussiver Elemente verzichtet werden sollte, um die Wahrnehmung eines ebenmäßigen, immersiven Umfelds nicht zu unterbrechen (Marks, 2017, S. 489–490). Die Besonderheit der Ambience in Videospielen besteht außerdem in ihrer Reaktivität. Sie erweitert die Vorstellung des Spielers auf nicht sichtbare Interaktionen in der imaginativen Ferne der Game-Welt – beispielsweise kann die Wahrnehmung entstehen, dass ein akustisches Signal der ambienten Ebene und die damit verbundene Aktion eine kausale Beziehung zu einer darauffolgenden, räumlich näheren Handlung aufweisen. Umgekehrt besteht die Möglichkeit einer reaktiven Ambience auf zentrale Aktionen, wie etwa eine dynamische Publikums-Ambience in Sport-Games, die sich den aktuellen Situationen entsprechend verhält (Bridgett, 2021).

Im Kontext von Videospielen bezieht sich der Begriff der Foley-Sounds auf alle Charaktergeräusche, die durch die Bewegung einer Figur onscreen entstehen. Hierbei muss in der Bearbeitung zwischen Spielercharakter und NPC unterschieden werden – während die Foley-Geräusche des gespielten Charakters gleichbleibend in ihrer Position sowie Lautstärke sind, muss für Foleys der NPCs anhand von Parametern wie Lautstärke, Filter und Positionierung ein dynamisches akustisches Verhalten implementiert werden (Bridgett, 2021, S. 12). Besonders für sich wiederholende Foley-Geräusche wie Schritte besteht die Gefahr eines eintönigen Spielerlebnisses. Um dieses Problem zu umgehen, sollten in der Implementierung Container mit einer Auswahl an Schritt-Samples für verschiedene Untergründe und Schrittgeschwindigkeiten bereitgestellt werden. Diese werden zusätzlich in Pitch, Lautstärke und Filter in Echtzeit randomisiert, wodurch ein abwechslungsreicher Höreindruck entsteht (Leenders, 2012, S. 40–41).

Insbesondere im Bereich des Survival-Horrors nehmen Waffen-Soundeffekte eine zentrale Rolle ein. Als interaktive Objekte müssen diese in Echtzeit akustisch auf die Aktionen des Spielers reagieren. Ihr Sounddesign bezweckt vor allem, dem Nutzer ein Gefühl für ihre Leistungsfähigkeit, Stärke und Energie zu vermitteln. Waffen-Sounds beinhalten ausgelöste Geräusche bei Handhabung und Einsatz der Waffe sowie Impact-Geräusche beim Treffen eines Ziels bzw. Gegners. Die Gestaltung dieser Aspekte erfolgt dabei z. B. für Schusswaffen in folgenden Schichten: Die Detonation stellt den ursprünglichen Schuss dar und enthält einen betonten Transienten, um die Impulsivität des Geräusches zu unterstreichen. Die tieffrequente Sub-Ebene sorgt für die Verstärkung von wahrgenommenem Gewicht und Impact. Auf der mechanischen Ebene werden individuelle mechanische Eigenschaften der Waffe dargestellt, etwa durch die Integration von Lade-, Verschluss- und Patronengeräuschen. Darüber hinaus stellen Parameter wie die auditive Darstellung der

Schussrate sowie eine umgebungsabhängige Hallfahne einen wesentlichen Bestandteil zur Optimierung der Authentizität des Objekts dar (Sinclair, 2020, S. 126–131). Durch eine kontextbasierte Gestaltung der Waffengeräusche, kann ebenfalls dem Charakter-Zustand Bedeutung verliehen werden. Dynamische und reaktive Modifikationen des ursprünglichen Soundeffekts – beispielsweise die Verstärkung tieffrequenter Anteile des Signals in besonders spannungsgeladenen Situationen – tragen zur akustischen Darstellung des emotionalen Zustands der gespielten Figur bei (Bridgett, 2021, S. 73).

Sounds innerhalb der UI- und HUD-Oberflächen liefern dem Nutzer für den Spielerstatus wichtige Hinweise. Da Menüs eine zentrale Rolle für den Handlungsverlauf und die strategischen Entscheidungen des Spielers einnehmen, müssen UI- und HUD-Sounds ein unmittelbares Feedback liefern, um eine eindeutige, zügige Unterscheidung zwischen relevanten und irrelevanten Informationen sowie positiven und negativen Auswahlmöglichkeiten zu gewährleisten. Ebenso empfiehlt sich eine Anpassung der auditiven Elemente dieser Oberflächen an die ästhetische und atmosphärische Ausrichtung des Spiels, um die Aufrechterhaltung des Immersionsgefühls sicherzustellen (Bridgett, 2021, S. 13).

Zur Ergründung der **musikalischen Gestaltung** von Videospielen ist eine Unterscheidung zwischen diegetischer und nicht-diegetischer Musik erforderlich. Diegetische Audioquellen wie Radios, Fernseher oder Live-Musik dienen neben der immersiven Umgebungsgestaltung der Informationsvermittlung über die Beschaffenheit des diegetischen Raums mittels ihrer akustischen Eigenschaften. Der überwiegend nicht-diegetische Soundtrack fungiert als Schnittstelle zwischen der Gefühlsebene des Hauptcharakters und der spielenden Person. Er repräsentiert den Zustand der Charaktere, die Atmosphäre der virtuellen Welt oder einer bestimmten Situation. Dabei kann es zum fließenden Übergang extradiegetischer Musik in die diegetische Sphäre und umgekehrt kommen. Klänge, die zunächst nicht als Ereignis in der Spielwelt vorhanden sind, werden somit aus den Gedanken des Spielers und der Charaktere in eine physikalische Manifestation innerhalb der Spiel-Umgebung überführt, um die Präsenz des Zustandes innerhalb der Diegese zu verankern (Bridgett, 2021, S. 13).

Musik wirkt auf die emotionale Wahrnehmung von Figuren, Situationen sowie auf die narrative Perzeption des Spiels ein. Einzelne Instrumente lenken den Fokus auf spezifische Charaktere, während orchestrale Arrangements die übergeordnete Handlung verdeutlichen (Sinclair, 2020, S. 13–14). Dieser Effekt findet vorwiegend für kurze filmische Zwischensequenzen Anwendung, in denen eine gewisse Stimmung der Figuren und deren Umgebung als auch ein Voranschreiten der Handlung offenbart wird. Musikalische Untermalung während des Gameplays beweckt, je nach

Spiel, unterschiedliche Wirkungen auf den Rezipienten. Im Falle eines Rennspiels dient optimistische Musik mit hohem Tempo und Lautstärke vorrangig der Unterhaltung, während Story- oder Strategiespiele anhand von unaufdringlicher, beständiger Vertonung subtil zur Erhaltung einer spezifischen Atmosphäre und Immersion beitragen (Marks, 2017, S. 237–238).

Die Gruppe **Stimme** umfasst sämtliche vokale Bestandteile, einschließlich Voice-Over, Dialoge, Barks (kurze, sprachliche Ausrufe), Grunts (nichtsprachliche Laute) und Walla (ambiente Hintergrundstimmen). Der Dialog zwischen Haupt- und Nebencharakteren ist auf das narrative System des Spiels abgestimmt, welches durch adaptive Reaktionen und Informationen den Handlungsverlauf unterstreicht und bestimmt. Ein wirksames Mittel zur Übertragung des Innenlebens der Hauptfigur auf die emotionale Wahrnehmung des Spielers besteht in der extradiegetischen Vokalisation der Gedanken des Charakters als sogenanntes Offscreen-Voice-Over. Die zentrale Funktion von Barks ist die Übermittlung variierender NPC-Dialoge, sowohl in direkter Interaktion mit dem Spieler als auch in hintergründigen Gesprächen untereinander. Barks verleihen der Spielwelt sowohl Tiefe als auch Authentizität und ermöglichen die subtile Vermittlung hilfreicher Informationen sowie Handlungsdetails. Die Varietät der Sprache ist ebenfalls bei Lauten physischer Anstrengung zu berücksichtigen, um dem gesteuerten Charakter Leben und Realismus einzuhauen (Bridgett, 2021, S. 16–17).

3.5 Ziele von Sounddesign in Videospielen

Das Sounddesign eines Videospiels, kann Sinclair (2020) zufolge in drei Hauptziele, welche die Qualität des Spielerlebnisses bestimmen, unterteilt werden: Informieren, Unterhalten und Eintauchen bzw. Immersion. Mit dem Zweck, durch Austausch von sachlichen und emotionalen Mitteilungen eine Beziehung zwischen Spieler und Anwendung zu etablieren, stellen diese Ziele Kernkomponenten bei der Entwicklung eines kohärenten und lebendigen Sounddesigns dar (S. 8).

3.5.1 Infomation

In 3D- und VR-Videospielen begrenzt sich das menschliche Sichtfeld im Raum und somit auch die mögliche Aufnahme von Informationen auf 90–120 Grad. Hinzu kommt, dass sich der visuelle Fokus innerhalb dieses Umfangs auf einen schmalen Bereich begrenzt. Daher übernimmt das Audio einen großen Teil der Übermittlung aller sichtbaren und nicht sichtbaren Vorgänge, die dem Hörer wichtige Informationen über den Zustand des Spiels mitteilen (Sinclair, 2020, S. 8–9).

Grundsätzlich ist es von großer Wichtigkeit, dem Spieler Informationen über die Beschaffenheit seiner Umgebung zu vermitteln. Dieser Designprozess wird als Environmental-Modelling bezeichnet und trägt ebenfalls zur Immersionsbildung bei. Dabei geht es weniger darum, umgebungsorientierte Klangeigenschaften perfekt zu simulieren, sondern mehr um die Erschaffung einer in sich schlüssigen, funktionalen und unterhaltsamen Spielumgebung. Jedoch besteht generell die Möglichkeit, ein akustisch genaues Abbild eines Raumes zu generieren. Beispielsweise kann unter Verwendung von Hall-Prozessoren und anhand von Absorptions-Koeffizienten, der im Game dargestellten Materialien, ein realistisches Reflexions-Verhalten errechnet werden.

Eine Möglichkeit zur akustisch passenden Abbildung eines Raumes ist die Anwendung von weniger Pre-Delay für kleine Räume und mehr Pre-Delay bei großen Räumen sowie dichten urbanen Umgebungen zur natürlich wirkenden Simulation von Reflexionen. Dabei sollte das Level der Reflexionen den Eigenschaften der beschallten Materialien entsprechen, also für absorbierende Materialien wie Vorhänge oder Teppiche niedrig und für reflexive Materialien wie Fliesen, Marmor oder Zement höher ausfallen (Sinclair, 2020, S. 215–219)

Wird eine realistische akustische Abbildung der Distanz zu einer Schallquelle angestrebt, sind im Rahmen des Environmental-Modellings folgende Aspekte zu berücksichtigen:

Bei Veränderung der Distanz zwischen Rezipienten und Schallmitter ändern sich die Amplitude einer Audioquelle als auch die spektrale Verteilung und räumliche Wahrnehmung. Der Grad, mit dem sich hohe Frequenzen bei steigender Distanz zur Schallquelle absenken, ist durch die tatsächliche Distanz und Umgebungs faktoren wie Temperatur und Luftfeuchtigkeit bestimmt. Beispielsweise kommt es mit steigender Distanz zu einer allgemeinen Amplitudenabsenkung, einer verringerten Differenz von Amplituden Hoch- und Tiefpunkten sowie zu einer Reduktion hochfrequenter Signalelemente, was zusammen mit einem verstärkten Hallanteil in einem allgemein leiseren, dumpfen, undefinierten Soundereignis resultiert.

Zudem ändert sich mit der Distanz die Exaktheit der Lokalisierung von Schall im 3D-Bereich. Geräuschquellen wie beispielsweise ein Wasserfall können sich aus der Ferne meist präzise im Raum orten lassen. Je näher der Hörer dem Emitter kommt, umso zweidimensionaler bzw. breiter erscheint das Geräusch – es ist schwerer einem genauen Punkt zuzuordnen. Für eine realistische Umsetzung dieser akustischen Distanzempfindung empfiehlt es sich, SFX-Versionen für beide Geräusch-Zustände zu erstellen und mittels eines, an die Distanz zwischen Emitter und Spieler gekoppelten, Crossfades die Versionen adaptiv zu überblenden (Sinclair, 2020, S. 233–234). Weitere Aspekte, die dem Hörer Aufschluss über die Distanz zum Emitter geben und in der Implementierung berücksichtigt werden können, sind beispielsweise das Verhältnis von trockenem zu reflektiertem Signal sowie die ILD und ITD (Sinclair, 2020, 56–57).

Ein weiterer zentraler Bestandteil des informativen Tons sind Nutzer-Feedbacks. Klassische Feedback-Varianten wie Chimes oder Buzzer dienen der akustischen Rückmeldung über Erfolg oder Misserfolg einer Spieleraktion. Positive Klanghinweise geben dem Spieler einen angenehmen, musikalischen Höreindruck, wohingegen negative Hinweise als unangenehm und lärmend erscheinen. Diese Feedback-Cues erfüllen die Funktion, in einer geräuschintensiven Umgebung die Aufmerksamkeit des Spielers auf sich zu ziehen und ihren semantischen Inhalt verständlich zu vermitteln. Darüber hinaus können sie subtile Hinweise auf die Bedeutung eines Objekts oder einer Situation geben und unterschwellige Gefühle im Spieler hervorrufen. Der bewusste Einsatz des wechselnden Verhältnisses von Sound, Musik und Stimme kann den Spieler dazu bewegen, sich auf bestimmte Aspekte zu konzentrieren, oder ihn davon ablenken (Sinclair, 2020, S. 11–12).

In actionreichen Genres wie Horror, ist der Spieler darauf angewiesen, durch auditive Signale auf bevorstehende Gefahren vorbereitet zu werden. Häufig manifestiert sich diese Warnung in Form eines akustischen Sounds, also eines Geräusches, dessen Quelle nicht oder noch nicht sichtbar ist. Die Spielerhandlung wird somit bereits vor dem Eintreten eines visuellen Events beeinflusst, indem Art, Lautstärke und Position des Stimulus den Spieler zum Fliehen oder Weitergehen in eine bestimmte Richtung bewegen. Ebenso bezweckt der Einsatz von symbolischem Sound das gezielte Fokussieren des Hörers auf bestimmte Informationen über Charaktere, Stimmungen, Umgebungen und Objekte. Stille kann hierbei ebenfalls als bedeutungsvolles akustisches Feedback fungieren, indem sie dem Spieler signalisiert, dass ein Spielsegment abgeschlossen ist, es nichts mehr zu entdecken gibt und ein Fortschreiten innerhalb der Spielwelt erforderlich ist. Es ist beizufügen, dass über die klassischen Sound-Cues hinaus Musik, Geräusche und Stimmen allgemein dazu beitragen, Informationen über kulturelle, physikalische, soziale und historische Aspekte eines Umfeldes zu vermitteln (Collins, 2008, S. 130–132).

3.5.2 Immersion

Grundsätzlich bezieht sich der Begriff der Immersion auf das Gefühl des Eintauchens in eine Welt, die von der physikalischen Umgebung abweicht, wobei sich der Rezipient temporär nicht mehr des Mediums hinter der künstlichen Umgebung bewusst ist. Das Eintauchen in eine Welt kann über sensorische Empfindungen hinaus auch durch narrative Elemente erfahren werden. Im Gaming-Kontext wird Immersion in Abhängigkeit des Spielsystems auf verschiedenen Ebenen betrachtet. Ein 3D-Videospiel, welches auf einem PC oder einer Konsole gespielt wird, ist als non-immersiv klassifiziert, ein VR-Spiel dagegen wird aufgrund seiner zusätzlichen Komponenten, die den Spieler fast komplett von seiner Umgebung abschneiden, als vollständig immersiv bezeichnet (Sinclair, 2020, S. 14). Bridgett (2021) zufolge wird das Eintauchen des Nutzers auf der Ton-Ebene durch die synchrone Übertragung der Soundeffekte zu den im Spiel durchgeführten Aktionen bewirkt.

Infolgedessen werden die akustischen Erwartungen des Spielers erfüllt, was die Aufrechterhaltung der Immersion unterstützt. Die Umsetzung dieser immersiven Erfahrung durch Sound hängt dabei von den Erwartungen und Vorgaben des jeweiligen Genres ab (S. 65). Auditive Bestandteile der Immersionsbildung lassen sich in zwei Kategorien unterscheiden: Zum einen gewährleisten Elemente, die dem Spieler nicht unbedingt bewusst auffallen, wie Atmosphäre und Score, ein vollständiges, gehaltvolles Modell der Spielumgebung. Neben musikalischen Aspekten sorgt die Vielfältigkeit der Geräusche als auch deren räumliche und zeitliche Randomisierung für die Immersionssteigerung einer Umgebung. Zum anderen tragen in sich konsistente und zur visuellen Erscheinung passende Sounds von interaktiven Objekten oder Charakteren zu einem stimmigen Gesamtbild bei. Dabei sollten jedoch situationsabhängige Variationen gegeben sein, um die Erwartung des Spielers zu bestätigen und die Immersion zu erhalten: Ein Monster sollte in seiner vokalen Gestaltung konsistente Merkmale aufweisen, sich jedoch in seiner Emotion, beispielsweise bei einem Schrei aus Schmerz oder aus Wut, klar unterscheiden. Ungereimtheiten im Sounddesign führen schnell zu einem Bruch der Spieler-Immersion, da sie die Verlässlichkeit von Hinweisen verringern und eine stärkere Orientierung an visuellen Anhaltspunkten erzwingen. Daher sollten unpassende auditive Signale, Inkonsistenzen sowie Unterbrechungen vermieden werden. So ist es beispielsweise förderlich, atmosphärische Musik bei Tod der Spielfigur des Nutzers nicht zu unterbrechen, sondern über den Spielverlauf kontinuierlich weiterzuführen, um eine beständige Klangerfahrung zu bewahren (Sinclair, 2020, S. 15–16).

3.5.3 Unterhaltung

Neben der Gestaltung einer kohärenten und informativen Spielumgebung erfüllt Game-Audio eine zentrale narrative und emotionale Funktion. Sound steuert die Art und Intensität der affektiven Wirkung auf die spielende Person, wobei die Definition einer unterhaltsamen Audio-Szenerie von der gegebenen Spielsituation und Zielgruppe abhängig ist. Über den Verlauf des Mediums lenkt der wechselnde Fokus auf einzelne Bestandteile des Mixes die Aufmerksamkeit und Empfindung der Empfänger, wodurch gewünschte emotionale Aspekte einer Szene gezielt hervorgehoben werden können. Besonders Musik dient hierbei zur Beeinflussung der Spielerempfindung und Handlungsgestaltung in Videospielen (Sinclair, 2020, S. 12–13).

Wie bereits erläutert, intendiert das Genre Horror vorrangig die Induktion von Angst durch das Eintauchen in gefährliche oder beklemmende Situationen von einer sicheren Umgebung aus, wodurch die Rezipienten Unterhaltung erfahren. Das folgende Kapitel stellt einen Bezug zwischen dem Horror-Genre und den erläuterten Bestandteilen sowie Funktionen des Sounddesigns in

Videospielen her. Dabei werden technische wie auch psychologische Ansätze in der Tongestaltung dargestellt, die innerhalb der Videospiel-Kategorie zur gezielten Erzeugung von Emotionen wie Angst und Unsicherheit beitragen.

4. Audio in Horror-Videospielen

Die Effektivität der gefühlsmäßigen Wirkung des Horror-Genres lässt sich auf diverse Einflussfaktoren zurückführen, die von Audio bis hin zur strukturellen Gestaltung der Handlung reichen. Ein zentraler Wirkungsmechanismus des Horrors resultiert aus dem Interpretationsspielraum des Mediums. Die gezielte Zurückhaltung von Informationen regt die Imagination des Spielers dazu an, verstörende Vermutungen und Erklärungen für das Geschehen der fiktiven Welt zu konstruieren. Häufig ist das Horror-Setting in einer familiären Umgebung verankert, die durch das Eindringen bösartiger Entitäten destabilisiert wird. Die sich daraus formende alternative Realität legitimiert – im Gegensatz zu anderen Genres – die Integration individueller Mechanics wie übernatürlicher Elemente oder besonderer Fähigkeiten. Meist wird anhand psychischer Zustände des Charakters begründet, warum dieser einen gefährlichen Ort nicht einfach verlassen kann, was aufgrund der Beschaffenheit der Horror-Spielwelten vom Spieler als plausibel beurteilt wird. Dem Spieler werden limitiert Informationen über die Umwelt und deren Geschehen vermittelt, um eine vollständige Orientierungslosigkeit gegenüber bevorstehenden Gefahren zu vermeiden. Insbesondere psychologische Horror-Spiele konstruieren die Hintergrundgeschichte mittels subtiler Hinweise wie Dialoge, Graffiti an Wänden, Tagebucheinträge oder Flashbacks, wobei der Spieler durch informative Lücken stets zu seiner eigenen Interpretation gelenkt wird, was zur Erhaltung einer angespannten Atmosphäre beiträgt. Darüber hinaus nutzt das Horror-Genre oftmals die Angst vor Dunkelheit und Nebel sowie vor den potenziellen Feinden innerhalb der verschleierten Bereiche (Rouse, 2009, S. 14–16). Neben den erläuterten grundlegenden Zwecken des Sounddesigns in Videospielen begibt sich der Ton spezifisch im Horror auf eine psychologische Ebene, um die Wahrnehmung des Rezipienten durch den Einsatz von Falschinformationen und Mehrdeutigkeit zu manipulieren. Soundeffekte werden dabei entgegen den Erwartungen der Spielenden angewendet, sodass etablierte Regeln des Sounddesigns und der Musik im Kontext des Mediums keine Anwendung mehr finden (Roberts, 2014, S. 144). Das folgende Kapitel bietet einen Überblick über angewandte und mögliche auditive Gestaltungsmittel dieser Art.

Die beispielhafte Nennung von *Silent Hill 2* bezieht sich im folgenden Abschnitt ausschließlich auf das Videospiel von 2001, lediglich das fünfte Kapitel dieser Arbeit behandelt die Neuadaption des Originals.

4.1 Timing und Lautstärke

Toprac & Abdel-Meguid (2011) differenzieren in der Emotionsäußerung auf auditive Stimuli zwischen Furcht und Angst. Dabei definieren sie Furcht als eine emotionale Reaktion auf eine wahrgenommene Bedrohung, während Angst als ein genereller Gefühlszustand ohne einen klar identifizierbaren Stimulus auftritt (S. 189).

Das Timing von Soundeffekten kann über die Funktion einer immersiven, in sich schlüssigen Spielumgebung hinaus manipuliert werden, um bestimmte emotionale Absichten umzusetzen. Das Timing von SFX wird dabei als der Grad an Synchronisation zwischen einem Sound und dessen referenziertem sichtbaren Objekt definiert und ist in drei Ebenen gegliedert: Im Falle plötzlicher Begegnungen mit Feinden wird durch die Verwendung einer präzisen Synchronisation von auditiven und visuellen Reizen ein Überraschungs- bzw. Schreckmoment ausgelöst und die Wahrhaftigkeit der Situation verstärkt, was eine gesteigerte Furchtreaktion erzielt. Gestaltet sich das Timing jedoch leicht versetzt, äußert sich die emotionale Reaktion aufgrund der Ungewissheit über die Eigenschaften des Emitters in Anspannung. Die dritte Kategorie beschreibt SFX, die zufällig, ohne jeglichen visuellen Zusammenhang abgespielt werden, wie Feuerknistern, Windpfeifen oder bebende Erde. Auf diese Kategorie geht die vorliegende Arbeit gesondert in den Kapiteln 4.5 und 5.2.4 ein.

Ebenso spielt die Lautstärke bzw. das Dynamikverhältnis aller tonalen Ebenen zueinander eine wichtige Rolle auf der SFX-Ebene des Horrors. Häufig steht dabei Audio, welches eine emotionale Komponente vermitteln soll, im Mix vor atmosphärischen Klängen, um die Aufmerksamkeit des Rezipienten auf sich zu ziehen. Beispielsweise werden Musik und ambiente Komponenten nach einem lauten Event stark in ihrem Pegel reduziert oder verstummen komplett, mit dem Zweck, das vorherige Geschehen in seiner emotionalen sowie narrativen Bedeutung zu unterstreichen. Erhöhte Lautstärke von Soundeffekten wirkt besonders effektiv in der Induktion plötzlicher, schockierter emotionaler Reaktionen sowie der Hervorhebung von Leere in einer stillen Umgebung. Ambiente Soundeffekte dienen hingegen mit einem generell niedrigen Pegel der Immersions- und Stimmungsbildung und signalisieren dem Spieler gefährliche Umstände, wodurch Anspannung und Angst ausgelöst werden (Toprac & Abdel-Meguid, 2011, S. 180–181; Roberts, 2014, S. 145–149). Die Autoren resümieren ihre Analyse in der Erkenntnis, dass laute Soundeffekte, welche einen hohen Grad an Synchronisation aufweisen, vorrangig die Intensität des Furchttempfindens steuern, wohingegen leise, spukhafte und mehrdeutige Stimuli ohne visuelles Gegenstück subtil ein Gefühl der Angst beim Hörer erwecken. Es ist zu vermerken, dass die Furchtreaktion bei synchronen audiovisuellen Ereignissen primär durch deren visuelle Anteile verursacht wird, jedoch eine Intensivierung durch die akustischen Elemente erfährt (Toprac & Abdel-Meguid, 2011, S. 182).

4.2 3D-Audio

Wie zuvor erläutert, beschreibt räumliches Audio für den Hörer eine maßgebliche Komponente zur Lokalisierung von Entitäten innerhalb der Spielwelt. Die Richtung, aus welcher der Hörer mit einem akustischen Reiz beschallt wird, steht in direktem Zusammenhang mit der wahrgenommenen Unheimlichkeit des Audios. Mit diesem Verhältnis befasst sich eine Nutzerstudie von Ekman und Kajastila (2009), in welcher raubtierähnliche Soundsamples auf Breit- und Punktschallquellen in verschiedenen Lautsprecher-Setups abgespielt sowie jeweils hinter und frontal zur Hörerposition platziert wurden (S. 3–4). Die Auswertung der Nutzerbewertungen ergab, dass Punktschallquellen von hinten tendenziell als unheimlicher eingestuft werden im Vergleich zu identischen, frontalen Quellen. Allerdings wird diese Empfindung bei breiten Schallquellen ausgeglichen, da hier die Unterscheidung zwischen vorderen und hinteren Stimuli erschwert ist. Besonders für hinter dem Hörer emittierte Geräusche ist ein signifikanter Anstieg der Unheimlichkeitswahrnehmung bei breiten Schallquellen festzustellen (Ekman & Kajastila, 2009, S. 5).

Die Autoren bezeichnen Emotion im Kontext von Videospielen in ihrer Herangehensweise als „a dynamic, parametrically controlled feature“ (Ekman & Kajastila, 2009, S. 5). Die Studie führt mit der systematischen Nutzung von Richtungsparametern ein Beispiel für eine mögliche parametrische Emotionssteuerung in Videospielen auf und schlägt eine Adaption der Erlebnisintensität basierend auf dem Nutzerverhalten oder einstellbaren Präferenzen vor. Unter Berücksichtigung der bereits behandelten sowie der folgenden auditiven Methoden stellt die richtungsabhängige Wahrnehmung eine von vielen Komponenten dar, die bei der Entwicklung und Optimierung einer solchen parametrischen Steuerung zu berücksichtigen sind (Ekman & Kajastila, 2009, S. 5).

Brinkman et al. (2015) schildern in ihren Untersuchungen zu emotionalen und räumlichen Wahrnehmungsfaktoren verschiedener Audiosysteme folgende Beobachtungen: Vor allem für dreidimensionales Audio lässt sich auf Grundlage gemessener Attribute wie Lokalisation, Distanz, Externalisation, Bewegung und Raumwahrnehmung eine gesteigerte spatiale Perzeption im Vergleich zu Stereo und Dolby Surround verzeichnen. Zudem zeigte sich für das 3D-Hörbeispiel der Ermittlung generell ein erhöhter SUD-Wert (Subjective Units of Distress), dessen Ausmaß innerhalb der Skala in Korrelation mit den individuellen Ängsten der Testpersonen variierte. Die Autoren kommen in ihrer Untersuchung zu dem Schluss, dass Sound als alleinstehender Stimulus Angstreaktionen hervorrufen kann, wobei der Grad der Räumlichkeit eines Audiosystems Einfluss auf die subjektive Hörerfahrung ausübt.

Im Kontext einer VR-Umgebung lässt sich durch Sound eine Steigerung der emotionalen Immersion sowie der Reaktion feststellen. Es ist jedoch zu berücksichtigen, dass die Studie nur ein einzelnes Hörbeispiel (das Geräusch einer fliegenden Wespe) verwendet. Das wiederholte Hören des Samples

könnte, wie in Kapitel 2.1.3 dargelegt, durch Vorbereitung auf den akustischen Stimulus die Valenzreaktion auf den Reiz abgeschwächt haben. Die Argumentation der Autoren, dass zwischen Stereo und 3D-Sound hinsichtlich der Verbesserung individueller Hörerfahrungen kein signifikanter Unterschied zu erwarten sei – weder für Ton als einzelner Reiz noch für audiovisuelle Stimuli – ist daher ebenfalls kritisch zu hinterfragen (Brinkman et al., 2015, S. 46–48).

Zwar existieren nur wenige wissenschaftliche Belege zur Überlegenheit von 3D-Audio gegenüber Stereo hinsichtlich emotionaler Reaktionen, jedoch ist diese These insbesondere in interaktiven Horror-Medien durch Entwickler- sowie Spieler-Einschätzungen eindeutig widerlegt. Spieleentwickler Morisama Sato betont in einem Interview mit dem PlayStation Official Magazine die Unerlässlichkeit von 3D-Audio im Horror-Genre:

In horror, creating the sense of a presence beyond the visual scene on screen is essential, so I think 3D audio will be indispensable in horror games from now on. (Palumbo, 2021)

Sato vertritt die Meinung, Immersion stelle den wichtigsten Aspekt dieses Genres dar, insbesondere für explorative Werke im Bereich des Survival-Horrors. Der Einsatz von 3D-Audio zur Verstärkung visuell und narrativ detailreicher Gegebenheiten des Sub-Genres spiegele sich in einer als mehrschichtig empfundenen Umgebung wider und sei essenziell für eine intensive Wahrnehmung einzelner, unheimlicher Soundeffekte (Palumbo, 2021).

Eine spezifische Methode im Bereich des räumlichen Horror-Audios stellt die direktionale Soundmanipulation dar. Direktionale Soundeffekte weisen den Hörer häufig auf die Position und Entfernung einer feindlichen Präsenz außerhalb des Sichtbereiches hin. Zudem kann sich der Spieler meist nur durch strategisches Verstecken vor der Bedrohung schützen und gelangt so in einen anhaltenden Zustand von Terror, da er nie Gewissheit darüber hat, aus welcher Richtung der nächste Angriff zu erwarten ist oder ob eine Konfrontation tatsächlich beendet ist. Richtungsbasierte Manipulation von Sound beschreibt in diesem Kontext beispielsweise einen auditiven Hinweis, welchem der Spieler für gewöhnlich in die dementsprechende Richtung nachgeht. Wird daraufhin der tatsächliche Angriff des Gegners aus einer abweichenden Richtung ausgelöst, resultiert dies in einer Panikreaktion. Des Weiteren führt diese Technik zu einer verstärkten Immersion, da die Unvorhersehbarkeit der Bedrohung die volle Aufmerksamkeit des Spielers fordert. Zudem kann direktionale Mehrdeutigkeit den Spieler gezielt ablenken oder dazu bringen, kontinuierlich nach der Emissionsquelle zu suchen, und somit Verwirrung und Anspannung bewirken (Roberts, 2014, S. 144–145).

4.3 Vorwarnung

Auditive Vorwarnungen befähigen den Rezipienten nicht nur durch ihre Informationsebene, auf bevorstehende Ereignisse zu reagieren, sondern beschreiben ebenso eine signifikante Methode zur Bewirkung von Verunsicherung und Angst. Das Phänomen der antizipatorischen Angst im Horror durch Vorwarnung wurde bereits seit vielen Jahren in einigen Studien wie der von Opton et al. (1968) beobachtet. Diese belegt, dass die Existenz einer Vorwarnung über ein bevorstehendes Ereignis deutlich mehr Beunruhigung bei den Rezipienten verursacht als eine Variante ohne jegliche Ankündigung auf eine Szene. Für vorgewarnte, furchterregende Events erlangt die physiologische Stressreaktion in der Antizipationsphase einen bemerkenswerten Anstieg im Vergleich zur Reaktion im Moment der tatsächlichen Erfahrung der Szene (Opton et al., 1968, S. 206–207).

Der erläuterte Effekt kann ebenso auf auditive Warnungen in Videospielen übertragen werden. Diese stellen sich jedoch dank der Nichtlinearität des Mediums hinsichtlich ihrer Funktionen und Auswirkungen als weitaus vielschichtiger dar. Der folgende Abschnitt geht diesen Besonderheiten nach.

Die Verwendung akustischer Warnsysteme in Horror-Spielen ist primär auf das Wissen des gesteuerten Charakters abgestimmt. Dessen Umgebung mit ihren allgegenwärtigen Bedrohungen betont die Notwendigkeit von audiovisuellen Vorwarnungsmechanismen. Am Beispiel von *Fatal Frame* (Tecmo, 2001) ist ein Warnsystem zu erkennen, welches auf verschiedenen sensorischen Signalen basiert: einem sich farblich verändernden Item-Filament, Vibrationsfeedback des Controllers sowie einem simultan ausgelösten Herzschlag-Sound und gespenstischen Geräuschen. Dieses System tritt in Kraft, sobald sich der Spielcharakter in der Nähe einer Bedrohung, in diesem Fall Geister, befindet. Da der Warnmechanismus auch bei räumlicher Separierung zu nahen feindlichen Entitäten ausgelöst wird, entsteht eine antizipatorische Angst hinsichtlich dessen, was sich im nächsten Raum befinden könnte. Andere Werke, wie jene der *Silent Hill*-Reihe, stellen dem Charakter ein spezifisch als Warnsystem fungierendes diegetisches Objekt zur Verfügung. Das Radio in *Silent Hill* (Konami, 1999) stellt einen besonders detaillierten Mechanismus dar, da sich sein Output in Frequenz, Tonhöhe, Lautstärke, sowie der Distanz zum Feind, ab der es getriggert wird, an die Art der Bedrohung adaptiert. Zum einen ermöglicht das Radio somit eine präzise, oft unterbewusste Vorahnung über bevorstehende Konfrontationen mit Gegnern. Zum anderen generieren unvorhersehbare Ausfälle oder Manipulationen des Geräts, beispielsweise an unterirdisch gelegenen Orten oder beim Auftreten mehrdeutiger Signale, Verunsicherung und Unbehagen beim Spieler hinsichtlich der Korrektheit und Zuverlässigkeit des Warnsystems. Vor allem in einem solchen Szenario oder in einer Situation mit eingeschränkter Sichtweite fungieren feindliche Ausrufe oder Schrittgeräusche offscreen als Indikatoren für eine verborgene Bedrohung.

Daher vermitteln Audio-Hinweise sowohl Gewissheit über den bevorstehenden Konflikt mit dem Monster als auch Unsicherheit über den Ausgang der Situation.

Perrons (2004) Untersuchungen bestätigen die These, dass die Vorbereitung auf eine beklemmende Situation die emotionale Anspannung des Spielers gegenüber dieser verschlimmert (S. 4–6). Grimshaw (2011) fügt hinzu, um in dieser Hinsicht effektiv zu sein, müsse Vorwarnung vor allem unzuverlässig sowie in ihren Informationen über die Positionierung der Quelle begrenzt sein. Unter Berücksichtigung dieser Aspekte entstehe eine neue Funktion der Vorwarnungen: Das Kögern des Spielers durch Sound. Beispielsweise führt der Verzicht auf die Adaption akustischer Eigenschaften in Relation zur Charakterposition zu einer starken Limitierung der im Signal enthaltenen positionellen Informationen. Dies resultiert in einer gestörten Wahrnehmung der Distanz zur Bedrohung und somit in einer verstärkten Spannungswirkung der auditiven Vorwarnung sowie einer vorsichtigen Spielerhandlung (S. 205).

Aus den dargelegten Warn-Modellen der Horror-Spiele sowie der vorangegangenen Untersuchung des emotionalen Sounds entnimmt die Autorin folgende hypothetische Anwendungszwecke: Zunächst können, besonders in simulierten Gefahrensituationen, rhythmische Elemente wie ein Herzschlag eine physiologische ebenso wie eine emotional-empathische Kopplung der Spieler-Reaktion an den Zustand des Charakters bewirken. Des Weiteren erzeugen stark begrenzte oder inkorrekte Informationen in auditiven Vorwarnungen Spannung und Verunsicherung. Zu Beginn sehr akkurate, verlässliche Warnsignale bauen das Vertrauen des Rezipienten auf, um daraufhin, mittels plötzlichen Entziehens der akustischen Reliabilität, den Spieler den Halt auf dieser Wahrnehmungs-Ebene verlieren zu lassen.

Kockesen (2015) ermittelt diesen Effekt anhand eines Vergleiches von korrekten und inkorrekten Audio-Hinweisen über den Spielstatus in einem Survival-Horror-Spiel. Für die abweichende Version des Spiels mit inkorrekttem Warnsystem, wurde beispielsweise ruhige Musik verwendet, wenn sich Bedrohungen in Spielernähe befanden, sowie ein zufällig abgespieltes Monster-Knurren in kompletter Abwesenheit der Monster (S.6). Die Auswertung der Angst-Level-Bewertung ergab, dass die erste Einwirkung des Monster-Knurrens, primär in dessen abweichendem Zustand, jedoch auch in dem einer tatsächlichen Bedrohung, als am unheimlichsten von allen akustischen Warnungen empfunden wurde. Zudem konnte für das erneute Auslösen eines hochfrequenten Geräusches im inkorrekten Zustand ein erhöhtes Angstniveau beobachtet werden. Dies lässt sich darin erklären, dass die zufällig ausgelösten Signale die Versuchspersonen anfänglich eher verwirren, jedoch zu bestimmten Zeitpunkten im Spiel als scheinbar relevant bezüglich der Gefahrenvermittlung eingeschätzt werden. Ein signifikanter Anstieg in der Angsterfahrung zeigte sich insbesondere bei Testpersonen, welche zunächst die Version mit akkurate Sound-Cues spielten und anschließend

jene im Zustand des inkorrekten Warnsystems. Die Autorin der Studie führt dieses Ergebnis darauf zurück, dass die Hörer sich eine Interpretation der zunächst zuverlässigen auditiven Hinweise aneignen, welche sich allerdings im weiteren Verlauf als unzuverlässig herausstellen. Dieser Ablauf führe zu einem Zusammenbruch des Vertrauens in die Sound-Ebene sowie in Gefühlen von Überraschung und Angst (Kockesen, 2015, S. 9–12).

Maßgeblich für die emotionale Wirkung in Horror-Videospielen ist demnach weniger die Relevanz der gegebenen Signale, sondern vielmehr deren Inkonsistenz, Reihenfolge und Kontext sowie die daraus resultierenden Schlussfolgerungen der Spieler. Besteht eine Kopplung der Warnsystem-Signale an bekannten Gefahren, entsteht neben dem Spielergefühl von Angst auch konkrete Furcht.

4.4 Transdiegetisches Audio

Über die als klar diegetisch oder nicht-diegetisch identifizierten Sounds hinaus besteht eine Überschneidung der beiden Kategorien im sogenannten transdiegetischen Bereich. Jørgensen (2007) definiert transdiegetisches Audio als eine temporäre Überschreitung der gewohnten Grenze zwischen diegetisch und nicht-diegetisch. Die Autorin differenziert die transdiegetische Ebene anhand der Art der Soundquelle in interne und externe Komponenten. Interne transdiegetische Elemente umfassen diegetisches Audio, dessen Kommunikation weniger auf die Entitäten innerhalb der Spielwelt, sondern vielmehr auf die Spieler in der externen Realität, der Extradiegesse, ausgerichtet ist. Es erfolgt beispielsweise eine Reaktion der diegetischen Charaktere auf Aktionen des Rezipienten, welche sich häufig in Form von sprachlichen Hinweisen über den psychischen Zustand der Figur manifestiert. Konträr dazu bezeichnet die externe transdiegetische Ebene extradiegetische Sounds, die eine hohe Relevanz für die internen Geschehnisse aufweisen. Diese Signale haben den Anschein, sich mit diegetischen Charakteren zu verständigen, sowie dem Spielenden Informationen und Warnhinweise zu vermitteln, die meist auf bevorstehende, wichtige Ereignisse hindeuten. Ein Beispiel hierfür stellt adaptive Hintergrundmusik dar (S. 113–116). Insbesondere die Zuordnung von Ambience zu einem klar diegetischen oder extradiegetischen Raum stellt eine Herausforderung dar, da sie nicht nur Umgebungsgeräusche vermittelt, sondern darauf abzielt, die generelle Atmosphäre einer Umgebung einzufangen. Musikalische Bestandteile und Geräusche einer Ambience sind nicht fest in einer Diegese-Ebene verankert. Vielmehr wird ihr wahrgenommener Ursprung durch das Zusammenspiel oder Gegenwirken mit der visuellen Spielumgebung beeinflusst. Transdiegetische Ambience trägt im Horror zur Induktion von Verunsicherung und Nervosität bei, indem sie entweder die Emotionen der Spielfigur metaphorisch reflektiert oder durch auditive Kohärenz mit der dargestellten Welt die Präsenz diegetischer

Geräuschquellen wie etwa Geister suggeriert, die dem Rezipienten noch nicht eindeutig offenbart wurden (Jørgensen, 2007, S. 111–114).

Kromand (2008) betont, dass besonders der Einsatz atonaler Ambience die Erwartung des Rezipienten an die üblicherweise als eindeutig extradiegetisch erkennbare Hintergrundmusik nicht erfüllt, was ein Gefühl von Unsicherheit bewirkt. Zudem erweist sich ambienter Sound im Horror häufig als situationsabhängig und unregelmäßig, wie Kromand am Beispiel von *BioShock* (2K Games, 2007) weiter ausführt. Die beständige Ambience von *BioShock* wird durch irreguläre tonale Elemente wie eine monotone, im Pegel anschwellende Note ergänzt, um mittels der antrainierten Assoziation des klassischen Horror-Stillmittels antizipatorische Angst beim Spieler zu evozieren. Das Crescendo-Element wird in einem zeitlichen Intervall und unabhängig von der narrativen Handlung aktiviert, was zu einer gesteigerten Anspannung führt, wobei der Sound in bestimmten Situationen Relevanz für die Vorbereitung auf herannahende Feinde hat, in anderen Fällen vergeht die Warnung jedoch ohne konkrete Bedrohung. Diese Faktoren können zwar als Antizipation des Spielgeschehens aufgefasst werden, sie bewirken jedoch vorrangig ein Misstrauen des Spielers gegenüber der Soundquelle und ihrem konkreten Ursprung. Ein weiterer Mechanismus, welcher die Manipulation des Spielers mithilfe von Ambience verfolgt, ist für *Silent Hill 2* zu beobachten. Hier signalisiert das vom Charakter mitgeführte Radio mittels statischen Rauschens die Präsenz von Feinden, versagt jedoch vereinzelt. Beim Betreten mancher Räume wird eine dem Radiosignal ähnliche Ambience angewendet, in anderen Räumen wird das Rauschen, wie ursprünglich erwartet, durch ein Monster darin getriggert. Außerdem erzeugt das zeitweise Hinzumischen von Vogelgezwitscher und quietschenden Rädern in das Rauschen den Anschein einer vom Radio unabhängigen Ambience (S. 17–18).

Kromand (2008) erweitert Jørgensens (2007) Modell der transdiegetischen Ebene, indem er aufzeigt, dass der Verlust der Fähigkeit, Audio als klar diegetisch oder nicht-diegetisch einzustufen, nicht nur eine zeitliche Überschreitung, sondern einen Zusammenbruch der Diegese-Barriere bewirkt. Manipulierte Warnsysteme sowie atonale und unmelodische Ambience erzeugen durch ihren transdiegetischen Effekt ein anhaltendes Gefühl der Unsicherheit beim Rezipienten (S. 18). Ein weiteres Phänomen, das mehr als transdiegetischer Prozess charakterisiert werden kann, ist die sukzessive Offenlegung von zunächst extradiegetischem Audio als diegetische Klangquellen im Verlauf der Wahrnehmung. Dies kann als narratives Gestaltungsmittel fungieren oder die Unsicherheit der Figuren hinsichtlich der Quelle zum Ausdruck bringen (Bridgett, 2021, S. 82).

Die Erzeugung einer unterschwelligen, intern transdiegetischen Wirkung kann besonders durch die Integration musikalischer Elemente innerhalb der Diegese realisiert werden. Als Bestandteil des Set-Dressings einer Szene legt der Charakter mittels der diegetischen Klangelemente um sich herum

dem Spielenden seine Persönlichkeit dar oder vermittelt mit der Musik assoziierte kulturelle Identitätsaspekte. Diegetische Musik trägt durch ihre Kommunikation mit der Extradiegesse wesentlich zur Stimmungsbildung des Hörers bei. Beispielsweise versetzt der Transfer einzelner Score-Bestandteile in den 3D-Raum als auch die Bewegung dieser aufgeteilten Audioquellen um die Spielerposition herum den Hörer in Zweifel hinsichtlich des tatsächlichen Ursprungs der Quelle innerhalb der diegetischen Ebenen. Besonders atonale und fremdartige Quellen werden als vom Score separierte Elemente wahrgenommen, was den Spieler dazu veranlasst, in Unbehagen die Realität zu hinterfragen (Bridget, 2021, S. 130–131).

4.5 Simulation veränderter Bewusstseinszustände

In der Sparte der psychologischen Horror-Games steht meist das mentale Befinden des Charakters im Zentrum des Geschehens und stellt damit einen wichtigen, zu vertonenden Aspekt für die Spannungsbildung dar (Picard, 2009, S. 109). Eine effektive gestalterische Technik zur psychologischen Beeinflussung der Charakter- sowie Erlebniswahrnehmung des Spielers besteht in der Simulation akustischer Halluzinationen. Auf dem Gebiet der Audiologie definieren sich diese als die Wahrnehmung von Geräuschen, Musik und anderen Klängen ohne das Vorhandensein externer akustischer Stimuli. Akustische Halluzinationen können unabhängig von der mentalen Verfassung des Betroffenen auftreten, werden jedoch insbesondere in ihrer verbalen Erscheinungsform mit Erkrankungen wie Schizophrenie assoziiert (Musiek et al., 2007, S. 32).

Im Kontext von Videospielen bedienen sich Sounddesigner vorwiegend mittels Stimmen, die aus dem Nichts zu kommen scheinen, der Angst des Unbekannten, um so dem Rezipienten eine auditive Halluzination zu simulieren. Tonale Stimuli ohne sichtbaren Emitter, die außerdem nicht als extradiegetisches Element erkennbar sind, hindern den Spielenden daran, eine zeitliche sowie räumliche Beziehung zwischen den Reizen herzustellen (Toprac & Abdel-Meguid, 2011, S. 182).

Die angewendeten Halluzinations-Effekte betonen die verlassene Spielumgebung und narrative Gegebenheiten, wie am Beispiel von *Silent Hill* in Form von Geräuschen wie zuknallenden Türen, Flüstern, Schluchzen oder dem Betteln der Tochter des Protagonisten, gerettet zu werden. *Silent Hill: Homecoming* (Konami, 2008) nutzt ebenfalls vermehrt vokale Halluzinationen und fokussiert diese mittels des Einsatzes von Babygeschrei, Kinderstimmen und Schlafliedern auf den psychologischen Kernpunkt der Geschichte – die Kindheit. In Spielen der *Fatal Frame*-Reihe weisen vokale Elemente wie gespenstische Stimmen, Geflüster und Schreie auf übernatürliche Elemente hin. Auditive Halluzinationen wie diese enthalten oftmals provokative Inhalte oder Warnungen. Sie

beschreiben einen effektiven Mechanismus zur Repräsentation und Induktion negativer emotionaler Reaktionen oder dienen der Nachbildung tatsächlicher Halluzinationen oder Gemütszustände der Spielfiguren (De Lima & Demarque, 2013, S. 20–23).

Die Wirksamkeit der Übertragung des audiologischen Phänomens auf Videospiele lässt sich anhand einer Nutzerstudie von De Lima und Demarque (2013) belegen. Die Studie implementierte auditive Halluzinationen, ähnlich denen der zuvor analysierten Fallbeispiele, innerhalb eines simplen Horror-Spiels und stellte diese Version einer ohne die verbalen Stimuli gegenüber. Die Auswertung der selbstberichteten emotionalen Reaktionen ließ dabei für das Szenario mit Präsenz akustischer Halluzinationen eine deutliche Steigerung der Angst sowie des Angstverhaltens erkennen (De Lima & Demarque, 2013, S. 23–25). Aufgrund ihres geringen Umfangs zieht die Nutzerstudie akustische Halluzinationen nicht im Kontext einer vielschichtigen Soundlandschaft, sondern als isolierte akustische Reize zur Bewertung heran und beschränkt sich auf eine einseitige visuelle Umgebung ohne bedrohliche Komponenten. Daher ist es möglich, dass in der Studie ermittelte Ausmaß der affektiven Reaktionen auf die ersten Momente eines Videospiels zu beziehen, jedoch bedarf es für eine umfangreiche Analyse des Kontexts einer audiovisuell komplexen Spielumgebung. Da das Auslösen der auditiven Reize in Abhängigkeit vom Eintreten des Spielers in bestimmte Regionen implementiert wurde, kann nur vermutet werden, dass eine willkürliche Emittierung der sprachlichen Samples die Angstreaktionen erhöht hätte.

Unter dem Konzept veränderter Bewusstseinszustände versteht man neben Träumen und Halluzinationen ebenso die Wahrnehmungsmanipulation externer Sinnesreize, beispielsweise im Zuge eines Rauschzustandes. Weinel und Cunningham (2017) veranschaulichen die Formen, in denen akustisch abgewandelte Wahrnehmung auftritt, anhand einer Skala. Das Ausmaß, mit welchem die Wahrnehmungszustände von den akustischen Eigenschaften des ursprünglichen Reizes abweichen, beschreibt den Grad an Intensität der wahrgenommenen Halluzination, wobei zwischen folgenden Abstufungen zu unterscheiden ist: Zunächst besteht die **normale**, unveränderte akustische Wahrnehmung eines aus der externen Umgebung emittierten Stimulus. Unter **erweiterter** Wahrnehmung versteht sich die Verstärkung bestimmter akustischer Aspekte in der Perzeption eines externen Reizes, wobei häufig eine als angenehm und detailreich beschriebene Hörerfahrung entsteht. Besteht eine drastische, interne akustische Modifikation eines Soundereignisses, spricht man von **Verzerrung**. Mögliche Symptome sind beispielsweise die Entstehung von auditiven Effekten wie Verzerrung, Echo oder Filter. Wie zuvor erläutert beschreibt die höchste Stufe der Skala, die akustische **Halluzination**, die vollständig interne Entstehung akustischer Reize in Form von Musik, Geräuschen, Stimmen oder Stille (Weinel & Cunningham, 2017, S. 2).

In ihrer Forschung demonstrieren die Autoren die Implementierung von drei Effekten aus dem Spektrum der akustisch veränderten Wahrnehmung in eine Spielumgebung, mit dem Ziel, eine Grundlage zur adaptiven Simulation alternativer Bewusstseinszustände von Charakteren in Videospiele zu schaffen.

Selektive auditive Aufmerksamkeit: Dieser Vorgang definiert die gesteigerte Aufmerksamkeit bzw. Vereinnahmung des Hörers gegenüber einer spezifischen akustischen Entität, wobei nicht beachtete Soundquellen eine Abschwächung in der wahrgenommenen Lautstärke erfahren. In der Implementierung wird die Ausrichtung der Spielerkamera mit der Fokusverlagerung des Spielers und dessen Avatar gleichgesetzt. Mit zunehmender Dauer der Betrachtung eines Soundobjekts steigt der Wert für Aufmerksamkeit, welcher parallel die Amplitude der fokussierten Quelle erhöht und die aller übrigen Schallquellen mittels einer Hüllkurve abschwächen lässt. Dieser subjektive Aufmerksamkeitseffekt geht bei Abwendung der Kamera vom jeweiligen Objekt wieder in den Normalzustand über.

Erweiterte Wahrnehmung: Ein erweitertes Hörerlebnis kann insbesondere für musikalische Reize beobachtet werden und sich in Form von veränderter Intensität, Klangfarbe oder spektralen Eigenschaften der Soundquelle äußern. Dieses Phänomen wird in seiner emotionalen Wahrnehmung als genießbar beschrieben, dient aber auch als Vorbote von Bösem und kann Angst oder Paranoia hervorrufen. Im Kontext interaktiver Medien arbeitet diese Methode mit einem Erweiterungs-Wert, bei dessen Änderung ein Crossfade-Wechsel zwischen den Sound-Loop-Variationen der Stufen Dumpf, Mittel und Hell aktiviert wird. Die Filterung sowie Verstärkung bestimmter Frequenzen innerhalb der Erweiterungs-Ebenen simulieren eine verstärkte oder gedämpfte auditive Wahrnehmung. Eine Umsetzung in Videospiele zur Akzentuierung musikalischer Eigenschaften bedarf aufgrund der Verdreifachung der benötigten Sampleversionen einer Speicheroptimierung, idealerweise einer Generierung in Echtzeit.

Spatiale Unterbrechung: Veränderte Bewusstseinszustände können unter anderem Störungen in der räumlichen Perzeption von Schall verursachen. Ein Signal, dessen räumliche akustische Eigenschaften in Kontrast mit Merkmalen einer Umgebung stehen, führt aufgrund der Desorientierung des Hörers zu Überwältigung und Verwirrung. In Videospiele erfolgt z. B. die Modulierung der Distanz durch die Oszillation des Audios auf einem Animations-Pfad um die feste Position der ursprünglichen Klangquelle herum. Insbesondere im Horror-Genre stellen spatiale Unterbrechungen durch die Bewegung der Quelle auf komplexen, nicht-periodischen Oszillationsstrukturen sowie der Zusatz von Echo- und Halleffekten einen Mechanismus zur

Imitation eines psychotischen Zustandes oder einer verzerrten Größenwahrnehmung der Spielumgebung dar (Weinel & Cunningham, 2017, S. 2–4).

In Hinblick auf zukünftige Entwicklungen stellen sich die Autoren, mittels vom Entwickler vorgefertigter, einstellbarer Parameter, eine Adaption grafischer sowie akustischer Elemente an die Avatar-Zustände vor, um das emotionale Befinden des Charakters zu verdeutlichen. Beispielsweise könnte ein Input-Parameter steuern, ob die Reizwahrnehmung der Hauptfigur intern oder extern erfolgt. Die situativen Zustände des Charakters könnten durch die emotionalen Parameter Valenz und Erregung beschrieben werden, um darauf basierend die Simulation veränderter Bewusstseinszustände direkt zu steuern. Die vorgeschlagene avatarbasierte Simulation bietet im Vergleich zur bewegungsbasierten Soundgenerierung eine hohe Anpassungsfähigkeit an verschiedene Spielsituationen als auch eine detailliertere Repräsentation psychischer Zustände. In Zukunft könnte die Präzision der dargelegten Mechanismen mithilfe von Technologien wie Eye-Tracking oder Biofeedback optimiert werden (Weinel & Cunningham, 2017, S. 5–6). Da die Autoren die Frage nach der genauen Funktionsweise sowie Implementierung der steuernden Parameter wie dem Erweiterungs-Wert oder dem Valenz-Erregungs-Modell offenlassen, verlangt die adaptive, auditive Darstellung von Bewusstseinszuständen in Videospielen ausführlichere Forschung.

Besonders im Genre psychologischer Horror-Spiele bewirkt die auditive Simulation veränderter Bewusstseinszustände beklemmende Gefühle beim Spieler. Über Audio-Halluzinationen hinaus, besteht ein weiteres psychologisches Phänomen, dessen Transfer auf die auditive Ebene interaktiver Anwendungen ein vielversprechendes Gestaltungsmittel beschreibt.

4.6 Audio Uncanny Valley

Der Robotiker Mori (2012) stellte 1970 die Hypothese des Uncanny Valley (Unheimliches Tal) auf, welches die Menschlichkeit eines Objekts und die Affinität des Betrachters zu dieser Entität in Relation zueinander setzt. Die Visualisierung dieses Verhältnisses weist, wie in der folgenden Abbildung zu erkennen, an einem spezifischen Bereich auf der Menschlichkeits-Achse einen starken Einbruch der Affinitäts-Kurve auf – das sogenannte Uncanny Valley (S. 98).

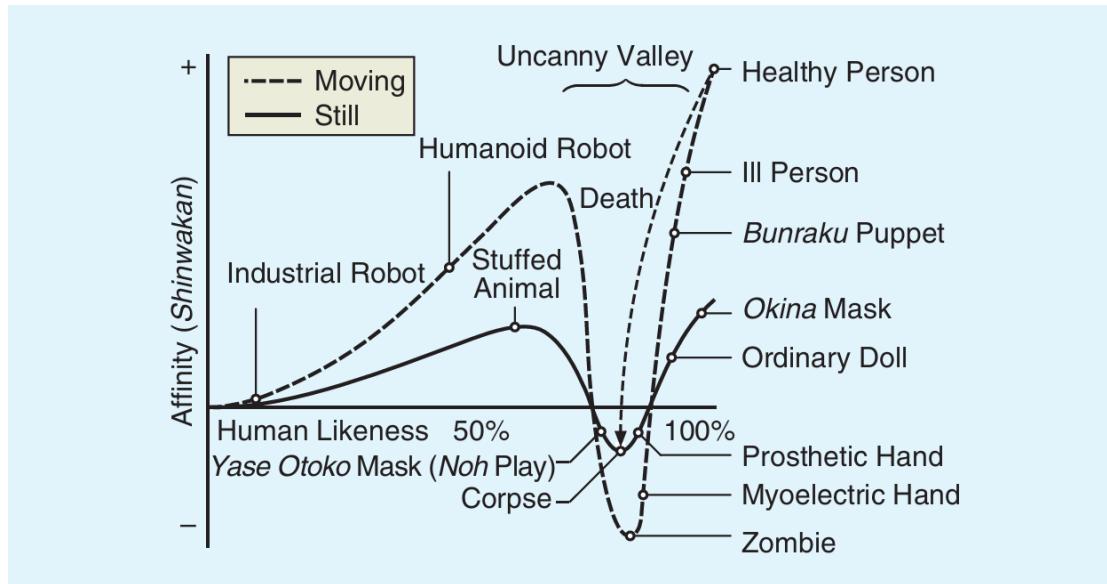


Abbildung 2: Kurve des Uncanny Valley

Quelle: Mori, 2012, S. 99

Der Autor beschreibt dieses Phänomen als unheimliches Gefühl, welches im Zuge von mangelnder Affinität während der visuellen Perzeption einer nichtmenschlichen Erscheinung mit einem bestimmten Grad an optischen humanoiden Anteilen entsteht. Dieser Effekt ist darin zu begründen, dass die im Tal des Graphen verorteten Gestalten Tod und Gefahr symbolisieren. Mittels Bewegungen und Gestiken, die aufgrund zeitlicher Abweichung von menschlicher Motorik meist dehumanisiert wirken, kann die Affinität reduziert und die wahrgenommene Unheimlichkeit verstärkt werden (Mori, 2012, S. 99–100).

Die Analyse der genauen psychologischen Zusammenhänge dieses Effekts sowie sämtliche, durch das Uncanny Valley hervorgerufene, emotionale Reaktionen liegen außerhalb des Rahmens dieser Arbeit. Daher bezieht sich der folgende Abschnitt auf mögliche Verfahren in der Audioproduktion von Videospielen, welche das Prinzip des Uncanny Valley als Ansatz zur Erzeugung diverser negativer Emotionen heranziehen.

Grundsätzlich führt die Verfremdung eines ursprünglich als vertraut aufgefassten Audios zur Erzeugung einer unheimlichen Perzeption (Grimshaw, 2009, S. 25). In Horror-Medien lässt sich

dieser Effekt anhand der Verfremdung ikonischer Klänge, wie beispielsweise dem Tonhöhen-Layering eines Telefonklingelns, beobachten. Der vertraute Stimulus eines klassischen Telefonklingelns bleibt erkennbar, erhält jedoch durch auditive Verzerrung auf den zusätzlichen Ebenen eine befremdliche, unheimliche Note, welche den Soundeffekt in einen Horror-Rahmen einbettet. Der Prozess einer gewohnten auditiven Reizwahrnehmung, die im zeitlichen Verlauf eine Entfremdung erfährt, kann ebenso innerhalb des Genres, mit dessen eigenen charakteristischen Elementen stattfinden (Ball, 2006, S. 44).

Die akustische Gestaltung trägt des Weiteren in Wechselwirkung mit der visuellen Ebene eines Videospiels zum Uncanny-Valley-Effekt bei. Dieser wird durch eine niedrige Auflösung des Audios im Vergleich zur bildlichen Auflösung, eine überspitzte Artikulation der Lippen, parallel zu vokalen Reizen sowie einen Mangel an Synchronisation von Lippen zu Stimme bei fotorealistischen Spielcharakteren erzielt (Grimshaw, 2009, S. 25). Tinwell und Grimshaw (2009, S. 13) weisen außerdem auf die Korrelation zwischen unmenschlichen Aspekten der Stimme eines Charakters und dessen wahrgenommener Unheimlichkeit sowie Seltsamkeit hin. Als stimmlich unmenschliche Attribute können eine monotone, langsame Sprache, eine unpassende Intonation sowie eine allgemein unstimmige Wahrnehmung der Stimme im Vergleich zur visuellen Erscheinung des Charakters aufgeführt werden. Je stärker diese Aspekte in einer Stimme vorhanden sind, umso intensiver wird ein Charakter als fremd und merkwürdig wahrgenommen.

Die Übertragung des Uncanny Valley auf das Sounddesign in Horror-Spielen ist nicht auf eine spezifische Form der auditiven Verfremdung beschränkt, jedoch sind bestimmte Faktoren zu berücksichtigen, um die gezielte Induktion eines unheimlichen, befremdlichen Gefühls zu gewährleisten. Zunächst sollten sich wiederholende Samples vermieden werden, da sie dem Rezipienten mit der Zeit vertrauter werden, wodurch die intendierte Entfremdungswirkung aufgehoben wird. Ebenso entscheidend darüber, ob mittels der angegebenen Methoden eine unheimliche Wirkung erzielt wird, ist der Rahmen, in welchem sie Anwendung finden. Audiovisuelle Elemente des Uncanny Valley können außerhalb eines Horror-Kontextes beispielsweise zur Belustigung beitragen. Daher ist in Videospielen ein klarer, vom Spieler akzeptierter Horror-Rahmen nötig (Grimshaw, 2009, S. 24).

Wie bereits aufgezeigt, erhält sich die Spielerimmersion durch die Synchronität zwischen den SFX und den entsprechenden visuellen Stimuli. Insofern stellt sich für die Anwendung des Uncanny Valley in Videospielen die Frage, wieso auditive Verfremdung, die oftmals mit einer Diskrepanz zwischen Bild und Ton verbunden ist, keinen Immersionsbruch verursacht. Am Beispiel eines First-Person Survival-Horror-Games, argumentiert Grimshaw (2009), dass anhand der visuellen Präsenz-

Hypothese sowie der Rahmen-Hypothese ein Immersionsbruch gehemmt und der Angsteffekt verstärkt wird. Erstere beschreibt das Hineinversetzen des Spielers in die Realität der Spielwelt. Die Rahmen-Hypothese legt ein klares Horror-Setting fest und gewährleistet folglich ein Bewusstsein des Rezipienten über die Absicht des Spiels, Angst zu erzeugen. Durch die kontextuelle Erwartungshaltung des Rezipienten werden unheimliche und unbehagliche Gefühlsregungen als Bestandteil der Spielerfahrung antizipiert (S. 24).

Trotz der Annahme, dass im Laufe der Zeit durch wiederholte unheimliche Aspekte Vertrautheit geschaffen wird, beschreiben Tinnwall und Grimshaw (2009) eine völlige Überwindung des Uncanny Valley als unmöglich. Dies sei vor allem darin begründet, dass sich im Zuge der stetig wachsenden Komplexität neuer Technologien ein anspruchsvollereres Urteilsvermögen der Nutzer entwickle, was in einer steigenden Empfindlichkeit für subtile Unstimmigkeiten resultiere (S. 23).

Die dargelegten Erkenntnisse bestätigen, dass die Erzeugung eines unheimlichen Effektes durch Verfremdung auch auf der akustischen Ebene realisierbar ist. Dieser Prozess tritt in Medien selten komplett unabhängig von visuellen Reizen auf, weshalb das Uncanny Valley in Videospielen als ein audiovisuelles Phänomen betrachtet werden muss.

4.7 Monster und der Sound der Einsamkeit

Je nach Genre und Medium verbergen sich unterschiedliche Intentionen hinter dem Einsatz von Monstern. Im Horror beschreibt die Erscheinung von Monstern klassischerweise eine Abnormalität, welche mit der natürlichen Struktur und den gewohnten existenziellen Regeln der positiven Figuren einer fiktiven Welt interveniert. Der ausschlaggebende Unterschied zu anderen Genres besteht im Horror in der negativen Reaktion der Charaktere auf abnormale Gestalten und Situationen (Caroll, 1990, S. 16). Diese stellt sich dabei nicht nur durch die Wahrnehmung einer Bedrohung und demzufolge ein Gefühl von Angst dar, sondern ist insbesondere von Emotionen wie Abscheu, Ekel und Übelkeit gezeichnet. Horrorbasierte Ungeheuer werden nicht nur durch Tödlichkeit, sondern besonders durch ihre abstoßende Wirkung charakterisiert, die häufig anhand von Merkmalen wie Schmutz, Verfall und Schleim symbolisiert wird (Caroll, 1990, S. 22). Die Spiegelung der Charakteremotionen auf den Rezipienten stellt einen fundamentalen Bestandteil des Horrors dar. Die Reaktionen der Figuren werden dabei, aufgrund des Bewusstseins über die fiktive Gegebenheit, nicht exakt auf die des Zuschauers repliziert, sondern angenähert (Caroll, 1990, S. 18). Im Rahmen von Videospielen besteht jedoch eine grundlegende Überschneidung der Spieler- und Charakterreaktion auf Monster, da der Rezipient aktive Einflussnahme auf das Verhalten des Avatars besitzt. Monster agieren sowohl auf der mechanischen als auch auf der psychologischen Ebene als

ein Hindernis und eine Bedrohung, welche das Ziel verfolgt, den Spieler zu eliminieren. Ihre Erscheinung ist häufig durch die Transgression zwischen menschlich und unmenschlich, lebendig und tot sowie hybride maschinelle, tierische oder übernatürliche Merkmale gekennzeichnet.

Sinclair (2020) unterteilt die auditive Expression von Kreaturen in primäre und sekundäre Sounds. Demzufolge bilden morphologische Eigenschaften eine Grundlage, um darauf basierend die sogenannte Hauptstimme bzw. die primären Sounds eines Lebewesens adäquat zum Ausdruck zu bringen (S. 137). Die akustisch erwirkte, beängstigende Rezeption von Kreaturen basiert zum einen darauf, dass der Spieler anhand von Eigenschaften der Primärsounds auf die Art und anatomische Beschaffenheit der Bedrohung schließt. Jedoch besteht der entscheidende Faktor vielmehr in der Assoziation des Stimulus mit erlernten Symboliken des Horror-Genres sowie den damit verbundenen negativen Gefühlen (Roux-Girard, 2011, S. 203). Ergänzend zur Hauptstimme bilden periphere Geräusche auf der sekundären Ebene wie etwa Atmung, Schlucken oder Hautgeräusche bei Bewegung, einen wichtigen Bestandteil in der Immersionsvertiefung und ermöglichen dem Spieler eine Konkretisierung über die Art der Bedrohung. Sekundäre Sounds sollten in ihrem Pegel bewusst hinter Primärgeräuschen platziert werden, um die intendierte Charakterisierung des Monsters nicht zu gefährden. Wie genau die verstörende, abstoßende Wirkung eines Monsters mittels Bild und Ton umgesetzt wird, variiert mit narrativen Aspekten sowie der vom Spiel intendierten psychologischen Wahrnehmung und Reaktion des Konsumenten (Sinclair, 2020, S. 137).

Spezifisch für psychologische Horrorspiele ist die metaphorische Darstellung der Psyche des Hauptcharakters durch Monster, wie am Beispiel des Babys in *Resident Evil: Village* (Capcom, 2021) belegt werden kann. In dieser Situation steht der Hauptcharakter Ethan, der nach dem Tod seiner Frau auf der Suche nach seinem entführten Baby ist, temporär unter dem Einfluss einer psychotropen Substanz, was ihn halluzinieren lässt. Die Szenerie beinhaltet symbolische visuelle Elemente wie eine lebensgroße Puppe, fallende Bilder einer Schwangeren sowie ein Krankenzimmer und die scheinbar aus einem Radio emittierende Stimme seiner Frau. Diese Aspekte versinnbildlichen das Schuldgefühl Ethans über den Verlust seiner Frau und Tochter, sowie die Mutation des entführten Babys, von der er nichts weiß (Jones, 2021). Die auditive Halluzination äußert sich zunächst in Form eines generischen, zweidimensionalen Babygeschreis (Timodus, 2021, 0:18). Kurz darauf begegnet der Spieler einer Halluzination eines riesigen, deformierten Babys, welches aus Fleisch und Knochen zu bestehen scheint. Insbesondere das primäre Audio des Monsters verstärkt, durch Verzerrung und tonhöhenmoduliertes Layering von stöhnenden Lauten, verstörendem Gekicher und Schreien, die Dehumanisierung der Erscheinung. Diese resultiert in Kontrast mit vereinzelten Rufen nach dem Vater in einer emotionalen Ambivalenz (Timodus, 2021,

1:30; Jones, 2021). Die anthropomorphe Physiologie des Babys wird durch nasse, fleischige Sounds und quietschendes Leder für die Haut sowie durch tieffrequente, nasse Stampf- und Schleifgeräusche unterstrichen (Timodus, 2021, 5:03). Die audiovisuelle Darstellung von Ethans Halluzination offenbart die unterbewusste Paranoia und Angst des Charakters und verweist den Spieler zudem durch unterschwellige Hinweise auf die Hintergründe der Handlung.

Silent Hill 2 spielt auf eine ähnliche Weise mit der Verdeutlichung der Hintergrundgeschichte und psychosexuellen Problemen des Hauptcharakters durch das Sounddesign verschiedener Monster. (Timodus, 2021). Diese wird gesondert im Analyse-Teil über das Remake behandelt. Im Allgemeinen kann die Erscheinung der Monster der Silent Hill-Reihe als eine Modalität des Uncanny Valley betrachtet werden, da diese eine Synthese aus menschlichen Eigenschaften mit Elementen wie Verwesung, Blut, verdeckten und entstellten Köpfen oder insektenähnlichen Komponenten aufweist. Durch den abrupten Wechsel von zweibeinigem Gang zu unnatürlichen Krabbel- oder Huschbewegungen sowie die deplatziert wirkende Verwendung mechanischer Geräusche für Wesen mit einer humanen Erscheinung, setzt das Werk den psychologischen Fokus auf den Ausdruck physischer Kontamination. Die betonte Vokalisation der Monster wie etwa Aufstoßen, Fauchen sowie nasse Bewegungs- und Impact-Geräusche tragen neben der Induktion eines Ekelgefühls zur Verstärkung grafisch nicht klar wahrnehmbarer Details, wie bestimmter Oberflächentexturen, bei. Das Verschwimmen der Grenze zwischen menschlichen und unmenschlichen Eigenschaften, mithilfe des Zusammenspiels von Bild und Ton, stellt in diesem Fall die Bedrohung von Verunreinigung in den Vordergrund, was in Gefühlen von Ekel und Furcht resultiert (Ekman & Lankoski, 2009, 189–190).

Ein alternativer Effekt kann im Falle der Monster in *Fatal Frame* beobachtet werden, welche sich als menschliche Leichen in Geister-Erscheinung mit unnatürlicher Körperhaltung manifestieren. Der emotionale Fokus liegt hier auf dem Leiden und Schmerz in der Vorgeschichte der Geister, welche dem Spieler sowohl durch visuelle Rückblicke als auch durch primitive, sich wiederholende Stöhnen- und Klageläute vermittelt werden. Der groteske Aspekt des Werkes setzt sich weniger aus Ekel, und mehr aus mentaler Kontamination, der Angst vor spirituellem Verfall und Wahnsinn, zusammen (Ekman & Lankoski, 2009, S. 188–190). In Verbindung mit plötzlichen Abbrüchen und Einsätzen von Klängen bei der Begegnung mit feindlichen Entitäten, zielt das Sounddesign der Monster in *Fatal Frame* primär auf die Gestaltung der subjektiven Charaktererfahrung der Bedrohung ab (Ekman & Lankoski, 2009, S. 192).

Eine bislang nicht weit verbreitete, jedoch in ihrer Wirkung als äußerst effektiv betrachtete Methode im Sounddesign unheimlicher Kreaturen, findet sich in der Mehrdeutigkeit der Körpergeräusche des gesteuerten Charakters und der Monster. Dabei weisen etwa die Schrittgeräusche des Spielenden auf einer bestimmten Oberfläche eine beachtliche Ähnlichkeit zu denen des Monsters auf. Die Bewegung des Charakters resultiert somit in einem verstärkten Zweifel des Rezipienten über die Echtheit des akustischen Reizes, sowie in einer gesteigerten Vorsicht und Anspannung, ohne dass die Voraussetzung für eine Gefahr gegeben ist (Roux-Girard, 2011, S. 206). Insbesondere im Rahmen des Survival-Horrors, welcher sich durch von Anthropomorphismus geprägte Welten und die Einsamkeit des Hauptcharakters auszeichnet, führen Soundeffekte, deren Erscheinung im Widerspruch zu einer leblosen Welt steht, zur Verstärkung einer paranoiden Stimmung. Geräusche wie Rascheln, Kratzen oder Schlurfen, die keiner Ursache zugeordnet werden können, suggerieren dem Spieler eine Bedrohung in einer scheinbar unbelebten Umgebung. Schrittgeräusche, welche zeitgleich mit denen des Hauptcharakters ausgelöst werden und leicht versetzt abklingen, veranlassen den Spieler dazu, sich ohne sichtbare Gefahr beobachtet und verfolgt zu fühlen (Ekman & Lankoski, 2009, S. 191–192).

4.8 Musik

Wie zu Beginn klargestellt, findet sich eine zentrale Bedeutung musikalischer Signale in der Wahrnehmung und im Erleben individueller sowie gemeinschaftlicher Emotionen. Während die Rezeption von Musik als einzelner Stimulus als frei und subjektiv bezeichnet werden kann, verfolgt ihre Implementierung in Videospielen konkrete Ziele innerhalb der Bereiche Information, Immersion und Unterhaltung. Gesondert auf das Genre des Horrors bezogen, beinhalten diese grundsätzlich die Induktion einer unheimlichen Stimmung, das Informieren in Form akustischer Hinweise sowie die Manipulation des Spielers mittels mehrdeutiger oder unzuverlässiger Klänge. Bei Erfüllen der Ziele ist ein emotional immersives, verstörendes Spielerlebnis gewährleistet (Roberts, 2014, S. 138). Die musikalische Gestaltung sowie deren Intention variieren mit den individuellen Gegebenheiten eines Videospiels und treten auf verschiedenen diegetischen Ebenen wie auch in Zwischenzuständen auf. Die folgenden Erläuterungen ermitteln daher spezifische, funktionelle Formen von Musik, die aufgrund ihrer psychologischen Wirksamkeit im Rahmen von Horrorspielen üblicherweise Anwendung finden.

4.8.1 Musik als Warnsystem und Leitmotiv

Der Einsatz von Musik in konfrontativen Situationen ist neben seiner emotional-metaphorischen Funktion als narratives, extradiegetisches Warnsystem definiert. Werden entsprechend spannungsbildende musikalische Sequenzen beim Erscheinen feindlicher Kreaturen ausgelöst, bezweckt dies eine Erregung negativer Gefühle wie Panik und Angst gegenüber der Begegnung beim Rezipienten. Umgekehrt werden diese Reaktionen mit dem Abklingen der Musik beschwichtigt, um dem Spieler ein Gefühl von vorübergehender Sicherheit und Abwesenheit der Gefahr zu vermitteln (Perron, 2018, S. 302). Der klangliche Verweis auf einen Zustand von Gefahr gestaltet sich typischerweise durch den Einsatz intensiver rhythmischer Elemente sowie Moll-Tonarten oder verminderter Akkorde. Die stetige Aufrechterhaltung einer solchen musikalischen Kulisse bewirkt ein starkes Vorantreiben der Spielerhandlung, was sich einerseits in einer stetigen Anspannung des Rezipienten äußert, da diesem die akustische Definitionsbildung eines sicheren Zustandes, wie etwa durch Stille oder ruhige Musik, verwehrt wird. Andererseits übt solch ein Handlungs-Appell einen motivierenden Effekt auf seine Hörer aus (Whalen, 2004; Roberts, 2014, S. 138).

Musik kann außerdem als Fusion von Warnsignal und Leitmotiv für eine bestimmte Art von Bedrohung in Erscheinung treten. Die klanglichen Hinweise fungieren dabei als Ankündigung und Bestimmung spezifischer Monster außerhalb des Sichtbereiches (Roberts, 2014, S. 145). Dieses Stilmittel lässt sich an den sogenannten Killern in *Dead by Daylight* (Behaviour Interactive, 2016) veranschaulichen: Jeder der Charaktere zeichnet sich durch ein eigenes musikalisches Thema aus, welches in Abhängigkeit von deren Erscheinungsbild und Vorgeschichte bestimmte Elemente enthält. Beispielsweise werden die Eigenschaften der Hexe durch tribale Elemente wie hölzerne Trommeln, Perkussion und Flöten verdeutlicht, das Leitmotiv des Clowns enthält hingegen Instrumente wie eine Marschtrommel, Akkordeon, Xylophon sowie Fagott. Die von einer dissonanten, verzerrten sowie meist treibenden Instrumentalisierung geprägten Themen setzen bei Verfolgung eines Spielers ein, wobei in Korrelation mit verringriger Distanz eine erhöhte Anzahl und Intensität der Instrumente sowie eine Verdopplung des Tempos erzeugt wird (Dead by Daylight Wiki, o. D.). Grundsätzlich sollten solch klangliche Leitmotive vor dem ambienten Soundtrack positioniert werden, um dem Hörer ein klar wahrnehmbares Signal über die sich nähernde Gefahr zu vermitteln und so gegebenenfalls eine effektive Kommunikation mehrerer Spieler untereinander zu ermöglichen. In anderen Szenarien kann der gezielte Einsatz von räumlicher Mehrdeutigkeit durch musikalische Themen den Spieler dazu veranlassen, in Verwirrung nach der Position der Soundquelle zu suchen. Die Ablenkung des Spielers, welcher sich bereits aufgrund weiterer Angriffe in einer Stresssituation befindet, erzeugt ein intensiviertes Anspannungsgefühl (Roberts, 2014, S. 145).

4.8.2 Darstellung von Charakteremotionen und Umgebungen

Eine weitere Aufgabe musikalischer Komponenten in Videospielen besteht in der Verstärkung visueller Anteile einer Szene sowie in der Repräsentation emotionaler Figuren- und Spielerreaktionen. Roberts (2014) beleuchtet musikalische Dynamiken beim Wechsel zwischen verschiedenen Umgebungen und Zuständen am Beispiel von *Silent Hill*. Im ersten Werk der Reihe zeichnet sich der Übergang von weiten, dunklen, atmosphärischen Außenbereichen zu labyrinthartigen, klaustrophobischen Innenräumen in seiner musikalischen Struktur ab. Dabei dominieren in freien Umgebungen ambiente, mechanische Klänge, die sich mit beengenden räumlichen Zuständen zunehmend in dissonanten Klängen, krachenden Geräuschen und pulsierenden Bassuntertönen verdichten. Musik kann so beispielsweise den Spieler in Angst und Konfusion über den weiteren Verlauf des Geschehens rätseln lassen (S. 139).

Huckvale (1990) verweist auf die symbolische Darstellung emotionaler sowie physiologischer Charakterzustände durch Musik im Horror-Kontext. Demzufolge werde die Verwendung von Tremolo mit Zittern oder Nervosität und Staccato mit nervöser Schreckhaftigkeit assoziiert, wobei ein Crescendo eine allgemeine Steigerung dieser Wirkungen herbeiführe (S. 16). Wiederholte, hohe Streichernoten sowie eine Tempobeschleunigung durch perkussive Elemente stellen ein effektives Mittel dar, um das Durchleben von Panikattacken eines Charakters sowie die daraus resultierenden Bewegungsunstabilitäten des Avatars während der Flucht vor einer Bedrohung klanglich zu repräsentieren (Perron, 2018, S. 301). Das Anschwellen negativer, chaotischer Gefühle, wie etwa beim Betreten zunehmend verstörender Umgebungen, kann durch eine Steigerung überlappender musikalischer Sequenzen und Lautstärke sowie durch den Einsatz von Atonalität realisiert werden. Insbesondere in Perspektiven, in denen der Spielende hauptsächlich von hinten auf die Spielfigur blickt, dienen musikalische Hinweise als eines der wenigen Mittel zur Abbildung der Charakteremotionen (Whalen, 2004).

4.8.3 Manipulation: Musik und Stille

Unter Berücksichtigung der Immersionserhaltung arbeiten Sounddesigner häufig mit manipulativen Ansätzen, die entgegen der Erwartung des Nutzers fungieren. Beispielsweise werden in Verfolgungsszenen, welche mit der Erwartung an einen lauten, treibenden Score in Verbindung stehen, leise, dumpfe Musikstücke angewendet. Die somit beim Spieler entstehenden Gefühle von Verunsicherung und Entkörperlichung lassen sich dadurch erklären, dass die visuell bedingte Erwartungshaltung gegenüber dem Ton nicht erfüllt wird (Roberts, 2014, S. 149).

Akira Yamaoka, Komponist und Sounddesigner von *Silent Hill 2*, verfolgt mit seinem Score eine vielschichtige, kontrollierte Wiedergabe und Manipulation psychologischer Aspekte. Im Interview mit ToneArtistLive (2014) verweist er darauf, dass jeweils mehrere Versionen eines Stückes komponiert wurden, welche in Abhängigkeit von den Gefühlszuständen der Charaktere sowie den Aktionen des Nutzers ausgelöst werden (09:17–10:23). Yamaoka bestätigt außerdem die Hypothese, dass ein maßgebliches Ziel von Videospiel-Scores darin besteht, die Spielenden in eine bestimmte Richtung zu lenken. Dies kann sowohl in Form einer positiven Rückmeldung über die ausgeführten Bewegungen und Interaktionen auftreten als auch gewollt auf die Irreführung des Spielers abzielen (ToneArtistLive, 2014, 21:48–22:27). Seinen erwartungswidrigen Ansatz hinsichtlich der Musik verdeutlicht Yamaoka am Beispiel des unvorhersehbar fluktuierenden Verhältnisses zwischen Musik und Stille:

The normal way is to use strong music for an attack scene or a surprise scene, but I normally use that type of music and also make sudden silences, too, or use very soft music. So I try not to make a pattern that the user will get used to. I want to make variations that users cannot anticipate.
(Kohler, 2007)

Der Einsatz von Stille in *Resident Evil 7: Biohazard* (Capcom, 2017) stellt eine vergleichbare Strategie dar. Indem ein vertrautes akustisches Umfeld, bestehend aus Schrittgeräuschen der Hauptfigur und einer unaufdringlichen Ambience, abrupt durch Stille unterbrochen wird, verliert der Hörer seine klanglichen Orientierungshilfen. Der somit erzeugte Zustand von Verunsicherung bildet eine Grundlage, um den Spieler durch ein plötzliches Geräusch in Angst zu versetzen (Resident Evil, 2017a, 00:22). Die semantische Auffassung des Spielers zur Abwesenheit akustischer Stimuli variiert abhängig von ihrem Zeitpunkt und ihrer Einsatzhäufigkeit über den Verlauf des Werkes. In Szenarien, die einen Zustand von Sicherheit vermitteln sollen, kann Stille anstelle der vertrauten, atmosphärischen Klangkulisse stehen, um beispielsweise die Abwesenheit von Gefahr zu signalisieren. Jedoch führt diese Ruhe tendenziell zu einem verstärkten Misstrauen und antizipatorischer Angst beim Hörer. Die schlagartige Intrusion von Stille, meist durch unangekündigte Angriffe, wird genutzt, um den antizipatorisch unheimlichen Effekt zu verstärken oder vom Überraschungsmoment Gebrauch zu machen, um den Spieler in Schreck zu versetzen (Whalen, 2004; Dennis, 2020).

Ein für den Horror charakteristisches Stilmittel zur Manipulation der Spielerwahrnehmung stellt die klangliche Juxtaposition dar. Häufig erscheinen Kinderlieder, welche ein Gefühl von Fröhlichkeit und Sicherheit vermitteln, innerhalb eines extremen, düsteren Kontextes. Durch Techniken wie etwa Verlangsamung und dissonante Verzerrung verfremdet, erzeugt die Anwendung ursprünglich

nostalgischer, positiver Musik in einer konträren virtuellen Umgebung ein Gefühl von Disharmonie und Beklemmung (Hill, o.D.). Innerhalb dieses Stilmittels wird oft durch eine Modifizierung der Instrumentalisierung und Texte die Hintergrundgeschichte eines Charakters aufgegriffen oder ein Bezug zum Thema Kindheit hergestellt (Resident Evil, 2017, 00:29). Die diskrepante Wirkung dieser Methode lässt sich am Beispiel des Einsatzes eines Kinderliedes an verschiedenen Stellen des Sci-Fi Survival-Horror-Spiels *Dead Space* (Electronic Arts, 2023) aufzeigen. Neben den mechanischen, futuristischen Soundeffekten und bedrohlichen, atmosphärischen Klängen der dystopischen Welt erzeugt ein verzerrt und tragisch interpretiertes Wiegenlied eine verstörende Stimmung. Die Bedeutung dieser unpassenden Gegebenheit wird dem Spieler zusätzlich durch das verstärkte Herzklopfen der Hauptfigur übermittelt (VG Moments, 2023).

Die vorangegangenen Ziele von Musik in Videospielen beschreiben nur einen Bruchteil des Umfangs an sachlichen, emotionalen und lediglich unterhaltenden Informationen, die sie vermittelt. Videospiel-Musik ist allgegenwärtig und im Gegensatz zur Gruppe der Soundeffekte nicht an visuelle Vorgaben gebunden. Sie bezieht sich auf spezifische Charaktere, Orte, Zustände und Handlungen sowie die Gefühle des Spielers zu diesen Instanzen. Bedauerlicherweise überschreitet eine ausführlichere Ergründung dieser zahlreichen klanglichen Gestaltungsmöglichkeiten den Rahmen dieser Arbeit.

5. Analyse von Silent Hill 2: Remake

5.1 Handlung

Die Handlung des Spiels beginnt mit James Sunderland, der am Waldrand der Stadt Silent Hill angekommen ist. In einem inneren Monolog erzählt er, dass er seine Frau Mary drei Jahre zuvor an eine tödliche Krankheit verloren habe. James habe vor kurzem jedoch einen Brief erhalten, dessen Handschrift, Name auf dem Umschlag und Text darauf hindeuten, von Mary geschrieben zu sein. Dieser enthält die Nachricht, sie würde an dem besonderen Ort der beiden – in Silent Hill – auf ihn warten. In Ungläubigkeit und Hoffnung, dass seine Frau doch am Leben ist, macht sich James auf in die verlassene Stadt.

Silent Hill ist stark vernebelt und wirkt komplett verlassen. James trifft jedoch auf verschiedene menschliche, entstellte Kreaturen, die ihn unaufhörlich bedrohen. Während er ums Überleben kämpft, enthüllt er durch das Erkunden verschiedener Orte und das Lösen von Rätseln die düstere Wahrheit seiner eigenen Geschichte und die der Stadt. Über den Verlauf des Spiels begegnet der Protagonist verschiedenen Personen wie Angela, die auf der Suche nach ihrer Mutter ist und Eddie, welcher mit Aggressionen und Essattacken zu kämpfen hat. Außerdem lernt er das Mädchen Laura kennen, welches im Krankenhaus mit Mary befreundet war und wegen eines Briefes ebenfalls nach ihr sucht. Schließlich begegnet James Maria, die wie eine Kopie von Mary aussieht, jedoch in einem knapperen Outfit und mit stark flirtendem Verhalten auftritt. Sie versichert James, dass sie nicht seine Mary sei. Von diesem Zeitpunkt an folgt sie James über den Großteil der Handlung.

Schließlich kommt es zu einem Wendepunkt, als James durch ein Video-Tape erfährt, dass er Mary damals mit einem Kissen erstickt hat. Dieser Vorfall liegt außerdem nicht drei Jahre, sondern drei Tage zurück. Je nach Ende, das der Spieler erreicht, variieren James' Gründe dafür von Gnade bis zu Hass.

Im Spielverlauf wird graduell klar, dass Silent Hill eine subjektive Hölle für die einzelnen Charaktere simuliert. Die Monster ebenso wie Maria sind Manifestationen von James' Unterbewusstsein und verkörpern auf unterschiedliche Weise seine Schuldgefühle, sexuelles Verlangen und den Wunsch nach Bestrafung dafür, seine Frau umgebracht zu haben. Die psychischen Probleme von Eddie und Angela können ebenfalls teilweise in metaphorischer Form von James wahrgenommen werden. Beispielsweise manifestieren sich Angelas Trauma, Schuldgefühle und Selbstmordgedanken aufgrund der Misshandlung durch ihren Vater in Form eigener Monster und Halluzinationen. Einige Theorien besagen, dass die Erschaffung der Monster von Silent Hill verursacht wird, andere führen sie auf Manifestationen von James' Psyche zurück. In jedem Fall besteht der Kernpunkt des Spiels in der Simulation einer halluzinationsähnlichen Erfahrung und dem Auflösen der Grenzen zwischen

Realität und Einbildung, wodurch die psychologischen Probleme von James für den Rezipienten allgegenwärtig sind.

Es ist außerdem zu berücksichtigen, dass die visuelle sowie auditive Gestaltung der Welt zwischen einem Normalzustand und der sogenannten Otherworld wechselt. Diese alternative Welt wird als bräunlich, zerfallen und verrottet visualisiert. Es ist dabei nicht belegt, ob es sich um eine weitere Halluzination handelt oder ob dieser Zustand die echte Welt darstellt und sich James die harmonischere Normalwelt nur einbildet.

Schlussendlich findet James Maria in Gestalt von Mary auf, die daran scheitert, ihn zu täuschen. Mary will James überreden, Maria hinter sich zu lassen und mit ihr zu sein. James lehnt dies ab, wodurch er gezwungen wird, gegen Mary in ihrer Monster-Form zu kämpfen und sie schließlich zu töten. In Abhängigkeit von den summierten Spielerhandlungen führt dies zu verschiedenen Enden, die von James' Selbstmord über ein glückliches Ende, in dem er mit Laura die Stadt verlässt, bis hin zu obskurer Comedy reichen. In seiner Essenz zusammengefasst, treten folgende Erkenntnisse über die psychologischen Hintergründe von James zum Vorschein: Er wollte nicht mehr mitansehen, wie seine Frau an der Krankheit leidet, kam nicht mit ihren Stimmungsschwankungen klar, fühlte Frustration über die fehlende Sexualität der Beziehung und war Mary schlicht und einfach leid. James' psychisch krankes Verhalten und die daraus resultierenden Gedanken und Weltwahrnehmung, bringen den Spieler dazu, die Perspektive einer moralisch verwerflich handelnden Person einzunehmen, die sich nach und nach an ihre Geschichte erinnert und sich ihrer Fehler bewusst wird. Das Werk setzt sich damit nicht nur kritisch mit mentalen Krankheiten auseinander, sondern simuliert diese dem Spieler, mittels audiovisueller Symbolik, so roh und nah wie möglich.

5.2 Sounddesign

5.2.1 Ambience (Video 1)

Die Ambience von Silent Hill ist in drei Ebenen zu gliedern. Die Grundlage wird dabei aus einer ruhigen, stetigen Ebene von Umgebungsrauschen und tieffrequentem Grummeln sowie dezenten Umgebungsgeräuschen geformt. Bemerkenswert sind dabei die Variation und das Level an Realismus der texturierten Soundeffekte, welche besonders durch das Flacker-Surren von Lampen oder die Adaption von Regengeräuschen auf unterschiedlichen Materialien bemerkbar sind und zur Immersion der Umgebungen beitragen (00:00). Bereits nachdem die ersten Feinde vorgestellt werden und die Stadt somit als gefährlich verstanden werden kann, bildet sich die zweite Schicht, bestehend aus hochfrequenten Tönen im Crescendo, geisterhaften, atonalen Stutter-Effekten und Sägegeräuschen sowie metallisch quietschenden Geräuschen, die vergleichbar mit dem Knarzen eines Schiffskörpers sind. Die akustischen Sequenzen der zweiten Ebene schwanken in zufälligen Abständen und Kombinationen an und ab (00:27). Die nächste Ebene wird bei der Begegnung mit Feinden oder rein zufällig, jedoch deutlich seltener als Ebene zwei, ausgelöst. Sie stellt eine rhythmische Klangschicht mit hoher Lautstärke dar, die stilistisch an die grundlegende Ambience angepasst ist. Beispielsweise hetzt ein rhythmisches Ticken den Spieler bei Begegnung mit Feinden durch die Spielwelt, wohingegen in Situationen ohne sichtbare Feinde ein anschwellendes, metallisches Grollen mit basslastigem Impactsound für Verunsicherung und einen Schreckmoment sorgt (01:33). Nach dem Töten der Kreaturen treten ebenfalls atonale, rasselnde und rhythmische Versionen dieser Audiosequenzen auf, um die Nachwirkung und Folgenschwere der Aktion zu betonen. Zum einen dient dies der langsamen Abschwellung der Spannung, was eine verbesserte Erhaltung der Immersion erzielt. Zum anderen wird somit eine unterschwellige Fatalität und Dunkelheit von James' Aktionen und denen des Spielers vermittelt, wodurch ein gewisses Mitleid mit den eigentlich bedrohlichen Kreaturen entsteht.

Je nach Zeitpunkt des Einsatzes kann es dazu führen, dass diese Sequenzen mit einer Situation kongruieren, was in einer unvorhersehbaren und somit furchteinflößenden Spielererfahrung resultiert (02:00). Die Atmungs-SFX in der Otherworld des Krankenhauses werden in der dritten Ebene im Pegel stark erhöht, in ihren hohen Frequenzen gedämpft und sind allgemein deutlicher als Atmung zu erkennen. Das zusätzliche Herzklopfen schafft neben seiner unangenehm nahen Präsenz zum Hörer Angst und Verwirrung, da es sowohl James als auch der Quelle der Atmung zugeschrieben werden kann und auf ein bedrohliches Ereignis hinweist, welches jedoch nie stattfindet (02:12). Da sich die Atemgeräusche der ersten Ebene mit der Zeit für den Spieler als eindeutig extradiegetischer Bestandteil etablieren, wird die initiale Verunsicherung über den Ursprung der Soundquelle in den sporadisch ausgelösten Sequenzen der dritten Ebene aufgefrischt.

Während die Atmung in der ersten Ebene mechanischer und rhythmischer ist und mit dem Rest der Ambience verschmilzt, tritt sie hier in den transdiegetischen Bereich ein und ruft erneut Unsicherheit hervor.

Die Realisierung der ambienten Szenen weist in dem behandelten Werk ebenfalls auf veränderte Umweltzustände hin. Innerhalb kürzester Zeit folgt beispielsweise auf eine stark gedimmte auditive Atmosphäre ein Angriff, begleitet von tiefen sowie hohen Alarm-Sirenen, starken Windgeräuschen und metallischem Klappern, welche nach Betreten eines Gebäudes komplett abklingen. Die Klanggestaltung verdeutlicht in Situationen wie diesen Zustände von Gefahr und Sicherheit, wobei diese teils umschlagen, bevor eine Veränderung sichtbar ist, was den Rezipienten bereits vor einem visuellen Hinweis in Schock und Handlungsangst versetzt (02:35).

Ein zentrales, ambientes Stilmittel in *Silent Hill 2: Remake* ist die raumspezifische Ambience, welche die visuellen Elemente und die symbolische Bedeutung einer jeden Umgebung individuell verdeutlicht. Häufig werden Wind- und überlagerte, synthetische Atemgeräusche verwendet, um verlassene, dunkle Räumlichkeiten darzustellen (03:24). Bass-Töne und Grummeln wiederum deuten auf Verwesung hin (03:48). Die Veränderung durch Verrottung, Schmerz und Gewalt in der Otherworld wird dabei z. B. durch dissonante, geläutete Töne und das Knacken eines kaputten Tonaufnahmegeräts verdeutlicht (04:09). Musik stellt sich als eher seltener Anteil der ambienten Kulisse heraus und wird in instrumentaler Form und nostalgischer Lofi-Ästhetik realisiert, wobei sie gleichzeitig als Bestandteil des Set Dressings, wie etwa in Form eines Radios, als transdiegetisches Element auf den Aspekt der Vergangenheit hinweist (04:34). Soundeffekte, die auf die ursprüngliche Geräuschkulisse des nun ausgestorbenen Raums hinweisen, bilden einen Übergang von Ambience zu einer symbolischen Audio-Halluzination. Je nach Auffassung erzählen sie somit eine Geschichte über einen Ort oder bringen die Gedanken des Charakters darüber zum Ausdruck (04:42). Ähnliche symbolische Audio-Elemente finden sich in fortschrittbasierter, auslösbarer Soundeffekten, die gleichzeitig Teil der extradiegetischen Atmosphäre sind. Ein Beispiel hierfür ist das Uhrenrätsel, welches mit jedem hinzugefügten Zeiger eine andere klangliche Reaktion hervorruft, von lautem Ticken und Zeigerdrehen bis hin zu Kuckucks-Uhr-Effekten (05:00). Die sich vom Rest der Ambience abhebenden, sehr lauten Uhrensequenzen unterbrechen die angenehme tonale Stimmung, weisen aber gleichzeitig auf den Fortschritt des Spielers hin. Diese trickreiche Methode sorgt daher nicht nur für Motivation beim Rezipienten, sondern ebenso für Misstrauen, Angst und Unsicherheit über die Korrektheit der Handlung, selbst in einer Erfolgssituation.

Wie zuvor am Beispiel der Atemgeräusche im Krankenhaus festgestellt, bezweckt die Ambience eine Personifizierung der scheinbar verlassenen Stadt, indem diegetische Geräusche in der

atmosphärischen Gestaltung aufgegriffen werden und der jeweiligen Umgebung so ein Gefühl von Eigenleben einhauchen. Zu Beginn des Spiels tritt dies in Form von metallischem Knarzen verschiedener großer Bauten und an den verschlossenen Zugängen der Stadt auf – untermauert von wummernden Windgeräuschen oder einem tiefen, sehr langsam pulsierenden Ton (05:35). Der metallische Körperschall geht graduell in die nicht-diegetische Ebene über, wo er in zahlreichen Momenten als auditive Grundlage der Ambience auftritt (05:49). Des Weiteren verläuft dieses Element später in ein klares, extradiegetisches Alarmgeräusch, welches im Gegensatz zu den diegetischen Warnungen deutlich langsamer pulsiert und raumspezifisch ist. Diese Alarmton-Atmosphäre findet sich hauptsächlich in Räumlichkeiten mit einem Gefängnis-Thema wieder, wie etwa in der Zelle im Woodside-Apartment oder bei James' Abstieg in das Toluca-Gefängnis (06:02). Besonders im Falle des Letzteren gestaltet sich das Sounddesign durch seinen räumlichen Hall und zeitliche sowie lautstärkemäßige Änderungen des Signals als unregelmäßiger in seinem Verlauf und kann als transdiegetisch wahrgenommen werden, wodurch der Spieler darüber verwirrt ist, welche Relevanz der Warnung zuzuordnen ist. Da der transdiegetische Alarm meist hervortritt, wenn die Hauptfigur allein ist, können diese Töne als Warnung der Stadt vor James selbst und dessen Darstellung als Eindringling und Verbrecher interpretiert werden. Die ambienten Alarmklänge bilden nicht nur die Spielwelt, sondern besonders das Innenleben des Hauptcharakters, als einen konstanten Ausnahme- und Gefahrenzustand ab. Es wird dem Spieler eine stetig präsente Gefahr – James selbst – vermittelt.

Die hohen und tiefen Alarmtöne und Sirenen sowie die dumpfen, entmenschlichten Atemgeräusche gelten als besonders häufige Merkmale der Silent Hill-Ambience. Sie unterstreichen die Dystopie der Welt und hauchen gleichermaßen toten Orten auf eine unangenehm menschliche Art Leben ein, wo keines sein sollte. Der Spieler wird durch den Wechsel zwischen einem anhaltenden Alarmzustand und der Suggestion einer geisterhaften Präsenz einem permanenten Gefühl von unterbewusster Anspannung und Bedrückung ausgesetzt.

5.2.2 Warnsystem (Video 2)

Das einzigartige Warnsystem der Silent Hill-Serie wird durch ein Taschenradio realisiert, welches James stetig bei sich trägt. In Gegenwart von Feinden sendet es ein schrilles, statisches Rauschen, ähnlich einem Radio-Tuning-Soundeffekt, aus. Das Signal gestaltet sich in seinem Pegel sehr nah und deutlich und in seiner Räumlichkeit als 2D-Sound, der sich in der Extradiegesse zu befinden scheint (00:00). Daher kann das Radio als internes transdiegetisches Element bezeichnet werden. Es wird als diegetisches Objekt initiiert, über das sich Spieler und Figur bewusst sind, geht dann allerdings in die extradiegetische Ebene über, um den Fokus auf die Informationsvermittlung an den Spieler zu legen. Der vorrangige Nutzen dieses informativen Systems besteht zudem in seiner psychologischen Wirkung wie der Verunsicherung und Verängstigung des Hörers, was durch einige Maßnahmen in der Spielmechanik hervorgerufen wird. Zum einen ist der Radio-Sound nicht linear mit dem Abstand der Feinde und variiert zufällig. Er kann beispielsweise anspringen, wenn sich der Feind noch relativ weit in der Ferne befindet, und dem Spieler dadurch das Gefühl geben, die Bedrohung übersehen zu haben (00:32). Insbesondere da sich Silent Hills Kreaturen häufig im Nebel, der Dunkelheit oder unter Autos verstecken, erweist sich das Warnsystem des Spiels als sehr gekonnt darin, anhaltende Paranoia über Distanz und Position der Feinde im Hörer zu triggern. In anderen Fällen wird die Irreführung weiter intensiviert, indem Räume, die keine Feinde beinhalten, eine Ambience emittieren, die dem Radio-Signal gleicht (01:04).

Außerdem kann basierend auf der Akustik des Systems keine Aussage über die Art der Kreaturen getroffen werden. Der einzige Unterschied zwischen den Warnungen vor Kreaturen besteht darin, dass sich die sogenannten Mannequins in einer Starre verstecken und das Radio dabei versagt, sie zu ermitteln (01:37). Es kommt also zu einem temporären Ausfall des Systems, auf den der Spieler durch keinerlei Hinweise schließen kann. In einigen Fällen lösen schwer sichtbare, jedoch relativ ungefährliche Kakerlaken-Monster den Sound aus, was den Spieler bis zum Bemerken der konkreten Bedrohung vorsichtiger werden lässt (01:45). Das auditive Warnsystem von *Silent Hill 2: Remake* erleichtert die Erkennung von Bedrohungen, auf welche die visuelle Ebene keinerlei Hinweise gibt. Dennoch stellen sich die Warnsignale als irregulär und unzuverlässig heraus, was die Präsenz oder Abwesenheit, Art sowie Distanz der Feinde betrifft. Die zu Beginn etablierten Regeln des Prinzips werden vom Nutzer angenommen und bilden einen Rahmen von Sicherheit, bis dieser aufgrund der bewussten Aussparung oder Manipulation von Hinweisen unvorbereitet auf Feinde stößt oder sich seine Vorhersagen als falsch erweisen. Die einzige Möglichkeit, die dem Spieler somit bleibt, liegt in einem enorm wachsamen Vorgehen sowie dem Erleben eines stetigen Anspannungsgefühls.

Über das Radio hinaus äußern sich die Warnsignale des Spiels beim Auslösen diegetischer Komponenten wie der Betätigung einer Uhr, eines Lichtschalters oder dem Öffnen eines Türensystems. Die folgenschwere Bedeutung dieser Vorgänge wird mittels eindringlicher Alarmtöne realisiert, wobei die Geräuschkulisse oft rasant von eintönigem Hintergrundrauschen auf eine laute Ambience mit klingelnden, dröhnen Alarmen wechselt (02:01). Während der Spieler zu diesem Zeitpunkt häufig noch mitten in einer Zwischensequenz oder Aktion eines Rätsels hängt, wobei er in seiner Sicht sowie Bewegung für ein paar Sekunden stark eingeschränkt ist, wird durch akustische Sounds die schnell auf ihn zukommende Gefahr angekündigt. Dies äußert sich meist durch die charakteristischen Ausrufe von Monstern und deren Körpergeräusche beim Kollidieren und Durchbrechen der Umgebung. Die frühzeitige auditive Alarmierung dient zum einen dem informativen Zweck, dem Spieler durch die Positionsbestimmung der feindlichen Wesen strategischen Handlungsfreiraum zu geben. Allerdings dient sie, vor allem in Kombination mit der temporären Handlungsunfähigkeit, als ein effektives Mittel zur Induktion einer Panikreaktion, da es dem Spieler nicht möglich ist, sich im Raum neu zu orientieren und augenblicklich auf die Bedrohung zu reagieren.

Ein auditiv reichhaltiges Beispiel hierfür ist die sogenannte Ruined Area im Labyrinth. Der Spieler ist gezwungen, mehrere Räume hintereinander zu passieren, einen Angriff der Feinde darin auszulösen und diesen zu überleben, bis sich die Tür des nächsten Flurs öffnet. Bei Berühren der geschlossenen Tür wird schlagartig eine Kakophonie ausgelöst, bestehend aus mehreren überlagerten Alarmtönen, einem Hit-Soundeffekt und Geklapper, welches an die Geräusche der Mannequin erinnert. Da sich diese jedoch zu diesem Zeitpunkt nicht im Sichtfeld des Spielers befinden, lässt sich vermuten, dass das Klappern dazu dient, im Hörer Furcht vor einem übersehnen Mannequin auszulösen, da dies keine Seltenheit innerhalb des Gameplays darstellt. Außerdem ist gegen Ende des Signals ein im Tempo erhöhtes Uhrenticken zu hören, bis ein metallisch hallendes Geräusch das Öffnen der Tür ankündigt (03:12). Die visuelle Gestaltung der Räume in Form von roter Beleuchtung trägt zur furchteinflößenden, angespannten Stimmung des Soundscapes bei. Es wird allerdings berücksichtigt, dem Hörer anhand des Tickens und Türengeräusches einen absehbaren Zeitraum sowie das erfolgreiche Überstehen der Konfrontation und die räumliche Position der Ausgangstür zu vermitteln, um dessen Orientierungssinn und Motivation zu bewahren.

5.2.3 Monster (Video 3)

Die Entitäten von Silent Hill repräsentieren einen signifikanten Anteil der psychologischen Bedeutung des Spiels und können in zwei Kategorien eingeteilt werden: Kreaturen, die in vermehrter Zahl erscheinen und als einfache Monster zu bezeichnen sind, und spezifisch wichtige Charaktere, die meist Boss-Monster darstellen und nur an bestimmten Stellen des Spiels aufzufinden sind. Zunächst erfolgt die Betrachtung der einfachen Monster aus der auditiven Perspektive. Diese besitzen visuell eine Vielzahl menschlicher Eigenschaften, die durch die Ergänzung bestimmter audiovisueller Aspekte einen Grad an Anthropomorphismus erlangen.

Beispielsweise ist die Lying Figure eine menschliche Gestalt mit Absatzschuhen und einer fleischigen, bis über den Kopf gezogenen, Zwangsjacke. Die Gestalt bewegt sich meist schnell kriechend vorwärts und erhebt sich gelegentlich, um in einem langsam, zombieartigen Gang zu gehen. Lederige Körpergeräusche sowie krächzende, aufstoßende, teils raubtierähnliche Laute verdeutlichen diese visuellen Eigenschaften. Die beabsichtigte Wirkung dieser Kreatur liegt in der Induktion von Ekel, indem die Lying Figure zur Abwehr erbricht, was auf der Ton-Ebene durch ein Speiherausch und zischende Säure realisiert wird (00:00). Dem Spieler wird somit unmissverständlich klargemacht, dass es sich hierbei nicht nur um etwas Ekelerregendes, Krankhaftes handelt, sondern auch um etwas chemisch Gefährliches, das vermieden werden sollte.

Die Mannequins sind aus zwei unteren Körperhälften zusammengesetzte, puppenartige Gestalten in Menschengröße. Wie bereits betont, regen sie das Warnsystem nicht an und sind häufig in der Umgebung versteckt, wodurch der Spielende meist unvorbereitet mit ihnen konfrontiert wird. Vereinzelt bauen die Entwickler akustische Vorwarnungen ein, die jedoch aufgrund vorangegangener ähnlicher Ereignisse nicht automatisch mit der Kreaturen-Art assoziiert werden und demzufolge zu einem Schock-Moment und Vorsicht beim Nutzer führen (00:21). Die Körpergeräusche der Mannequins äußern sich in klappernden, zwitschernden und insektenartigen Surrgeräuschen, und obwohl die Monster keinen Kopf besitzen, stoßen sie vor allem als Kampfreaktion kindliches Lachen und Stöhnen aus, welches teilweise sehr sexualisiert wird (00:41). Im ersten Beispielvideo (00:21) fällt besonders der Sound einer Kolbenflöte auf, der zusammen mit spielzeugartigem Rasseln bei Erscheinung des Mannequin auftritt. Die Juxtaposition von auditiven kindlichen Merkmalen im Horror-Kontext wird hier vertieft in einen verstörenden Kontext gerückt, indem Eigenschaften von Unschuld und Sex zugleich durch dasselbe Monster verdeutlicht werden. Die clownartige Vertonung kann zusätzlich als Verhöhnung des Spielers oder von James' Unterbewusstsein an sich selbst gedeutet werden.

Ab dem Toluca-Gefängnis begegnet James einer auf allen Vieren und über Wände krabbelnden Version der Mannequins, welche in ihrer morphologischen Gestaltung an eine Fusion aus Spinne

und Pferd erinnert. Das Sounddesign als eine Mischung von Pfeifen, Klappern und Insektengeräuschen bleibt bestehen, die menschliche Vokalisation wird jedoch durch Fiepen und Quielen ersetzt (00:58). Das Gefängnis bietet nur begrenzte Sicht und Helligkeit, wodurch der Spieler besonders auf sein Hören zurückgreifen muss. Da jedoch enorm viele Feinde in der Umgebung herumwandeln und die Mannequins sich mit rapidem Tempo sowie von oben auf den Spieler zubewegen, ist dieser dazu gezwungen, ständig in Bewegung zu bleiben und dabei durch das Gemisch von atmosphärischen Kreatur-Sounds und dreidimensionalen Bewegungsgeräuschen einzuschätzen, wo sich die Gefahr im Raum befindet.

In Momenten, in denen der Spieler sich nicht für Konfrontation, sondern für Flucht entscheidet, spielt besonders die Richtungsabhängigkeit der Kreatur-Sounds eine Rolle in der raumbasierten Induktion von Angst. Insbesondere im Falle einer beschränkten Sicht, wenn sich der Spieler auf das Entkommen konzentriert, können die Geräusche der Feinde hinter dem Charakter in ihrer Position und Lautstärke deutlich und über einen längeren Zeitraum wahrgenommen werden. Ohne die Gefahr zu sehen, wird durch diesen übersteigerten Effekt die gefährliche Nähe und schnelle Verfolgung der Monster vermittelt (01:06).

Eine weitere herumwandelnde Bedrohung sind die Bubble Head Nurses: Krankenschwestern mit einem verbundenen Gesicht, knappen Outfit und hohen Schuhen. Deren zuckende Motorik und zombiehafter Gang stellt sich auf der primären Sound-Ebene durch gutturale, nasse, ächzende Grunts dar und menschliche, schmerzvolle Leidensschreie, wenn die Kreaturen von James besiegt werden (01:19). Die Nurses verdeutlichen demnach ein weitgehend humanes Monster, das in James ein Schuld- und Schamgefühl über sein sexuelles Verlangen auslöst sowie seine Ängste vor Krankheit und Dahinsiechen verdeutlicht. Das Audio dient der Kontrastbildung zwischen Sexualität und Krankheit wie auch dem Leiden, welches in James' Geschichte damit einhergeht.

Als Mandarins werden Monster einer weiblichen Erscheinung mit fleischigen Mündern anstelle von Gesicht und Armen bezeichnet. Sie bewegen sich an ihren Armen hängend unter Bodengittern entlang und greifen James mit ihren langen Zungen, Mündern und ohrenbetäubenden Schreien an. Da sich die Mandarins meist zügig und im Dunkeln unter dem Spieler bewegen, werden besonders die schweren Schrittgeräusche der Kreatur hervorgehoben und bezwecken, im Gegensatz zum Sounddesign anderer Monster, eine direktionale eindeutige Wahrnehmung, um dem Hörer einen genauen Eindruck über die Position der Gefahr zu ermöglichen (02:19). Die Mandarins senden außerdem einen brüllenden Schrei aus, der James für kurze Zeit in seiner Bewegung stoppen lässt, da er eine temporäre Hörschwellenabwanderung erleidet, die durch einen starken High-Cut-Filter und schrille, dissonante Pfeiftöne simuliert wird (02:29).

Dieser Effekt wird nur ausgelöst, wenn sich die Kreatur in nächster Nähe befindet, und ist in der Regel gefolgt von einem Angriff. Dieser wird durch zwei SFX angekündigt, die aus einem tiefen, raubtierartigen Grollen und nassen Körpergeräuschen bestehen. Da der Angriff, wie aus der Abbildung zu entnehmen ist, direkt nach den Impact-Geräuschen erfolgt, ist es dem Spieler also möglich, lediglich aus auditiven Signalen ein Muster zu erkennen und dementsprechend zu reagieren (02:45).

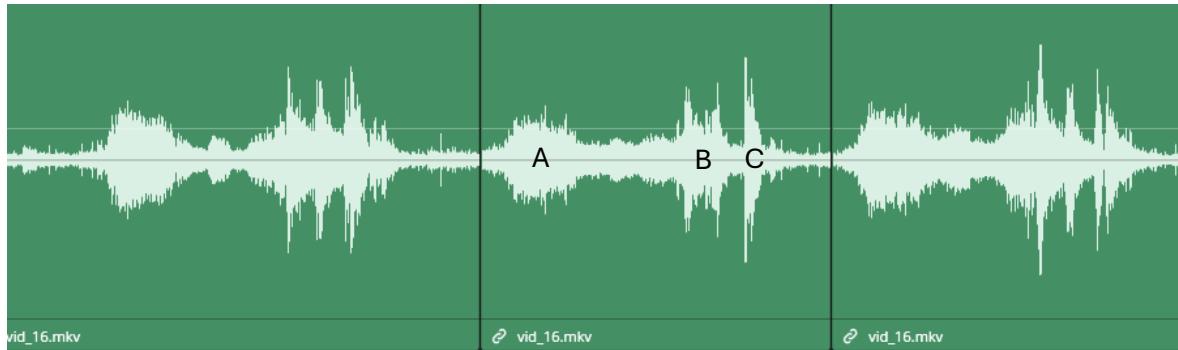


Abbildung 3: Gegenüberstellung der Audio-Signale bei drei Angriffen des Mandarins
 A: Brüllen des Monsters/ Tinnitus-Effekt, B: Zweifache Vorbereitung/ Körpergeräusche, C: Angriff
 Quelle: Eigene Darstellung

Obwohl der Fokus im Sounddesign der Mandarins primär auf informativen Aspekten liegt, stellt vor allem die tierische Vokalisation eine befremdliche Komponente dar. Das präsente Audio steht einer starken visuellen Einschränkung gegenüber, wodurch simultan mit der Angst vor dem Unbekannten gespielt und ein Gefühl von stetiger Verfolgung induziert wird.

Pyramid Head ist das essenzielle Boss-Monster des Werkes: ein muskulöser Mann mit einem pyramidenförmigen Metallhelm, der seinen Kopf komplett überdeckt. Er trägt eine Metzgerschürze und schleift ein riesiges Messer hinter sich her. Pyramid Head soll sowohl Kontrolle und Schuld als auch Bestrafung symbolisieren, was sich durch sein Misshandeln und Töten der übrigen Monster sowie in seinen Begegnungen mit James manifestiert. Im Zuge der auditiven Analyse der Figur wird eine starke Abgrenzung zu den einfachen Monstern deutlich. Pyramid Head erscheint häufig nicht einfach so, sondern kündigt sich über mehrere Szenen hinweg mit dumpfem Grollen und Schritten, sowie scharfen, metallischen Schleifgeräuschen an. Je nachdem, wie weit entfernt Pyramid Head vom Spieler wahrgenommen werden soll, wird das Signal beispielsweise stark gedämpft und einige Zeit später verschärft wiedergegeben, um die herannahende Gefahr zum Ausdruck zu bringen. Obwohl die Kreatur noch nicht anwesend ist, offenbaren sich die Schritt- und Schleifgeräusche sehr klar und laut, so dass der Rezipient die Richtigkeit seiner Wahrnehmung in antizipatorischer Angst hinterfragen muss (03:00). Zu anderen Zeitpunkten passiert sein Erscheinen schlagartig und still und wird durch tierisches leises Knurren, Röhren sowie kurze Sirenen töne untermauert (3:51). Sein allgemeines Sounddesign kann als atonaler Score bezeichnet werden, der auf basslastigen Schlägen

und metallischen Elementen, die im Rhythmus Ähnlichkeit zu Maschinengeräuschen aufweisen, beruht. Als zentrales Objekt bewegt sich der Sound seines Messers sowohl als ambiente, extradiegetische und unangenehm hochfrequente Töne innerhalb der Klangkulisse als auch in Form von schrillen, schleifenden Foley-Sounds in der Diegese. Die Präsenz des Monsters ist dabei fast immer von Sirenen untermauert, die in Tonhöhe und zeitlichem Verlauf variieren (04:01). Zusammenfassend lässt sich sagen, dass dieser Charakter durch seine Vertonung mit einem bedrückenden und unangenehmen Alarmzustand verbunden wird. Der Kontrast zwischen tieffrequenten, dunklen Impulsen und schrillen Sirenen- sowie Kratz-Tönen verdeutlicht seine Rolle als mechanische, jagende und womöglich gefährlichste Bedrohung im ganzen Spiel.

5.2.4 Auditive Halluzinationen und Jumpscares (Video 4)

Das wesentliche Thema in *Silent Hill 2: Remake* besteht in der Auseinandersetzung mit der psychologischen Befindlichkeit des Hauptcharakters. Da die Ursache und Beantwortung der Existenzfrage der audiovisuellen Halluzinationen auf unterschiedlichen Interpretationen beruhen, ist es schwer zu sagen, wieso und wann genau James halluziniert, welche Aspekte der Normalität angehören oder physikalisch real, also diegetisch sind. Jedoch verleiht genau diese Gegebenheit dem Sounddesign des Videospiels seine so faszinierende und unheimliche Atmosphäre. Da ein Großteil der Klänge in der Spielwelt von Wahnfiguren emittiert wird, zählen strenggenommen alle akustischen Anteile der Monster und teilweise die der Ambience zu Halluzinationen. Aus Gründen der Übersichtlichkeit geht dieser Abschnitt gesondert auf besonders prägnante Halluzinationseffekte ein, welche die Autorin folgendermaßen definiert: Sound-Events, die sich durch ihr seltenes Vorkommen klar von der Ambience abgrenzen, keinem sichtbaren Emitter zugeordnet werden können oder aufgrund ihres Grades an Surrealismus eindeutig als auditives Stilmittel einer Horror-Halluzination eingestuft werden können.

Bereits in den ersten Minuten des Spiels begegnet der Rezipient Ereignissen mit ungeklärter Ursache. Zunächst wird ein Geräusch rollender Steine an einem Steinhang ausgelöst, wobei die Bewegung der Steine sichtbar ist, lediglich die genaue Ursache bleibt ungeklärt und kann daher zunächst als zufälliges Naturereignis abgetan werden. Kurz darauf wird ein tief gepitchedes, leises Raubtierknurren ausgelöst, wodurch in Verbindung mit dem vorherigen Vorfall sofort die Vermutung entsteht, von einer Kreatur verfolgt zu werden (00:00). Im weiteren Verlauf begegnet der Spieler beim Passieren bestimmter Stellen Geräuschen wie dem Knarzen von Bäumen oder Gerumpel und Knarzen in Gebäuden. Häufig folgt auf die akustischen Ankündigungen ein Ereignis wie zum Beispiel das Einstürzen einer Treppe, wodurch die Geräusche sich von einer vermuteten Halluzination zu einem begründeten, jedoch gefährlichen Ereignis wandeln (00:22).

Nach und nach werden den erläuterten Objektgeräuschen menschliche Komponenten hinzugefügt, wie etwa hochfrequente Stimuli, die sich zwischen einem Schrei und Geräusch verorten lassen. Ab diesem Zeitpunkt beginnt James das erste Mal aktiv auf die Geräusche zu reagieren, was auf Spielerseite die Empathie zum Charakter verstärkt und die Behauptung bestätigt, dass die Signale zunehmend menschliche Eigenschaften annehmen (00:43). Vor allem in Umgebungen, die weniger Monster beinhalten, wird somit Unsicherheit und Anspannung induziert, da die Geräusche in der unmittelbaren Nähe und in Verbindung mit einer bestimmten Ursache vermutet werden. Ein im Spiel verstreut vorkommender Halluzinations-Sound, welcher sich schlussendlich als außerordentlich bedeutungsvoller Bestandteil der Narrative herausstellt (siehe Kapitel 5.2.7), ist das Nach-Luft-Schnappen einer Frauenstimme. Das Geräusch wird an bestimmten, jedoch nicht weiter relevanten Orten emittiert und gestaltet sich in manchen Situationen sehr leise oder versinkt leicht in ambienten Klängen (00:54). Zunächst führt dies dazu, dass der Spieler sich nicht sicher über seine eigene Hörwahrnehmung ist, was ihm einen halluzinationsähnlichen Zustand simuliert. Mit der Zeit etabliert und verdeutlicht sich das Signal jedoch innerhalb beängstigender Situationen und beginnt in einem handlungsrelevanten Kontext zu erscheinen.

Im Laufe des Spiels setzt sich die Zunahme von menschlichen Eigenschaften, der Unvorhersehbarkeit sowie der Zusammenhangslosigkeit der Halluzinationen weiter fort. Dem Spieler wird deutlich der symbolisch und physisch erfahrene Abstieg in die Dunkelheit von Silent Hill und James' Psyche vermittelt. Dabei wird der Spieler auf zynische Weise mit der Vergangenheit der Hauptfigur konfrontiert, wie beispielsweise im Aufzug der Otherworld des Krankenhauses. Hier wird ein Radioquiz ausgelöst, in dem ein Moderator vor einer hörbaren Zuschauermenge James als Guest der Show vorstellt und ihm mit Bestrafung bei falschen Antworten droht. Die Verfremdung der motivierten Moderations-Stimme durch einen Lofi-Radio-Effekt und eine Pitch-Shifting-Doppelung geben der Sprache des Moderators einen unheimlichen, dämonischen Beigeschmack (01:00). Das Quiz dient dem Lösen eines optionalen Rätsels, hat also keinerlei drastische Folgen. Die satirische, unheimliche Zurschaustellung und Beschuldigung von James erzeugen allerdings das eindeutige Gefühl, dass diese Situation surreal und etwas nicht in Ordnung ist.

Diese Art auditive Halluzination erfährt eine Vertiefung der kryptischen Aspekte. Diese offenbaren sich als wörtlich unverständliche, kurze Sprachsequenzen wie etwa ein laut wahrnehmbares Flüstern mit hervorgehobenen S-Lauten oder eine tief gepitchte Männerstimme mit einem großen Raumhall (01:35). Eine mögliche Erklärung für die Verwendung der Stimmen ist die Darstellung der ehemaligen Bewohner der Räume – die tiefe Vokalisation in der Gefängniszelle weist aufgrund der

Kult-Symbole an der Wand auf die geisterhafte Präsenz eines Mitglieds einer ominösen Glaubensgemeinschaft hin.

Im dunklen Hof des Gefängnisses befinden sich lediglich Rätselstationen, die von einer weiten, leeren Fläche umgeben sind. Die minimalistische Ambience wird in unregelmäßigen Abständen durch das Geräusch eines galoppierenden Pferdes unterbrochen. Es variiert in seiner Position und Bewegung im 3D-Raum, wobei der Eindruck eines am Spieler vorbeitrappelnden Tieres entsteht. Diese Methode bringt den Hörer dazu, in der Dunkelheit vergeblich nach der Quelle zu suchen und beim Lösen der Rätsel durch die undefinierbare Ablenkung leicht die Fassung zu verlieren. Hier sticht besonders die Absurdität der Art des Geräusches hervor, die keine tiefere Bedeutung aufweist und lediglich dem Auslösen von Verwirrung und einem unbehaglichen Gefühl dient. Außerdem weisen die Hufgeräusche aufgrund der matschigen Bodentextur Ähnlichkeit zu denen des Charakters auf, womit dem Nutzer klar kommuniziert wird, dass sich die Soundquelle im selben Raum befindet. Das Getrappel springt oft mit dem Vorwärtsbewegen von James an und klingt ebenso parallel zu dessen Schrittgeräuschen ab, wodurch das paranoide Gefühl verstärkt wird, von einer Entität verfolgt zu werden, die sich bewusst im Verborgenen halten möchte (01:53).

Gegen Ende des Spiels entwickeln sich audiovisuelle Halluzinationen, die an die Inszenierung eines Spukhauses erinnern. Ein eindrucksvolles Beispiel hierfür stellt der Korridor dar, welcher jeweils auf die drei Durchgänge der Alarmsystem-Räume aus Video 2 folgt und sich dabei mit jeder Phase in seiner Repräsentation intensiviert. Auf der Audio-Ebene äußert sich dies in Form eines metallischen Schleifgeräusches, welches an das Messer von Pyramid Head erinnert, das bereits etablierte Nach-Luft-Schappen einer Frauenstimme, gefolgt von einer Stille-Pause, auf die ein dissonanter, Impact-Sound, ähnlich eines Alarmtons folgt, wobei rot flackerndes Licht emittiert wird. Der Nachhall in der Gestalt von geisterhaftem Rauschen und einem schrillen Crescendo wird abrupt beendet (02:14). Die zweiten und dritten Versionen des Korridors sind bereits zu Beginn in einen roten Schein getaucht, wodurch eine Verbindung zum vorher erfahrenen Zustand hergestellt wird. Das Sounddesign äußert sich dieses Mal in Form eines metallischen Schleifens und eines Schlägeräusches, woraufhin ein Erdbeben ausgelöst wird, wodurch das Material der Wände zu Boden stürzt. Unterstützt wird dieses Bild durch eine Masse an Rüttel- und Bröckelgeräuschen sowie einen Sound, ähnlich eines tieffrequenten Schiffshorns (02:27).

Während metallische Geräusche für den Nutzer bereits als bedrohliche und symbolische Sound-Elemente etabliert sind, begegnet dieser der auditiven Halluzination der einatmenden Frauenstimme zum ersten Mal in Kombination mit zusätzlichen Jump-Scare-Geräuschen. Dadurch, dass auf ein bekanntes Geräusch, auf welches keine Konsequenz erwartet wird, ein Element folgt, welches der Spieler mit einer konkreten Bedrohung verbindet, spielen die Entwickler geschickt mit

Gewohnheitsdynamiken. Der Rezipient kann sich nicht mehr auf die erlernten Bedeutungen bestimmter Signale verlassen und ist dazu gezwungen, diese ständig zu hinterfragen und neu zu definieren, was zu einer fortlaufend erneuerten Bildung antizipatorischer Angst führt. Auf die emotionale Ebene von James' Charakter bezogen, beschreibt die Halluzination das Zusammenbrechen einer Welt, einer stabilen Verfassung und ein steigend bedrohliches, dunkles Ausmaß seiner Wahnzustände oder weist anhand der auditiv gewaltigen und maschinellen Elementen auf eine größere Macht hin, die James' verstörende Erfahrung lenkt. Ab dem Punkt, an dem die Halluzinationen ihren Zusammenhang verlieren und an schizophrene Erscheinungen erinnern, entsteht eine intensivierte unsichere Spielererfahrung, die mit einem Gefühl von Wehrlosigkeit einhergeht.

Im Lakeview-Hotel lenken die auditiven Halluzinationen das Thema auf Trauer (02:46). Das Schreien von Mary nach James erinnert an die mit ihr verknüpften Gefühle von Verlust und der Hoffnung darauf, dass sie noch am Leben ist (02:56). An diesem Ort treten ebenfalls vermehrt Jumpscares auf, die durch laute Geräusche von aufschwingenden Türen, herabfallenden Gegenständen oder einfallenden Wänden realisiert werden (03:02). Die Jumpscares intendieren hauptsächlich die Versetzung des Spielers in einen Schreckmoment, was in diesem Abschnitt des Spiels besonders mithilfe von vermehrten Stille-Phasen hervorgehoben wird. Das fortlaufende Thema einer zusammenbrechenden Welt dient außerdem als symbolisches Gestaltungsmittel, welches mit Fortschreiten in der Handlung an Bedeutung zunimmt und durch die gewaltigen, tieffrequenten Geräusche in seiner Wahrnehmung verstärkt wird.

5.2.5 Sonstige Soundeffekte (Video 5)

Während die vorangegangenen Kapitel spezifische Sounddesign-Methoden des psychologischen Horror-Genres beleuchteten, widmet sich dieser Abschnitt der Ergänzung der im Spiel häufig erscheinenden Motive, bei denen die auditive Gestaltung eine maßgebliche Rolle für die emotionale Auffassung der Rezipienten spielt. *Silent Hill 2: Remake* beschäftigt sich auf einer emotional metaphorischen Ebene stark mit dem mentalen Abstieg des Hauptcharakters in die dunkelsten Bereiche seiner persönlichen Geschichte. Während James wortwörtlich immer tiefer in die Abgründe der Stadt und seiner Gedanken springt, sind bei Näherung zum Abgrund ein starkes Rauschen und Grummeln, schrille als auch tiefe metallische, industrielle Klänge sowie James' panisches, verstörtes Atmen zu hören, wodurch dem Spieler ein Gefühl für die Absurdität und enorme Tiefe des Falls verdeutlicht wird. Der Prozess des Abstiegs wird von einer dissonanten, gelayerten, quäckenden Ambience begleitet, die zwischen menschlichen Klageschreien und einer an- und abschwellenden Sirene liegt (00:00). Die nasalen, alarmierenden, klagenden Klänge

erzeugen so ein klaustrophobisches Uncanny-Gefühl, welches durch die industriellen Soundkomponenten eine Verschärfung des Anthropomorphismus erfährt. Ein weiteres Horror-Thema, von welchem das Werk zahlreiche Male Gebrauch macht, ist das Motiv des Ekelns. Ekel spielt außerhalb der audiovisuellen Monster-Gestaltung für das Lösen von Rätseln eine Rolle, wobei James wiederholt dazu gezwungen ist, in ekelrege Substanzen zu greifen. In Situationen, in denen die Konsistenz einer Substanz gar nicht oder nur begrenzt sichtbar ist, führen matschige, nasse Schleim- und Blubbergeräusche sowie Säurezischen die Texturdetails eines gefährlichen oder abstoßenden Objektes direkt an den Betrachter heran (00:55). Die Angst des Unbekannten, in Zusammenspiel mit derartigen Geräuschen, unterstützt die von Ekel und Überwindung geprägte Charakterreaktion und induziert infolgedessen Angst sowie ein verstärktes Empathie- und Ekelempfinden des Rezipienten.

Das Werk beinhaltet zudem SFX, die zwar nicht spezifisch dem Horror-Genre zuzuordnen sind, jedoch zur Qualität der Spielererfahrung beitragen. Die stimmliche Gestaltung des Hauptcharakters bildet einen maßgeblichen Bestandteil der emotionalen Immersion. Beispielsweise zeigt der Protagonist nach einem feindlichen Angriff verstärkte und zitternde Atemgeräusche auf sowie realistisch anmutende Grunts in Situationen körperlicher Anstrengung (01:30). James reagiert in Momenten der Interaktion, wie etwa mit einer verschlossenen Tür, und äußert Verwunderung oder Entsetzen angesichts außergewöhnlicher Begebenheiten in Form von kurzen Aussagen an sich selbst (01:45). Diese vokale Rückmeldung haucht dem Charakter Leben ein und dient als internes transdiegetisches Kommunikationssignal an den Spieler, um diesem die Gedanken seiner Spielfigur zu vermitteln. Dies führt zu einer in sich schlüssigen, immersiven Spielerfahrung und verstärkt vor allem in stressreichen Situationen die Verbindung zwischen Spieler und Charakter. Die auf Oberfläche und Geschwindigkeit basierten Schrittgeräusche sowie die Kleidungsgeräusche von James zeichnen sich durch einen hohen Grad an Variabilität und Realismus aus, was die Bildung und Erhaltung der Immersion auf eine subtile Weise gewährleistet (01:55).

Wie im vorangegangenen Kapitel über Ambience dargelegt, zeichnet sich das Voranschreiten und Lösen innerhalb eines Rätsels durch akustische Feedbacks aus, welche selten helle, positiv konnotierte Signaltöne beinhalten. Die auditive Rückmeldung manifestiert sich meist als am Objekt orientiertes, realitätsnahes und subtiles Signal oder eine spannungsbildende, atmosphärische Klangsequenz (02:09). Der Verzicht auf ein affirmierendes Ton-Schema bei Erfolgserlebnissen intendiert die Induktion von Gefühlen wie Verunsicherung und Wachsamkeit beim Rezipienten. Insbesondere Sound-Emitter wie Uhren unterbrechen unmittelbar das Erfolgserlebnis, z. B. durch zügige Zeigergeräusche und laute Uhrentongongs. Das Auffinden eines Rätselobjektes wird hierbei von einem drängenden, sich im Pegel steigernden Uhrenticken sowie einer dunklen, atonalen

Klangkulisse begleitet. Die auditive Zuspitzung bereitet den Nutzer auf einen bevorstehenden Boss-Kampf vor, welcher das Erfüllen der Aufgabe zu einem nur kurz währenden Gefühl von Hoffnung und Erfolg wandelt (02:25). Die unheilvolle Stimmung in den als sicher und positiv aufgefassten Momenten bildet einen konstanten emotionalen Spannungszustand. Der Spieler erfährt durch seine Aktionen das komische, unsichere Gefühl, etwas Schlechtes ausgelöst zu haben. Da die Rätsel auf unterschiedliche Arten gelöst werden können, was sich wiederum auf die Narrative und den Ausgang des Videospiels auswirkt, stellt dieser tonale Mechanismus eine indirekte Warnung über die Folgenschwere der Entscheidungen dar.

Ein weiteres Beispiel für eine semantische Umkodierung ist der Speicherpunkt des Spiels. Dieser wird durch ein auffälliges, rot leuchtendes Quadrat visualisiert. Bemerkenswert ist, dass der Spieler beim Speichervorgang aus der Position des Quadrates James direkt anblickt. Diese Inszenierung schafft eine temporäre, transdiegetische Schnittstelle zwischen Charakter- und Spielerrealität und verweist auf die extradiegetische Präsenz des Spielers (02:56). Das Durchbrechen der vierten Wand wird außerdem durch Äußerungen der Charaktere umgesetzt (03:10). Die Entwickler des Werkes wählen einen raffinierten Weg, um die Elemente des UI-Designs ohne starke Einbrüche in der Immersion einzubauen, indem sie den Checkpoint als intern transdiegetisches Objekt definieren, welches mittels Charakterdialoge integriert wird. Speicherpunkte sind für gewöhnlich als Ort und Zeitpunkt definiert, an dem der Spieler eine emotionale, strategische Verschnaufpause bekommt. Jedoch erwirken die tonalen Komponenten des roten Quadrates einen gegenläufigen Effekt. Das pulsierende, überlagerte, tieffrequente Klanggemisch, welches teilweise an Insektenbrummen erinnert, sowie ein tieffrequenter, halliger Impact-Sound bei Betreten des Menüs führen zu einem Gefühl von Verunsicherung und Misstrauen.

Insbesondere im Zuge der Analyse der Rätsel- und Checkpoint-SFX fällt folgendes Schema auf: Vom Rezipienten als positiv oder sicher eingestufte Orte oder Situationen werden anhand der audiovisuellen Repräsentation bewusst mit einer konträren Stimmung versehen. Dies führt dazu, dass der Spieler stets einen physisch sowie mental unsicheren Charakterzustand erfährt und durch rapide umschlagende Situationen zu einem zügigen Voranschreiten innerhalb der Spielwelt getrieben wird.

5.2.6 Musik (Video 6)

Das Videospiel beinhaltet als offiziellen Score zahlreiche Lieder. Dieser Abschnitt behandelt aus Gründen der Übersichtlichkeit ausgewählte Stücke, die sich vor allem auf die psychologische Bedeutung der Handlung beziehen. Innerhalb des Gameplays wird extradiegetische Musik vorrangig zur Darstellung negativer oder melancholischer Emotionen mit niedriger Valenz eingesetzt. In *Silent Hill 2: Remake* spielt vor allem das Klavier eine zentrale Rolle in der Musik, wobei es vorrangig der Aufmerksamkeitslenkung auf einzelne Charaktere dient. Die dynamische Bandbreite des Klaviers ermöglicht zahlreiche emotionale Ausdrucksmöglichkeiten. Die Begründung für die emotionale Stimmung des Instruments, die häufig Trauer und Melancholie beinhaltet, liegt in einer Vielzahl an Aspekten der Klangeigenschaften, Spieltechniken und Assoziationen des Hörers mit bestimmten Komponisten wie Chopin oder Rachmaninoff (Flores, o. D.).

Meist bestehen die Stücke des Videospiels aus einer monophonen Klaviermelodie, die später allmählich durch ein Layering mit orchesterlichen Elementen ergänzt wird. Jedoch begegnet dem Spieler größtenteils die minimalistische Anfangssequenz der Stücke.

Als tiefgründiges Beispiel der affektiven musikalischen Gestaltung mittels Klavier dient das Stück „Anam Cara“, was auf Gälisch übersetzt Seelenfreund bedeutet und in der Tonart F-Moll geschrieben ist. F-Moll wird als düstere, traurige Tonart beschrieben, die mit Gefühlen von Tod, Verlust, tiefer Depression und Elend assoziiert wird (The Music Studio, 2023). Zum ersten Mal hört der Nutzer das Lied beim Auslösen einer Jukebox, wobei die diegetische Musik in eine klangliche Untermalung von James' Erinnerungen über die gemeinsame Zeit mit Mary in Silent Hill übergeht, was ihn offensichtlich emotional stark berührt (00:00). Das Überschreiten der diegetischen Ebene wird zu Beginn durch einen verwackelten Plattenspielereffekt realisiert. Mit Einsetzen der Erinnerung stabilisiert sich die Musik in einem sanfteren Lofi-Effekt, die Geräusche der Jukebox werden ausgeblendet und die Lautstärke der Musik erhöht, wobei noch das Plattenknistern zu hören ist. Die kurze, von Melancholie geprägte Stimmung bildet die Trauer und Hoffnung der Hauptfigur ab, während sie gleichzeitig subtil die bevorstehende, noch weit entfernte Enthüllung über die Wahrheit des Geschehens ankündigt. Der ungewollte Abbruch des Songs, der James abrupt in die Realität zurückholt, fungiert als Metapher für den Verlust von Mary sowie das unweigerliche Zerbrechen jeglicher Hoffnung auf das Zurückkehren in die bittersüßen Momente mit James' Seelenfreund. Der Song erscheint mehrmals in verschiedenen Kontexten sowie Tonlagen und wird aus diversen Quellen emittiert. Anam Cara ist zu hören, nachdem Maria das erste Mal scheinbar stirbt und James wieder allein ist (01:13) und nachdem James durch Laura herausfindet, dass Marys Tod noch keine drei Monate her sein kann und er auf dem Klavier des Hotels spielt, in dem die beiden während ihrer Krankheit zu Besuch waren (01:56). Außerdem symbolisiert die kaputte

Jukebox im Hotel durch den Abbruch der Wiedergabe und ihre Platzierung in einer überfluteten Räumlichkeit einen weiteren Zusammenbruch von James' Träumen über ein mögliches Wiedersehen mit Mary (02:15).

Während im Mittelteil des Spiels überwiegend atonale, industrielle Musik und Ambience dominieren, die sich mit monströsen Halluzinationen, dem Thema der Bestrafung und einem psychischen Abstieg auseinandersetzen, finden sich im Kapitel des Lakeview-Hotels meist melodische, minimalistische Klänge wieder. Sobald sich James der Eingangstür des Hotels nähert, sind leise Klaviertöne von links zu bemerken, die jedoch eher wie ein extradiegetischer Anteil wirken. Bevor sich das Klavier als diegetisches, von Laura gespieltes Element herausstellt, jagt diese James und dem Spieler jedoch erst einen Schrecken durch einen tiefen, dissonanten Akkord ein. Die beruhigende, gleichzeitig bedrückende Wirkung des Klaviers wird unterbrochen, worauf durch eine Zwischensequenz mit eindeutig nicht-diegetischem Klavier sofort eine erleichternde, emotionale Stimmung hergestellt wird. Die Klang-Ebene stellt das Hotel als Ort der Sicherheit, Melancholie und emotionaler Verbindung zu Mary dar (02:23).

Das Stück „Promise Reprise“ fungiert als eine Art Leitmotiv für Angela. Als James ihr begegnet, kämpft sie mit Selbstmordgedanken aufgrund ihrer Misshandlungserfahrungen. Sie wirkt erschöpft und apathisch. Das Stück symbolisiert mittels der verträumten, melancholischen, auf- und absteigenden Melodie, Glockenspielakzenten und unterstützenden Streichern die kindliche und verträumte Seite von Angelas Emotionen (03:16). Die verwendete Tonart G-Moll repräsentiert Aspekte, die sich in Angelas kindlichen Schuldgefühlen, der Sehnsucht nach ihrer Mutter und Beziehung zu ihrem Vater widerspiegeln: verträumte Melancholie, Trauer, Versagen und Unbehagen (The Music Studio, 2023; Buja, 2016). Die Melodie stirbt schließlich in einem Ritardando aus, als Angelas dunkle Gedanken stärker zum Vorschein kommen und sie James scheinbar für eine Bedrohung aus ihren Halluzinationen hält, wobei sie emotional zusammenbricht. Dieser Zustand äußert sich durch eine dunkle Ambience mit hohen, dissonanten Pads und tiefen Streichern mit Tremolo-Effekt die in einen schrillen Ton übergehen (04:05). Das Stück Promise Reprise bildet Angelas Verletzlichkeit, und Verlangen nach einer normalen Kindheit ab, während die zweite Sequenz durch klangliche Aspekte ihren traumatischen, von Nervosität und Angst geprägten Zustand widerspiegelt.

Eine spätere Begegnung mit Angela läuft auf eine visuelle Repräsentation ihres Selbstmordes hinaus. Als sie ihre dunklen Gedanken über ihre Kindheit ausdrückt, findet sich in dem unterstützenden Klavierstück der Szene das Glockenspiel wieder (04:37). Als Angela schließlich zwischen Flammen die Treppe hinaufsteigt, wechselt die auditive Untermalung zu hochfrequenten, sich überlagernden Tönen, gemeinsam mit tiefen Impact-Soundeffekten (04:48).

Dieser kontrastreiche Übergang verweist darauf, dass Angela nicht von den Erinnerungen ihrer Kindheit loslassen kann, und betont ihre verwirrte, verlorene Seele. Die entstandenen Depressionen und Schuldgefühle sind so groß, dass sie schließlich keinen anderen Ausweg mehr sieht. Die hochfrequente, angespannte Klangkulisse stellt den Aufstieg in ihren Tod, und nach Ansicht der Autorin einen gewissen Stille-Moment dar, welcher, unterstützt von den bombenartigen Impact-Sounds, ihr Schicksal als unabwendbar definiert.

5.2.7 Ende (Video 7)

Bestimmte Szenen des Spiels entfalten die gefühlsmäßige Wirkung des Audios vor allem durch das Zusammenspiel der einzelnen Sound-Elemente, wie es etwa für die bedeutungsschweren Darstellungen am Ende des Werkes der Fall ist. Daher zieht der folgende Abschnitt ausgewählte Szenen heran, um das spielweite Ineinandergreifen der ermittelten Audio-Schichten Musik, Sound und Stimme exemplarisch zu analysieren.

James sieht sich im Lakeview-Hotel eine Videoaufnahme über die glücklichen Erinnerungen mit Mary im Hotel an. Der Spieler hört Marys Stimme aus dem Video mit einem VHS-Effekt und parallel eine geheimnisvolle Glockenspielmelodie, wobei es zum langsamen Anschwellen eines dissonanten, später schrillen Tongemisches und dem Einsetzen einer mysteriösen, minimalistischen Klaviersequenz kommt. Mit dem Husten von Mary setzt ein intensives Grummeln ein und der Schnitt wechselt zu undeutlichen Aufnahmen, die zeigen, wie Mary von James umgebracht wurde. Die Szene wird auditiv durch dumpfe, verzerrte und laute Schritt- und Atemgeräusche von James, stark verzerrte Kampf- und Erstickungsgeräusche sowie schnelles Herzklopfen von Mary realisiert (00:00). Die Bearbeitung der Charaktergeräusche stellt eine emotionale Distanz zwischen Spieler und Hauptfigur her. Es ist anzunehmen, dass mit der starken Verzerrung des Audios die verstörende Wirkung des Geschehens gedämpft und an die Grundatmosphäre des Spiels angepasst wird. Diese Auffassung variiert selbstverständlich in Abhängigkeit von der emotionalen Wahrnehmung des jeweiligen Nutzers. Die entstellte Ästhetik des Videos lässt die Szene absurd und surreal wirken, sodass Soundeffekte wie das Herzklopfen oder Marys Schreie vom Spieler vermutlich in unterbewusstem Wunschdenken und Verleugnung nicht eindeutig definiert werden und Verunsicherung über das Unbekannte, nicht eindeutig Kategorisierbare stiften. Das im Kapitel der auditiven Halluzination behandelte Einatmen, welches mehrmals über das Werk verteilt ausgelöst wird, ist ohne Verfremdungseffekte und deutlich über der verzerrten Klangmasse zu hören (00:43). Es symbolisiert Marys letzten Atemzug, ihren Schreck als gedankliche Vokalisation formuliert und verleiht den bereits erlebten Halluzinationen so nachträglich einen unheimlichen Schauder-Effekt.

Mit James' erschütterter, erstarrter Reaktion auf das Video setzt ein extradiegetisches Stück mit mysteriöser, zügiger Klaviermelodie und einem natürlich gestalteten Schlagzeugbeat ein, welches die schnellen Gedanken, Verleugnung und mentale Instabilität der Hauptfigur verdeutlicht. Als Laura die Wahrheit über Marys Tod erfährt und die Trauer der beiden hervorsticht, wird dies anhand des Einsatzes einer Violine symbolisiert (01:02).

Kurz darauf findet sich James in kompletter Stille in der Otherworld wieder und erfährt eine Halluzination in Form von Marys Stimme, die ihn über das Radio verzweifelt zu sich ruft (02:04). Hiermit wird deutlich, dass er entgegen allen Beweisen den Tod seiner Frau sowie seine eigene Schuld leugnet und aussichtslos in seinem psychischen Wahn gefangen ist. Dieser Hochpunkt seines mentalen Zerfalls wird sowohl mittels der stark verrottenden Umgebung als auch der pulsierenden, dissonanten, atonalen und industriellen Synthesizer-Klänge der Szene deutlich, die zwischen Ambience und Musik liegen (02:39). Passend zu James' Zustand erweckt dieses Soundscape ein klaustrophobisches, beunruhigendes, kränkliches Gefühl, welches den Verlust seines Verstandes als auch die Labyrinth-Muster seines Denkens und seiner persönlichen Silent Hill-Version abbildet. In einem einsamen, scheinbar endlosen Gang wird eine weitere Halluzination in Form von Marys Stimme in einem vergangenen Dialog mit James während ihrer Krankheit ausgelöst (03:15). Aufgrund des bewusst eingesetzten Halls sowie des narrativ informativen Inhalts der Sprache grenzt sich diese von anderen Halluzinationen ab und kann als intern transdiegetischer Sound betitelt werden, der James' Erinnerung und seine innere Suche nach Rechtfertigung nach außen trägt. Ebenso bezweckt dieser Effekt die Vermittlung der vergangenen Handlung an den Spieler.

Je nach Entscheidungen des Spielers innerhalb des Videospiels können unterschiedliche End-Sequenzen, sogenannte Endings, erreicht werden. Die musikalische Gestaltung sowie das Sounddesign korrelieren mit der Atmosphäre und Bedeutung des entsprechenden Endings, wobei es vermehrt zum Einsatz leiser Synth-Töne und Klaviermelodien kommt, die eine melancholische, depressive oder mysteriöse Stimmung auslösen. Im „Leave“-Ending hat James eine imaginäre Begegnung mit Mary während ihres kranken Zustands. In dieser Version gesteht sich James seine Gefühle und Taten ein. Währenddessen sind eine ruhige, traurige Klaviermelodie und Pads zu hören, die kurz vor dem Moment des Geständnisses in einen tiefen Akkord fallen, um die Dramatik der Handlung und die Fatalität von James' Gefühlen zu unterstreichen (03:54). Nach dem überfüllten, industriellen und lauten Soundscape durch Monster, Zerstörung und psychischen Wahnsinn verursacht, kommen die Handlung sowie der Kampf für Spieler und Charakter zu einem bedrückenden, leisen Ende.

Ein außergewöhnlicher Audioeinsatz ist für das „Bliss“-Ending zu beobachten. Die Videoaufnahme von Mary, welcher der Spieler im Hotel begegnet, bleibt zu Beginn unverändert, wird jedoch im Gegensatz zur ersten Version von ruhigen Streichern und stellenweise einem Klavier begleitet, welche die Aussage von Mary über Silent Hill als friedlichen, heiligen Ort betonen. Die Kamera wird abgestellt und es ist zu sehen, wie James ins Bild läuft. Jedoch ist seine Stimme im Gegensatz zu Marys nicht mit einem schmalbandigen, verrauschten VHS-Effekt realisiert, sondern bleibt gänzlich unverfremdet, wodurch sie eher wie ein Voice-Over erscheint (04:47). Die Diskretion, die James' Stimme zwischen sich, Marys Vokalisation und der visuellen Ebene entstehen lässt, bewirkt beim Hörer das Gefühl, dass etwas nicht stimmt – einen Uncanny-Effekt. Es ist zu sehen, dass James nicht mehr in der normalen Welt vorhanden ist. Sein Transfer in die Aufnahme des Tapes verdeutlicht auf einer symbolischen Ebene, dass er schlussendlich in seiner Hoffnung, seinen Erinnerungen und Illusionen versinkt und schließlich physisch in der eigentlichen Realität an den Folgen seiner psychisch verdrehten Wahrnehmung stirbt. Die Gestaltung der Stimme veranschaulicht, dass sein Übergang in seine Erinnerungen nur eine weitere Halluzination ist, mit der er von der echten Welt aus kommuniziert. James' abweichende Stimme könnte andererseits ebenso als ein unvollendeter Übergang in eine andere Sphäre gedeutet werden.

Nach all den intensiven Bemühungen des Spielers, James' Vergangenheit in einem mental bedrückenden Labyrinth von Rätseln aufzudecken, verwehren die Endsequenzen ein befriedigendes Erfolgserlebnis. Stattdessen lässt das Werk den Nutzer mit ernüchternden Erkenntnissen, James' Tod oder neuen, unbeantworteten Fragen zurück.

Das Sounddesign fungiert aktiv als Bestandteil dieser Wirkung. Es bietet keine Erlösung für die angesammelte emotionale Spannung, sondern zieht seine Hörer schlussendlich nur weiter in die Gefühlsebene hinein – ein Effekt, der über die Grenzen des Spiels hinaus nachhallt.

6. Fazit

Zusammenfassend lässt sich sagen, dass Sounddesign auf die emotionale Ebene in Videospielen – besonders für die Gradwanderung zwischen einem unheimlichen und genießbaren Spielerlebnis in Horror-Spielen – einen signifikanten Wirkungsmechanismus in der kontrollierten Beeinflussung von Stimmung und Immersion darstellt. Dieses Toleranzfenster des Horror-Audios umfasst keine gestalterischen Regeln oder Grenzen und setzt sich mehr aus strategischen, abstrakten Ansätzen zusammen.

Das Ziel der vorliegenden Arbeit lag in der Ergründung theoretischer und angewandter Methoden sowie der Auswirkungen des psychologischen Sounddesigns in Horror-Videospielen. Hierfür wurden sowohl fundamentale Konzepte der Psychoakustik als auch des Game-Audios ergründet und auf das Horror-Genre spezifiziert. Die akustische, metaphorische Verdeutlichung von Charakter- und Umgebungszuständen durch Soundeffekte und Musik stellt sich als ein Kernpunkt psychologischer Horrorspiele heraus. Ein ebenso signifikanter Aspekt besteht im gezielten Aussparen und Manipulieren akustischer Informationen sowie in einer durch tonale Verfremdung erwirkte Entmenschlichung von Kreaturen. Diese Methoden führen insbesondere zu den Gefühlen Angst, Verunsicherung und Schock. Zur spezifischen, einheitlichen Betrachtung der Sounddesign-Methoden wurde das Videospiel *Silent Hill 2: Remake* als vertiefendes Beispielwerk herangezogen, welches sich durch seine starke Nutzung von Ton zum Ausdruck psychologischer Aspekte auszeichnet, wobei die Gefühle von Charakter und Spieler ineinander übergehen. Hier stechen insbesondere ein hohes Level an auditiven Halluzinationen sowie Kreatur- und SFX-Sounds, die an der Uncanny-Schwelle liegen, heraus. Die intensive Nutzererfahrung resultiert außerdem daraus, dass das Audio selbst in sicheren Situationen ein unterschwelliges Gefühl der Bedrohung vermittelt.

Die Kapitel der Psychoakustik und Game Studies bilden ein Fundament für das Verständnis des emotionalen Sounds in interaktiven Medien. Die Bestandteile des Gesamtkonzepts weisen allerdings eine Variation und mögliche Tiefe auf, die den Umfang dieser Arbeit überschreiten. Es ist eine detailliertere Erforschung einzelner Sound-Komponenten erforderlich, um deren emotionale Auswirkungen auf spezifische auditive Merkmale und einen audiovisuellen Kontext zurückführen zu können. Die erforschten Zusammenhänge bieten einen Leitfaden, um beispielsweise Entwickler von Sounddesign in Horror-Spielen bei technischen sowie gestalterischen Entscheidungen oder der Optimierung ihres emotional-strategischen Ansatzes zu unterstützen.

In Hinblick auf die Zukunft lässt sich festhalten, dass großer Entwicklungsbedarf in der Implementierung von Sound besteht, um die Simulierung alterierter Bewusstseinszustände sowie die Übertragung psychologischer, visueller Phänomene auf die Audio-Ebene realitätsgetreu nachzubilden. Die gezielte Anwendung sowie die nachweisbare emotionale Wirkung dieser Konzepte im Rahmen von Videospielen benötigen ausführliche empirische Forschungen. Mit dem Fortschritt visueller Technologien ist außerdem eine detailliertere Steuerung der Spielererfahrung durch Unterschiede oder Übereinstimmung von Bild und Ton zu erwarten. Durch den Anspruch an ein stetig wachsendes Level an Realismus und Immersion entsteht allerdings die Gefahr, ein zu abgestimmtes, emotional überreizendes Spielerlebnis zu erschaffen, weshalb diese Entwicklung kritisch zu hinterfragen ist. Nach Meinung der Autorin besteht ein Lösungsansatz hierfür in einem vermehrt strategischen als auch experimentellen Vorgehen. Ein zeitlich unvorhersehbarer, insgesamt harmonischer Einsatz von Soundeffekten in Kombination mit einer, an visuelle Komponenten adaptierte, Echtzeitgenerierung von Sound (Game Developers Conference, 2017, 28:44) bieten die Möglichkeit, bei jedem Durchlauf ein gleichermaßen unvorhersehbares, schlüssiges Soundscape zu schaffen, welches unheimliche Fiktion und Realismus vereint.

Literaturverzeichnis

- Asutay, E., Bergman, P., Genell, A., Kleiner, M., Tajadura-Jimenez, A., & Västfjäll, D. (2012). Emoacoustics: A Study of the Psychoacoustical and Psychological Dimensions of Emotional Sound Design. *Journal of the Audio Engineering Society*, 60(1/2), 21–28. https://www.academia.edu/1650139/Emoacoustics_A_Study_of_the_Psychoacoustical_and_Psychological_Dimensions_of_Emotional_Sound_Design
- Ball, S. M. (2006). *The Uncanny in Japanese and American Horror Film: Hideo Nakata's Ringu and Gore Verbinski's Ring*. (Masterarbeit, North Carolina State University). <http://www.lib.ncsu.edu/resolver/1840.16/708>
- Bayerisches Landesamt für Umwelt. (o. D.). *Die Wirkung von Lärm auf den Menschen*. LfU Bayern. https://www.lfu.bayern.de/laerm/laerm_allgemein/wirkung_auf_menschen/index.htm#:~:text=Bei%20Pegeln%20ab%2040%20bis,L%C3%A4rmwirkungsforschung%20von%20einer%20gesundheitlichen%20Beeintr%C3%A4chtigung Abgerufen am 02. März 2025.
- Bridgett, R. (2021). Leading with Sound: Proactive Sound Practices in Video Game Development. Routledge.
- Brinkman, W. P., Hoekstra, A., & Van Egmond, R. (2015). The Effect Of 3D Audio And Other Audio Techniques On Virtual Reality Experience. In B. K. Wiederhold, G. Riva & M. D. Wiederhold (Hrsg.), *Annual Review of Cybertherapy and Telemedicine 2015* (Bd. 219, 44–48). IOS Press. <https://doi.org/10.3233/978-1-61499-595-1-44>
- Buja, M. (2016, 11. März). *How You Should Feel in the Key of G Minor*. Interlude. <https://interlude.hk/feel-key-g-minor/> Abgerufen am 06. April 2025.
- Caroll, N. (1990). The Philosophy of Horror: Or, Paradoxes of the Heart. Routledge. <https://doi.org/10.4324/9780203361894>
- Clasen, M. (2023). The New Science of Recreational Fear. *Anglo Files*, 2023(207), 36–41. [36Fileshttps://cc.au.dk/fileadmin/dac/Projekter/Recreational_Fear_Lab/2023_AngloFiles_207_RF_MC.pdf](https://cc.au.dk/fileadmin/dac/Projekter/Recreational_Fear_Lab/2023_AngloFiles_207_RF_MC.pdf)
- Collins, K. (2008). *Game Sound: An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design*. The MIT Press.
- Dead by Daylight Wiki (o. D.). *Theme Music*. Dead by Daylight Wiki. https://deadbydaylight.fandom.com/wiki/Theme_Music Abgerufen am 23. März 2025.

- Dennis, A. (2020, 11. Februar). *The Sound of Silence: The Use of Silence In Games*. Medium. <https://anthonydennis-24878.medium.com/the-sound-of-silence-the-use-of-silence-in-games-9d53f1aaef25> Abgerufen am 24. März 2025.
- Dobie R. A., Van Hemel S., Committee on Disability Determination for Individuals with Hearing Impairments & National Research Council. (2004). *Hearing loss: Determining eligibility for social security benefits*. National Library of Medicine. <https://www.ncbi.nlm.nih.gov/books/NBK207834/>
- Eerola, T. (2018). Music and Emotion. In R. Bader (Hrsg.), *Springer handbook of systematic musicology* (539–554). Springer Verlag. https://doi.org/10.1007/978-3-662-55004-5_29
- Ekman, I. & Kajastila, R. (2009, 11.–13. Februar). *Localisation Cues Affect Emotional Judgements – Results from a User Study on Scary Sound [Konferenzbeitrag]*. *AES 35th Conference on Audio for Games*, London, UK. https://www.researchgate.net/publication/233406208_Localisation_Cues_Affect_Emotional_Judgements-Results_from_a_User_Study_on_Scary_Sound
- Ekman, I., & Lankoski, P. (2009). Hair-Raising Entertainment: Emotions, Sound, and Structure in Silent Hill 2 and Fatal Frame. In B. Perron (Hrsg.), *Horror Video Games: Essays on the Fusion of Fear and Play* (S. 179–197). McFarland & Company, Inc. <https://www.lehmanns.de/shop/sachbuch-ratgeber/27197219-9780786454792-horror-video-games>
- Epps, D. (2023, 31. Oktober). *Why do we love horror games? Psychologists explain our morbid curiosity*. Digital Trends. <https://www.digitaltrends.com/gaming/why-do-we-love-horror-games-analysis/> Abgerufen am 05. März 2025.
- Erlich, N., V. Lipp, O. & Slaughter, V. (2013). Of hissing snakes and angry voices: human infants are differentially responsive to evolutionary fear-relevant sounds. *Developmental Science*, 16(6), 894–904. <https://doi.org/10.1111/desc.12091>
- Farnell, A. (2010). *Designing Sound*. The MIT Press.
- Fladt, H. (2012). *Der Musikversteher: was wir fühlen, wenn wir hören*. (2. Aufl.). Aufbau.
- Flores, K. (o. D.). *Why is Piano Music Sad?*. Treesidemusicacademy. <https://www.treesidemusicacademy.com/features/why-is-piano-music-sad> Abgerufen am 07. April 2025.
- Gabrielsson, A. (2001). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, 5(1), 123–147. <https://doi.org/10.1177/10298649020050S105>

- Goh, R. Z., Phillips, I. B., & Firestone, C. (2023). The perception of silence. *PNAS*, 120(29), Artikel e2301463120. <https://doi.org/10.1073/pnas.2301463120>
- Görne, Prof. T. (2017). *Sounddesign: Klang, Wahrnehmung, Emotion*. Hanser.
- Grimshaw, M. N. (2009, 02.–03. August). *The audio Uncanny Valley: Sound, fear and the horror game* [Konferenzbeitrag]. Audio Mostly, Glasgow, United Kingdom. S. 21–26. <https://vbn.aau.dk/en/publications/the-audio-uncanny-valley-sound-fear-and-the-horror-game>
- Harrer, G., & Harrer, M. E. (2008). Emotionale Vorgänge beim Hören von Musik und ihre Resonanz im Vegetativum. *Neuro Rehabil*, 14(5), 238–246. Hippocampus. https://www.hippocampus.de/media/316/cms_4a94f09faa8a1.pdf
- Hesse, H. (2003) *Musik und Emotion: Wissenschaftliche Grundlagen des Musik-Erlebens*. Springer.
- Hill, I. (o.D.). *Irony and Juxtaposition*. BSoA Digital Space. <https://digiタルspace.bradfordcollege.ac.uk/10626277/irony-and-juxtaposition/> Abgerufen am 24. März 2025.
- Hunicke, R., Leblanc, M., & Zubek, R. (2004). *MDA: A Formal Approach to Game Design and Game Research* [Konferenzbeitrag]. Association for the Advancement of Artificial Intelligence (AAAI), San Jose, Costa Rica.
- Huckvale, D. (1990), Twins of Evil: An Investigation into the Aesthetics of Film Music. *Popular Music*, 9(1), 1–35. <https://doi.org/10.1017/S0261143000003718>
- IMDb. (o.D.). *Silent Hill 2: Benutzerrezensionen*. IMDb. <https://www.imdb.com/de/title/tt0282975/reviews/> Abgerufen am 25. April 2025.
- Isbister, K. (2023, 18. Dezember). *How Games move us*. The MIT Press Reader. <https://thereader.mitpress.mit.edu/how-games-move-us/> Abgerufen am 05. März 2025.
- Jones, E. (2021, 26. Mai). *Resident Evil Village: About the Baby*. Medium. <https://evanjones-168.medium.com/resident-evil-village-about-the-baby-9c572384ffc8> Abgerufen am 23. März 2025.
- Jørgensen, K. (2007). On Transdiegetic Sounds in Computer Games. *Northern Lights: Film and Media Studies Yearbook*, 5(1), 105–117. Intellect Press. https://doi.org/10.1386/nl.5.1.105_1
- Kemp, S. (24. April 2024). *Digital 2024 April Global Statshot Report*. Datareportal. <https://datareportal.com/reports/digital-2024-april-global-statshot> Abgerufen am 15. April 2025.

- Kockesen, G. (2015). *Effects of Switching Audio Cues in Survival Horror Games on Level of Fear*. (Masterarbeit, Hogeschool van Amsterdam). <https://theses.liacs.nl/1241>
- Kohler, C. (2007, 19. Oktober). *Interview: Silent Hill Sound Designer Akira Yamaoka*. WIRED. <https://www.wired.com/2007/10/interview-silence/> Abgerufen am 24. März 2025.
- Kromand, (2008, 22.–23. Oktober). *Sound and the diegesis in survival-horror games* [Konferenzbeitrag]. Audio Mostly, Piteå, Sweden. S. 16–19. Nicht mehr verfügbar.
- Krumhansl, C. L. (1997). An Exploratory Study of Musical Emotions. *Canadian Journal of Experimental Psychology*, 51(4), 336–352. 10.1037/1196-1961.51.4.336
- Leenders, M. J. (2012). *Sound für Videospiele: Besondere Kriterien und Techniken bei der Ton- und Musikproduktion für Computer- und Videospiele*. Schüren.
- Lovecraft, H. P. (1927). *Supernatural Horror in Literature*. Internet Archive. <https://archive.org/details/lovecraft-1945-sup.-horror/page/n5/mode/2up>. Abgerufen am 26. April 2025.
- Marks, A. (2017). *Complete Guide to Game Audio: For composers, sound designers, musicians and game developers* (3. Aufl.). CRC Press.
- Mraz, R. (2021, 13. Mai). *How procedural audio brings sounds to life in video games*. Splice. <https://splice.com/blog/procedural-audio-video-games/> Abgerufen am 08. März 2025.
- Musiek, F., Ballingham, T. M., Liu, B., Paulovicks, J., Swainson, B., Tyler, K., Vasil, K., & Weihing, J. (2007). Auditory hallucinations: An audiological perspective. *The Hearing Journal*, 60(9), 32–52. <https://doi.org/10.1097/01.HJ.0000295756.39258.41>
- Neff, Dr. M. A. (o. D.). *Using the Arousal-Valence Model to Better Your Emotional Intelligence*. Neurodivergent Insights. <https://neurodivergentinsights.com/blog/arousal-valence-model> Abgerufen am 26. Februar 2025.
- Opton, E., Nomikos, M. S. & Averill, J. R. (1968). Surprise versus suspense in the production of stress reaction. *Journal of Personality and Social Psychology* 8(2), 204–208. <https://doi.org/10.1037/h0025274>
- Palumbo, A. (2021, 9. April). *PS5 Is an Amazing Fit for Horror, Its Tech Contributes Hugely to Immersion, Says RE Village Director*. Wccftech. <https://wccftech.com/ps5-is-an-amazing-fit-for-horror-its-tech-contributes-hugely-to-immersion-says-re-village-director/> Abgerufen am 15. März 2025.

- Perron, B. (2004, 14.–16. September). *Sign of a threat: The effects of warning systems in survival horror games* [Konferenzbeitrag]. COSIGN 2004 Proceedings. Split, Croatia. <http://www.cosignconference.org/conference/2004/papers>
- Perron, B. (2018). *The World of Scary Video Games: A Study in Videoludic Horror*. Bloomsbury.
- Picard, M. (2009). Haunting Backgrounds: Transnationality and Intermediality in Japanese Survival Horror Video Games. In B. Perron (Hrsg.), *Horror Video Games: Essays on the Fusion of Fear and Play* (S. 93–118). McFarland & Company, Inc. <https://www.lehmanns.de/shop/sachbuch-ratgeber/27197219-9780786454792-horror-video-games>
- Raffaseder, H. (2010). *Audiodesign*. (2. Aufl.). Hanser.
- Reybrouck, M., & Eerola, T. (2017). Music and Its Inductive Power: A Psychobiological and Evolutionary Approach to Musical Emotions. *Frontiers in Psychology*, 8, Artikel 494. <https://doi.org/10.3389/fpsyg.2017.00494>
- Roberts, R. (2014). Fear of the Unknown: Music and Sound Design in Psychological Horror Games. In K. J. Donnelly, W. Gibbons & N. Lerner (Hrsg.), *Music in Video Games: Studying Play* (S. 138–149). Routledge. <https://doi.org/10.4324/9781315882697>
- Roederer, J. G. (2000). *Physikalische und psychoakustische Grundlagen der Musik*. (3. Aufl., F. Mayer-Pfeiffer & S. Güss, Übers.). Springer.
- Rouse, R. (2009). Match made in Hell: The inevitable success of the Horror Genre in Videogames. In B. Perron (Hrsg.), *Horror Video Games: Essays on the Fusion of Fear and Play* (S. 13–23). McFarland & Company, Inc. <https://www.lehmanns.de/shop/sachbuch-ratgeber/27197219-9780786454792-horror-video-games>
- Roux-Girard, G. (2011) Listening to fear: A study of sound in horror computer games. In M. Grimshaw (Hrsg.), *Game Sound Technology and Player Interaction: Concepts and Developments* (S. 192–212). IGI Global. <https://doi.org/10.4018/978-1-61692-828-5>
- Sinclair, J. L. (2020). *Principles of game audio and sounddesign: Sounddesign and audio Implementation for interactive and immersive media*. Routledge. <https://doi.org/10.4324/9781315184432>
- SteamDB. (o. D.). *Steam Horror Game Releases by Month*. SteamDB. <https://steamdb.info/stats/releases/?tagid=1667> Abgerufen am 26. April 2025.

Suter, B., Bauer, R., & Kocher M. (Hrsg.). (2021). *Narrative Mechanics: Strategies and Meanings in Games and Real Life*. Transcript. <https://doi.org/10.14361/9783839453452>

The Music Studio (2023, 11. Oktober). *How Chords and Key Impact Emotion in Music*. <https://www.themusicstudio.ca/how-chords-and-key-impact-emotion-in-music/#:~:text=Fm%20Dark%20and%20funereal%2C%20evoking,death%2C%20loss%2C%20and%20misery> Abgerufen am 06. April 2025.

Tinwell, A. & Grimshaw M. (2009, 23.–25. April). *Survival horror games: An uncanny modality* [Konferenzbeitrag]. Thinking After Dark International Conference, Montreal, Canada. https://www.academia.edu/2768928/Survival_horror_games_An_uncanny_modality

Toprac, P., Abdel-Meguid, A. (2011). Listening to fear: A study of sound in horror computer games. In M. Grimshaw (Hrsg.), *Game Sound Technology and Player Interaction: Concepts and Developments* (S. 176–191). IGI Global. <https://doi.org/10.4018/978-1-61692-828-5>

Västfjäll, D. (2012). Emotional Reactions to Sounds without Meaning. *Psychology*, 3(8), 606–609. <https://doi.org/10.4236/psych.2012.38091>

Weinel, J., & Cunningham, S. (2017). *Simulating Auditory Hallucinations in a Video Game: Three Prototype Mechanisms* [Konferenzbeitrag]. Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences. London, UK. Artikel 18, S. 1–7. <https://doi.org/10.1145/3123514.3123532>

Weiss, P. P. (2015). *Wenn Design die Materie verlässt: Sound. Das Design der Emotionen, der Imagination und der Lebendigkeit*. BoD – Books on Demand.

Whalen, Z. (2004). Play Along: An Approach to Videogame Music. *The international journal of computer game research*, 4(1). <https://www.gamestudies.org/0401/whalen/>

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8(4), 494–521. <https://doi.org/10.1037/1528-3542.8.4.494>

Videos

- Blumstein, D. (2012, 14. Dezember). *The Sound of Fear: Dan Blumstein at TEDxUCLA* [Video]. YouTube. <https://www.youtube.com/watch?v=HQQmFocLDng>
- Game Developers Conference. (2017, 11. Juli). *The Sound of No Man's Sky* [Video]. YouTube. https://youtu.be/zKJ_XuQjiw?si=Q18ZWtzISOV9PzgO
- Resident Evil. (2017a, 22. Juni). *Resident Evil 7 biohazard Developer Interviews: Music and Silence* [Video]. YouTube. https://www.youtube.com/watch?v=sf8PYS_u3ic
- Resident Evil. (2017b, 10. August). *Resident Evil 7 biohazard Dev Interviews: Go Tell Aunt Rhody* [Video]. YouTube. <https://www.youtube.com/watch?v=LxiGLM4njpw>
- Timodus. (2021, 06. Mai). *Resident Evil Village (PS5) Mutated Baby* [Video]. YouTube. <https://youtu.be/fPSe13ECEio?si=FOn90MGvfLy6WMJ6>
- ToneArtistLive. (2014, 29. September). *Akira Yamaoka On Composing For Video Games* [Video]. YouTube. https://www.youtube.com/watch?v=_1_nYgE8rMM
- VG Moments. (2023, 30. Januar). *Issac hears Twinkle Twinkle Little Star in the Elevator - Dead Space Remake* [Video]. YouTube. <https://www.youtube.com/watch?v=yf8bY0wPIng>

Videospiele

- 2K Games. (2007). *BioShock* (Windows Version) [Video Game]. 2K Games.
- Behaviour Interactive. (2016). *Dead by Daylight* (PC-Version) [Video Game]. Behaviour Interactive.
- Capcom. (2017). *Resident Evil 7: Biohazard* (PC-Version) [Video Game]. Capcom.
- Capcom. (2021). *Resident Evil: Village* (Playstation 5 Version) [Video Game]. Capcom.
- Electronic Arts. (2023). *Dead Space* (Playstation 5 Version) [Video Game]. Motive Studio.
- Konami. (1999). *Silent Hill* (Playstation Version) [Video Game]. Team Silent.
- Konami. (2001). *Silent Hill 2* (PlayStation 2 Version) [Video Game]. Team Silent.
- Konami. (2008). *Silent Hill: Homecoming* (Playstation 3 Version) [Video Game]. Double Helix Games.
- Konami. (2024). *Silent Hill 2 (Remake)* (PC-Version) [Video Game]. Bloober Team.
- Tecmo. (2001). *Fatal Frame* (PlayStation 2 Version) [Video Game]. Tecmo.