

Die akustische Architektur des Computerspiels

Inaugural-Dissertation
zur Erlangung der Doktorwürde
der
Philosophischen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu Bonn

vorgelegt von

Sebastian Lorenz

aus

Köln

Bonn 2014

Gedruckt mit der Genehmigung der Philosophischen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Zusammensetzung der Prüfungskommission:

Prof. Dr. Marion Gymnich

(Vorsitzende)

Prof. Dr. Bettina Schlüter

(Betreuerin und Gutachterin)

Prof. Dr. Erik Fischer

(Gutachter)

Prof. Dr. Sabine Sielke

(weiteres prüfungsberechtigtes Mitglied)

Tag der mündlichen Prüfung: 20.12.2013

Inhaltsverzeichnis

Einleitung	1
1 Fundamente	7
1.1 Sound im Kontext von Software	8
1.2 Spiel und Erzählung: Von Zahlen zu Zeichen	25
1.3 Lernvorgänge und Spielweisen	41
2 Akustische Bausteine im Wandel der Zeit	57
2.1 Frühe Formen der Vertonung	58
2.2 Programmierter Sound	66
2.3 Reale Klänge in virtuellen Welten	75
2.4 Vergangenes, Gebliebenes und Wiederkehrendes	90
3 Zeitliche Konstruktionsprinzipien	101
3.1 Zeitstrukturen des Computerspiels	102
3.2 Ereignisbasiertes Sounddesign	110
3.3 Wahrnehmungshilfen zeitlicher Abläufe	114
3.4 Musik im Spiel	121
4 Räumliche Konstruktionsprinzipien	143
4.1 Die Vertonung von Spielräumen	145
4.2 Raumbasiertes Sounddesign	157
4.3 Navigation und Positionierung	164
4.4 Spiel- und Erzählformen in Klangräumen	169
5 Adaption kinematographischer Gestaltungsmuster	183
5.1 Anpassung filmischer Vertonungsprinzipien	185
5.2 Szenengestaltung	193
5.3 Gesprochene Sprache	204
6 Audio-Konfiguration	215
6.1 Anpassbare Hörerlebnisse	217
6.2 Multiple Darstellungsmodi	231

6.3 Die Simulation technischer Medien	235
Ausblick	247
Literaturverzeichnis	254

Einleitung

Every possible choice Stanley could make has been designed for him long before he ever set foot here. The narrator wanted to kill him. Stanley was already dead, from the moment he hit „start“. There is no salvation for either of these two, I'm afraid. The narrator had as little power over Stanley as Stanley did over the paths that he walked. But listen to me.. this story is not over. You can still save these two. You can stop the program before they both fail. Push Escape and press Quit. There's no other way to beat this game. As long as you move forward, you'll be walking someone elses path. Stop now and it'll be your only true choice. Whatever you.. [Metallisches Knirschen. Schwarzer Bildschirm.]

Mit dem Verstummen der Voice-Over-Stimme aus den Boxen kehrt die Sprachfähigkeit tausender Stanleys vor den Bildschirmen zurück. In Foren, Portalen und Blogs hinterlassen staunende Spieler¹ ihre Meinung über die am 27. Juli 2011 erschienene *Half Life 2* Modifikation *The Stanley Parable* und eine ungewöhnliche Spielerfahrung. Die Stimme des Erzählers und der oben zur Wort kommenden Erzählerin stehen in ständigem Konflikt mit den Entscheidungen der Spieler. Einige Berichte enthalten das Eingeständnis, dass sich nur schwer darüber berichten lässt und legen den Lesern nahe, für wenige Stunden Spieler zu werden.² Andere Beiträge zerlegen und analysieren die Mechanik und Inhalte des Spiels oder stellen Querverbindungen zu Filmen wie *American Beauty*

¹Der Begriff „Spieler“ meint im Folgenden stets Spielerinnen und Spieler gleichermaßen.

²So schreibt z. B. das Gee Magazin: „Jeder inhaltliche Erklärungsversuch würde die Qualität des Spielerlebnisses mindern. Also bitte, betrachtet eure Unwissenheit als Chance auf ein verstörendes Abenteuer, das seine Spuren hinterlässt. Lediglich die Hürde der Installation muss überwunden werden [...]“ (<http://www.geemag.de/2011/08/05/the-stanley-parable/>, 01.11.2012)

her.³ Große Einigkeit herrscht darüber, dass *The Stanley Parable* zum Nachdenken anregt und Einblicke darüber vermittelt, wie Computerspiele funktionieren. Es baut seine Kulissen auf, nur um sie vor den Augen und Ohren der teilnehmenden Spieler wieder niederzureißen. Es legt zentrale Prinzipien des Designs offen, die Spiele gewöhnlich kaschieren oder verbergen. Es ist ein Computerspiel, das etwas über das Computerspiel sagt.

Das Werk des 22-jährigen Modders David Wreden demonstriert durch seine Anlage zur Selbstreflexion beispielhaft, wie ausdrucksstark und elaboriert Computerspiele heute sind. In der Tat zeigt ein Rundumschlag durch Spieler- und Entwicklercommunities, Gesellschaft, Politik, und Wissenschaft, dass ihnen im Positiven wie im Negativen viel mehr zugetraut und zugeschrieben wird, als nur unterhaltendes Spielzeug zu sein. Kritiker verteufeln sie als Killerspiele, Suchtmittel und Ausbildungswerkzeuge eines militärischen Unterhaltungskomplexes.⁴ Befürworter sehen in ihnen eine Lernhilfe und ein Mittel zur politischen Bildung.⁵ Der Deutsche Kulturrat setzt sich seit Jahren dafür ein, das Computerspiel als Kulturgut zu positionieren. Die Game-Designerin Jane McGonigal sieht in der Anwendung von Spielprinzipien und Spielmechaniken auf nicht-spielspezifische Domänen – bezeichnet als *Gamification* – einen Weg zum Weltfrieden.⁶

All diese ernsthaften Auseinandersetzungen mit dem Spiel deuten auf die Maturität des Mediums und seinen Einfluss auf unser Leben hin. Die Oberflächeneffekte dieser Entwicklung sind unverkennbar. Wer öffentliche Verkehrsmittel benutzt, sieht mehr Fahrgäste, die auf ihren

³„I downloaded an interesting Half Life 2 mod, called The Stanley Parable. It's a fascinating little experiment into narrative gaming, and comes off as a kinda hybrid of Bastion, 1984 and American Beauty.“ (User Null, <http://www.theclawproject.net/forum/viewtopic.php?f=5&t=1660>, 01.11.2012). Einen direkten Bezug zum Film etabliert die Übernahme des Soundtracks von Thomas Newman im Intro der *Stanley Parable*.

⁴Bereits 1993 schrieb Bruce Sterling einen Beitrag über den Zusammenprall von Unterhaltung und militärischer Fortbildung in virtuellen Schlachtfeldern (Sterling, 1993). Seit 2002 befindet sich mit *America's Army* ein von der U.S. Army veröffentlichtes Computerspiel auf dem Markt, das offen für Rekrutierungszwecke eingesetzt wird.

⁵James Paul Gee listet in (Gee, 2003) 36 Lernprinzipien auf, die in der Auseinandersetzung mit Computerspielen auftreten. Im Zuge der *Serious Games* Bewegung entstanden in den vergangenen Jahren mehrere Minigames, die sich der kritischen Vermittlung politischer und wirtschaftlicher Zustände verschrieben haben. Siehe z. B. <http://www.newsgaming.com> und <http://www.molleindustria.org>.

⁶(McGonigal, 2012, 8)

Smartphones spielen als Zeitung lesen. Das Motto „work hard, play hard“ ist längst die Devise einer Arbeitswelt, in der die Arbeit zum Spiel und das Spiel zur Arbeit geworden ist.⁷ Computerspiele und ihre angelagerten Diskurse sind vielleicht der eindringlichste Ausdruck einer Software-Kultur, „[...] a culture where the production, distribution, and reception of most content – and increasingly, experiences – is mediated by software.“⁸ Um die Wünsche, Ängste und Mentalitäten unserer Informationsgesellschaft zu verstehen, ist eine Auseinandersetzung mit ihnen unerlässlich. Mit der Fähigkeit zur Selbstreflexion erreichen sie offenbar ein neues Stadium ihrer Expressionsfähigkeit. Wie kam es dazu, dass Computerspiele so ausdrucksstark geworden sind? Was sind die Eigenheiten ihrer Sprache?

Einen Ansatz zur Beantwortung dieser Frage gibt der Medientheoretiker Lev Manovich in seinen Arbeiten über die „Sprache der Neuen Medien“ und den daran anschließenden Überlegungen über die soziale und kulturelle Dimension von Software. Er beschreibt die Entwicklung des Metamediums Computer als einen zweigliedrigen Prozess, der auf eine „Medienhybridisierung“⁹ hinausläuft. Demnach werden im Zuge der Software-Entwicklung nicht nur die Inhalte verschiedener Medientypen vermischt, sondern auch ihre „techniques, working methods, and ways of representation and expression“¹⁰. Die Fülle dieser Variantenbildung tritt nirgendwo deutlicher hervor als beim Computerspiel. Übergeordnete Genres wie *real time strategy* (RTS), *first person shooter* (FPS) und *massively multiplayer online role playing game* (MMORPG) bezeichnen nur einen Bruchteil der etablierten Spezies des Computerspiels, die sich im Laufe seiner Geschichte herausgebildet haben. Ständig treten neue Kombinationen hervor, die sich nicht eindeutig einem Genre oder Spieltyp zuordnen lassen. Die zusammengesetzte Bezeichnung der *Stanley Parable* als *experimental narrative-driven first person game* ist kein Einzelfall, sondern die Regel eines Vorgangs der kontinuierlichen Ausdifferenzierung. Im Epizentrum dieser „deep remixability“¹¹ verortet Manovich einen wachsenden Bestand bausteinartiger Ausdrucksmittel, die sich im Digitalen kreuzen und zu neuen hybriden Formen verschmelzen.

⁷vgl. (Yee, 2006)

⁸(Manovich, 2008, 19)

⁹(ebd., 72)

¹⁰(ebd., 25)

¹¹(ebd., 25)

Both the simulated and new media types – text, hypertext, still photographs, digital video, 2D animation, 3D animation, navigable 3D spaces, maps, location information – now function as building blocks for new mediums.¹²

Eine Kategorie dieser Bausteine, die bislang weder in den Game Studies noch in den Software Studies eine bemerkenswerte Beachtung gefunden hat, umfasst *akustische Elemente* wie die eingangs zu Wort kommende Voice-Over-Stimme. Akustische Elemente scheinen gleichermaßen dazu geeignet zu sein, als Ausgangspunkt für eine Untersuchung der Ausdrucksfähigkeit des Computerspiels zu dienen und einen Beitrag zum Verständnis der softwarebedingten Medienhybridisierung zu leisten. Gesprochene Sprache und Musik gehören schließlich zu unseren ausdrucksstärksten und elaboriertesten Kommunikationsmitteln. Darüber hinaus bestimmen auch Geräusche in hohem Maße sowohl die von uns wahrgenommene Umwelt, als auch medial inszenierte Arbeiten wie Radiosendungen, Hörspiele, Filme und nicht zuletzt Computerspiele.

Wenn akustische Elemente als Bausteine des Computerspiels behandelt werden, dann lässt sich ihr Gesamtgefüge als *akustische Architektur* beschreiben. Die Metapher der Architektur eignet sich insofern gut, als sie nicht nur das fertige Bauwerk, sondern auch den organisatorischen und gestalterischen Prozess des Bauens mit einschließt. Sie eröffnet die Möglichkeit, von den Rohmaterialien und konzeptuellen Entwürfen bis hin zum gegossenen Fundament und den darauf gesetzten Etagen unterschiedliche Phasen des Bauvorgangs in den Blick zu nehmen. Darüber hinaus sind die Architektur-Metapher und die mit ihr verbundenen Bilder des Bauens in der Software- und Computer-spieleentwicklung seit Jahrzehnten in Gebrauch. Software-Architekten und -Ingenieure formen aus Code jene komplexen Systeme, die unsere Informationsgesellschaft definieren.¹³ Level-Designer konstruieren mit ihren Level-Architekturen die begehbaren Räume virtueller Spielwelten. Der Architektur-Begriff bietet sich deshalb an, um zwischen mehreren miteinander verbundenen, aber in der Regel separat behandelten Bereichen – Sound, Spiel und Software – eine Bresche zu schlagen.

¹²(ebd., 72)

¹³Mitunter beziehen sich Verfahren der Software-Entwicklung auch direkt auf Vorbilder aus der Baukunst. So adaptierten Erich Gamma, Richard Helm, Ralph Johnson und John Vlissides den Begriff des „Entwurfsmusters“ von dem Architekten Christopher Alexander (Gamma u. a., 1994, 12).

Die akustische Architektur umfasst von der digitalen Repräsentation akustischer Elemente über ihre Verknüpfungen mit Spielaspekten bis hin zur akustischen Begleitung konkreter Spieldurchläufe alle Stufen des Vertonungsprozesses von Computerspielen. Sie ist ein übergeordneter, für die Ziele dieser Arbeit konzipierter theoretischer Umriss eines bestimmten Teilbereichs des Computerspiels. In der Praxis sind auf verschiedenen Ebenen der akustischen Architektur von Programmierern, Ingenieuren und Game Designern über Sound Designer und Komponisten bis hin zur Gemeinschaft der Spieler verschiedene Personengruppen involviert. Ihr gemeinsames Zutun bestimmt die Form akustischer Elemente als bausteinartige Ausdrucksmittel, ihr Aufgehen in der Hybridität neuer Medien sowie ihre diskursiv verhandelte Bedeutung im Rahmen des Umgangs mit Computerspielen.

Das Ziel dieser Arbeit besteht darin, durch eine Analyse der akustischen Architektur einen Beitrag zum Verständnis des Computerspiels im Speziellen und unserer vom Spiel dominierten Software-Kultur im Allgemeinen zu leisten. Als interdisziplinäres Vorhaben knüpft sie an Ergebnisse, Methoden und Fragestellungen der Game Studies, Software Studies und der Sound Studies an. Der Aufbau orientiert sich an verschiedenen Anknüpfungspunkten der Architektur-Metapher. Nach und nach rücken verschiedene Stufen des Planungs- und Bauprozesses in den Fokus.

Kapitel 1 beginnt mit theoretischen Überlegungen zur Struktur des Computerspiels als einer speziellen Form von Anwendungssoftware und der Rolle akustischer Elemente im Rahmen der Interaktion.

Kapitel 2 skizziert die Entwicklung der akustischen Bausteine von 1971 bis 2011. Im Mittelpunkt stehen dabei technisch bedingte Veränderungen des Rohmaterials und die daraus resultierenden Umbrüche und Kontinuitäten der akustischen Gestaltung.

Kapitel 3 behandelt gängige Praktiken zur zeitlichen Einbindung akustischer Elemente. Ausgehend von der Frage, wie zeitlich ungewisse Abläufe des Spiels vertont werden können, werden ereignisbasierte Gestaltungsverfahren vorgestellt und diskutiert.

Kapitel 4 untersucht analog zu Kapitel 3 die räumlichen Konstruktionsprinzipien der akustischen Gestaltung. Neben den wichtigsten Vertonungsprinzipien der simulierten Räume des Computerspiels

werden verschiedene Spielformen in virtuellen Klangräumen analysiert.

Kapitel 5 thematisiert die Verknüpfung zwischen Filmsound und der Vertonung des Computerspiels. Mit Rückgriff auf die Erkenntnisse der vorangegangenen Kapitel wird ausgelotet, welche kinemato-graphischen Vertonungsstrategien das Computerspiel adaptieren kann und welche Änderungen sie dabei erfahren.

Kapitel 6 diskutiert die Möglichkeiten der Audio-Konfiguration. Sie bildet in der Metaphorik der akustischen Architektur so etwas wie die Inneneinrichtung. Nach der Betrachtung von rahmenden Konfigurationsmöglichkeiten wie Optionsmenüs, in denen die Spieler ihren eigenen Audio-Mix einstellen können, rücken „ingame“-Konfigurationsmöglichkeiten, wie die Wahl eines Radiosenders, in den Fokus.

1 Fundamente

Dass Computerspiele in der Mitte der Gesellschaft angekommen sind, ist nicht zu übersehen. Gespielt wird in der U-Bahn, auf der Party und im Altenheim. Litfaßsäulen und Fernsehwerbespots kündigen den Release des nächsten großen Hits an, und im Bundestag steigt die erste LAN-Session¹. Mit 25 Millionen Gamern in Deutschland² beschäftigt sich knapp ein Drittel der Bevölkerung mit interaktiver Unterhaltungssoftware.

Mit der hohen Verbreitung und Beliebtheit des Computerspiels tritt neben Radiosendungen, Hörspielen und Filmen ein neues massenmedial verbreitetes Format auf den Plan, in dem designer Klang in all seinen Facetten auftritt. Aktuelle Computerspiele enthalten von orchesterlicher Musik und elektronischen Soundtracks über aufwändig produzierte Geräuschkulissen bis hin zu Sprachaufnahmen bekannter Hollywood-Stars alle Formen organisierten Klangs, die unsere Audio-Kultur zu bieten hat. Ein Großteil der Audio-Inhalte liegt dabei in einer Form vor, die medienübergreifend verwertbar ist. Der lizenzierte Game-Soundtrack ist z. B. auch als CD oder digitaler Download erhältlich, während die Geräusche u. a. aus den gleichen Sample-Bibliotheken stammen, mit denen Filme vertont werden. Im Gegensatz zu den ähnlichen Inhalten sind die Zugangsweisen jedoch grundverschieden. Die Spieler begegnen Alice Cooper und Rammstein als „Warriors of Rock“³ mit Plastikgitarren. In Ego-Shootern pirschen sie lauschend durch aufwändig gestaltete Klanglandschaften und in Jump'n'Runs triggern sie tausende Male das gleiche Sprunggeräusch. Die Spieler konsumieren nicht nur vorgefertigte Präsentationen, sondern interagieren mit reaktionsfähigen Systemen.

Die akustische Architektur des Computerspiels ist dementsprechend kein Gefüge, das für sich selbst steht. Sämtliche Audio-Inhalte befinden

¹(Neeb, 2009)

²(BIU, 2011, 2)

³so der Untertitel des sechsten Teils der *Guitar Hero* Serie

sich in einem bestimmten medienspezifischen Kontext, der ihre Einbindungsstrategien, Funktionen und Zugangsweisen prägt. Bevor in den folgenden Kapiteln einzelne Aspekte der akustischen Architektur in den Fokus der Analyse rücken, beschreibt dieses Kapitel zunächst zentrale Charakteristiken dieses Kontextes. Nach der Verortung akustischer Elemente im technischen Kontext des Programms und im inhaltlichen Kontext von Spiel und Erzählung folgen Überlegungen und Beobachtungen darüber, wie die Spieler mit den für sie aufbereiteten Angeboten umgehen.

1.1 Sound im Kontext von Software

Computerspiele sind eine spezielle Form von Anwendungssoftware. In einigen Punkten teilen sie Gemeinsamkeiten mit benachbarten Anwendungsprogrammen, während sie in anderen Punkten eigene Charakteristika aufweisen. Die auffälligste Eigenschaft von Computerspielen besteht darin, dass sie nicht dafür ausgelegt sind, uns in konkreten Problemstellungen des Alltags und der Arbeitswelt zu unterstützen. Sie weisen uns nicht den Weg wie ein Navigationssystem und geben uns keine Mittel an die Hand, in kürzester Zeit einen Serienbrief an mehrere tausend Empfänger aufzusetzen. Statt dessen stellen sie freie, fiktive Entwürfe dar, die Erfahrungen schaffen. Die Bandbreite der Erfahrungen reicht dabei von der Unterhaltung über das Lernen bis hin zu Kontemplation und Reflexion. Computerspiele sind damit für Software gewissermaßen das, was fiktionale Literatur für die Schrift ist. Verglichen mit anderer Anwendungssoftware binden sie die Spieler durch ihre Ausgaben und die angebotenen Eingabemöglichkeiten weitaus stärker ein. Die Eingaben fordern die Spieler heraus und verlangen eine aktive Teilnahme. Die Ausgaben präsentieren aufwändig gestaltete Darstellungen, die an unsere Sinne und unser ästhetisches Empfinden gerichtet sind.⁴ Akustische Elemente sind wichtige Bestandteile dieser Ausgaben.⁵

⁴Durch diese starke Gewichtung der Darstellung stellen Computerspiele hohe Anforderungen an die Ressourcen des Computers. Die Auslotung der (vor allem grafischen) Darstellungsmöglichkeiten bis an ihre technischen Grenzen ist ein Grund dafür, warum Computerspiele als Motor der Hardware-Entwicklung gelten.

⁵Mittlerweile registrieren einige Titel auch akustische Eingaben, z. B. in Form von Sprachbefehlen. Sie sind jedoch (noch) die Ausnahmen der Regel. Im Folgenden

Während Sound in den meisten Formen von Anwendungssoftware eine untergeordnete optionale Rolle spielt, gelten akustische Elemente innerhalb von Computerspielen als integrale Bestandteile der Präsentation. An Stelle von wenigen akustischen Signalen, die Fehlermeldungen und abgeschlossene Prozesse begleiten, enthalten Computerspiele Geräusche, Musik und gesprochene Sprache, die in mehreren sorgfältig aufeinander abgestimmten Klangschichten organisiert sind. Neben akustischen Elementen greifen Computerspiele auf diverse weitere Darstellungselemente zurück. Sie präsentieren ihre Inhalte auch textuell, grafisch, haptisch, oder als multimediale Mischformen.

Ausführbare Baupläne

Ein existenzielles Merkmal jedes Computerprogramms ist seine Ausführbarkeit. Anstatt etwas vorher Aufgezeichnetes wiederzugeben, erzeugt es etwas durch die schrittweise Ausführung von Anweisungen. Der ausführbare Code ist wie ein Bauplan, nach dessen Anleitungen das Programm sämtliche vorformulierten Formen und Funktionen errichtet. Das Gebaute bleibt jedoch nicht stehen, sondern verschwindet wieder, sobald das Programm seine Ausführung beendet. Es tritt nur in seiner Prozesshaftigkeit in Erscheinung. Ausgeführte Programme ähneln damit der Aufführung einer Partitur. So bemerkt der Informatiker Donald Knuth:

My feeling is that when we prepare a program, it can be like composing poetry or music;⁶

Im Gegensatz zu einem Orchester, bei dem sich in Abhängigkeit von der Tagesform jede Aufführung voneinander unterscheidet, ermöglicht die Ausführung von Programmen exakt reproduzierbare Ergebnisse. Sobald jedoch wie bei Computerspielen Benutzer und Zufallsprozesse hinzutreten, entsteht ein variierender Verlauf, der sich erst während der Aufführung entscheidet. Lev Manovich beschreibt das dadurch Erfahrbare als *software performance*:

Instead of fixed documents whose contents and meaning could be full [sic] determined by examining their structure

beziehen sich akustische Elemente, wenn nicht anderweitig gekennzeichnet, stets auf die Ausgabe.

⁶(Knuth, 1974, 670)

(which is what the majority of twentieth century theories of culture were doing) we now interact with dynamic “software performances.” I use the word “performance” because what we are experiencing is constructed by software in real time.⁷

Der Konstruktionscharakter und die hohe Variabilität stellt die kritische Auseinandersetzung mit Software vor große Herausforderungen. Die ausgegebenen Darstellungen, zu denen akustische Elemente ihren Teil beitragen, besitzen keine fixierten Formen, die analysiert werden könnten. Sie gehen erst während der Laufzeit des Programms aus den im Code vorgegebenen Kombinationsmöglichkeiten eingebundener Darstellungselemente sowie den Eingaben der Benutzer hervor. Genau genommen handelt es sich nicht um *eine* Form, sondern eine Vielzahl von Formmöglichkeiten. Um die Ausdrucksfähigkeit des Computerspiels und die Rolle akustischer Elemente in den Blick zu nehmen, ist das Herausgreifen einzelner gebauter Formen deshalb nicht ausreichend. Stattdessen muss neben den Bauten auch der gesamte Konstruktionsvorgang berücksichtigt werden, um die sich öffnenden Möglichkeitsräume fassen zu können⁸.

Um Einblicke in den Konstruktionsvorgang zu erhalten, sind mehrere Vorgehensweisen möglich. Eine davon besteht in der Untersuchung der zu Grunde liegenden Plattformen.⁹ Die Plattform bezeichnet von der Hardware über das Betriebssystem bis hin zu Engines für spezielle Anwendungsbereiche den gesamten technischen Unterbau des Systems. Sie legt die Grenzen und Möglichkeiten der in ihrem Rahmen erstell- und ausführbaren Programme fest, die von Entwicklern und Benutzern ausgelotet werden.¹⁰ Die beliebteste Spieleplattform in Deutschland ist der PC¹¹, dem aus diesem Grund auch in dieser Arbeit die größte

⁷(Manovich, 2008, 15)

⁸Erschwerend kommt hinzu, dass nicht nur das Gebaute je nach Durchlauf unterschiedliche Formen annimmt, sondern auch die Baupläne einem stetigen Wandel unterliegen. Programme liegen selten in einer finalen Fassung vor, sondern durchlaufen mehrere Versionen. Wenn ein Programm durch Updates oder Mods keine Änderungen mehr erfährt, wird es in den meisten Fällen auch nicht mehr benutzt. Siehe auch (Mackenzie, 2006, 12).

⁹Die Reihe Platform Studies der MIT Press beschreitet diesen Weg, indem sie sich aus einer kulturwissenschaftlichen Perspektive mit populären Plattformen wie dem Atari VCS, der Nintendo Wii und dem Commodore Amiga auseinandersetzt.

¹⁰Interessanterweise gibt es kaum eine Plattform für Endanwender, die keine Spiele enthält. Selbst der ebook-Reader Kindle startet mit der Tastenkombination Shift+Alt+M das Spiel *Minesweeper*.

¹¹(BIU, 2011, 4)

Aufmerksamkeit zukommt. Die Betrachtungen technischer Hardware-Spezifikationen und Analysen des Quellcodes sind hilfreich, um die während der Programmausführung vollzogenen Konstruktionsprozesse zu verstehen. Für sich genommen lassen sie jedoch nur geringe Rückschlüsse auf das durch Software Erfahrbare zu. Innerhalb der Game Studies herrscht deshalb Einigkeit darüber, dass die aktive Nutzung des Computerspiels eine unerlässliche Zugriffsweise auf den Untersuchungsgegenstand darstellt.

Ein weiterer Ansatz besteht darin, den Konstruktionsvorgang in einzelne Bestandteile zu zerlegen. Noah Wardrip-Fruin beschreibt ihn dafür als *expressive processing*.¹² Nicht nur Darstellungselemente, sondern auch die zugrunde liegenden Prozesse rücken dabei als expressive Strukturen und Inhalte in den Fokus. Einige davon basieren auf den Prinzipien traditioneller Ausdrucksformen wie Malerei, Literatur und Film. Andere, wie Künstliche Intelligenz, sind genuine Erzeugnisse moderner Informationstechnik. Um diese ausdrucksvollen Verarbeitungsprozesse zu berücksichtigen, segmentiert Wardrip-Fruin das Programm in zwei Bereiche. Die *Oberfläche* besteht demnach in Anwendungsprogrammen aus den speziell für den Menschen entworfenen Komponenten. Sie umfasst die gesamte Ausgabe und alle Eingabemöglichkeiten. Die *Tiefenebene* hingegen besteht aus jenen Teilen des Programms, die nicht unmittelbar für die sinnliche Wahrnehmung vorgesehen sind. Sie enthält u. a. jene Daten und Prozesse, die für die Konstruktion der Oberfläche verantwortlich sind. Diese Segmentierung des Programms in eine Oberfläche und einen tiefer liegenden, nicht unmittelbar einsehbaren Kern ist auch für die Untersuchung des akustischen Anteils von Computerspielen hilfreich, da er genau so wie die von Wardrip-Fruin untersuchten textuellen Darstellungsartefakte aus Prozessen entsteht, die auf Daten zurückgreifen. Anstelle von Textbausteinen bestehen diese akustischen Daten aus aufgenommenen Klangsnipseln oder akustischen Parametern wie Lautstärken und Frequenzen.

Eine ähnliche Aufteilung des Programms in eine für die Benutzer vorgesehene und eine darunter liegende Schicht ist auch im Rahmen der Software-Entwicklung üblich. An Stelle von einer Oberfläche und einer Tiefenebene spricht man von einem *Frontend* und einem *Backend*. Das Frontend bezeichnet in der Regel eine höher liegende Kommunikationsschicht, während das Backend die dahinter verborgene Arbeitsschicht

¹²(Wardrip-Fruin, 2009)

enthält. Für Anwender ist diese Zweiteilung z. B. bei Schachprogrammen sichtbar. Das derzeitige Angebot umfasst mehrere Frontends und Backends, bzw. *Engines*, die dank einheitlicher Kommunikationsprotokolle beliebig miteinander kombinierbar sind.¹³ Das Frontend umfasst die Darstellung des Spielbretts und der Schachfiguren, während die Engine als Backend die Spiellogik und die programmierten Gegenspieler bereit stellt.

Mensch-Computer-Interaktion

Neben der Ausführbarkeit statt der Wiedergabe gilt die Interaktivität als definierendes Merkmal des Computerspiels, das es von anderen Medien wie dem Film abgrenzt. Der Game Designer Chris Crawford definiert Interaktivität als „a cyclic process in which two actors alternately listen, think, and speak.“¹⁴ Das Sprechen des einen Agenten geschieht zeitgleich mit dem Hören des anderen Agenten. Bei der Mensch-Computer-Interaktion handelt es sich offenbar um zwei ungleiche Kommunikationsteilnehmer, deren Hören, Denken, und Sprechen sich stark voneinander unterscheiden. Für den Computer sind diese drei Schritte genauestens festgelegt. Sie sind die Ergebnisse komplexer Design-Prozesse, die von der Registrierung der Eingabe (dem Hören) über die Verarbeitung der Eingabe im Rahmen eines Rechenmodells (dem Denken) bis hin zur Ausgabe (dem Sprechen) als ausführbare Prozesse und abrufbare Datensätze realisiert sind. Für die Benutzer bestehen die drei Schritte aus der Wahrnehmung der Ausgabe, deren Interpretation und der Eingabe ihrer Entscheidungen. Sie können sich ihren Teil denken, müssen aber immer diejenige Sprache sprechen, die das Programm versteht. Anstatt frei zu formulieren, wählen sie bereits vordefinierte Antworten.

Jeder Akt des metaphorischen Sprechens und Hörens passiert die Barriere der Benutzerschnittstelle. Sie lässt sich in zwei Bereiche aufteilen. Die Hardware-Ebene besteht aus den Ein- und Ausgabegeräten

¹³Beispiele für Schach-Frontends sind Arena (<http://www.playwitharena.com>) und XBoard (<http://www.gnu.org/software/xboard/>). Beliebte Schach-Engines sind Crafty (<http://www.craftychess.com>), Fruit (<http://www.fruit chess.com>), Houdini (<http://www.cruxis.com/chess/houdini.htm>) und Stockfish (<http://stockfishchess.org>).

¹⁴(Crawford, 2002, 5)

wie Maus, Tastatur, Gamepad, Bildschirm, Kopfhörer, Kameras, Mikrofonen und Sensoren. Die Software-Ebene umfasst jene symbolische Schicht, die den Benutzern über die Ein- und Ausgabegeräte die Interaktion mit dem Programm ermöglicht.

Welchen Stellenwert haben nun die Darstellung und insbesondere akustische Elemente als Teil der Darstellung innerhalb dieses Interaktionsgeflechts? In den Game Studies herrscht Einigkeit darüber, dass nicht die Darstellung, sondern die Teilnahme und das daraus resultierende Gameplay den Kern des Computerspiels bildet.¹⁵ Alexander Galloway bezeichnet es deshalb als „*action-based* medium“¹⁶ und Crawford schreibt:

Graphics, animation, sound, and music are all necessary to gaming, and they're all important, but they're not the schwerpunkt. Interactivity (sometimes called „gameplay“) is the real schwerpunkt of games.¹⁷

Janet Murray hebt diesen Aspekt gerade als zentrales Merkmal hervor, das den Computer vom Apparat des Films unterscheidet.

Just as the primary representational property of the movie camera and projector is the photographic rendering of action over time, the primary representational property of the computer is the codified rendering of responsive behaviours. This is what is most often meant when we say that computers are *interactive*. We mean they create an environment that is both procedural and participatory.¹⁸

Im Rahmen dieser Überlegungen nehmen die grafischen und akustischen Darstellungen eine untergeordnete Position ein. Sie stellen lediglich einen Teilaspekt des oben beschriebenen Interaktionszyklus dar. Die Beurteilung akustischer Elemente driftet in der Theoriebildung und der praktischen Erfahrung jedoch weit auseinander. In der Theorie sind akustische Elemente die Nebensächlichkeit einer Nebensächlichkeit. Einerseits fällt die Darstellung hinter den Aspekt der Teilnahme zurück, den Game Designer wie Wissenschaftler im Kontext des Gameplays stärker fokussieren. Akustische Elemente als Teil der Ausgabe rücken dadurch an die Peripherie. Andererseits tritt das Hörbare

¹⁵siehe z. B. (Mäyrä, 2009, 317)

¹⁶(Galloway, 2006, 3)

¹⁷(Crawford, 2003, 74)

¹⁸(Murray, 1997, 74)

in unserer visuell ausgerichteten Kultur hinter das Sichtbare zurück. Schon während der Konzeptionsphase vieler Spiele erfährt die akustische Präsentation nur geringe Aufmerksamkeit. In der Praxis hingegen sind es in hohem Maße akustische Elemente, die unsere Spielerfahrung prägen. Wenn Spieler über ihre zurückliegenden Spielerfahrungen berichten, nehmen akustische Elemente in ihren Schilderungen häufig eine besondere Rolle ein. Sie erinnern sich an Melodien, Soundeffekte und Sprachsamples. Akustische Elemente reichen in hohem Maße über das Computerspiel hinaus. So setzen sich die Fans von *Portal* z.B. mit dem Song des Nachspanns intensiv auseinander. Ob Rock, a capella, 8-Bit Version, Jazz-Cover, Electro Remix, oder wiedergegeben durch einen Faserlaser - auf Portalen wie youtube verarbeiten die Spieler „Still Alive“ von Jonathan Coulton in vielfältigen Formen weiter. Die Musik vieler Computerspiele geht auf diese Weise in das Repertoire der musizierenden Netzgemeinschaft über.

Im Kontext der interaktiven Spielsituation ist Sound anders gestaltet als in linearen Präsentationsformen wie in einer Filmsequenz. Das Spiel stellt Eingriffsmöglichkeiten bereit, die den Spielern über akustische und grafische Elemente vermittelt werden. Es treten nicht nur die Eingabegeräte hinzu. Auch ein Teil der Ausgabe konstituiert das Interface und ist ganz auf die Teilnahme, die Interaktion der Spieler mit dem Spielsystem, ausgerichtet. Um diesen Aspekt in der Beschreibung akustischer Arrangements zu berücksichtigen, müssen zwei zentrale Prinzipien des interaktionsbasierten Designs herangezogen werden: Feedback und Suggestion. Unter dem Feedback der Benutzerschnittstelle versteht man jegliche Formen der Rückmeldung auf vorangegangene Eingaben. Suggestive Darstellungen hingegen sind Wegweiser, die die Benutzer auf etwas hinweisen und bestimmte (Re-)Aktionen nahe legen.

Feedback

Akustisches Feedback ist das am weitesten verbreitetste Verτονungsprinzip in Anwendungsprogrammen. Ähnlich wie in einem Gespräch, in dem das Gegenüber durch kurze akustische Rückmeldungen seine bleibende Aufmerksamkeit signalisiert, bestärkt das akustische Feedback den Dialog zwischen Programm und Benutzer. Es bildet einen konstitutiven Bestandteil der Mensch-Computer-Interaktion. Speziell in Computerspielen bestätigt das Feedback nicht nur die Eingabe, son-

dern attestiert darüber hinaus eine Handlung in einer fiktiven Welt. In der Realität begleiten Geräusche unentwegt unsere Tätigkeiten. Die Integration akustischen Feedbacks innerhalb des Computerspiels trägt diesem Umstand Rechnung und lässt die symbolischen Manipulationen innerhalb des virtuellen Raums weitaus natürlicher erscheinen, als sie es eigentlich sind.

Akustisches Feedback lässt sich bezogen auf den oben skizzierten Interaktionszyklus in zwei Varianten unterteilen:

- Rückmeldungen über durchgeführte Aktionen der Benutzer und
- Rückmeldungen über Operationen des Systems.

Beide Varianten bestätigen etwas, das gerade geschieht oder vor kurzem geschehen ist. So besitzen nahezu alle Aktionen, die die Benutzer in einem Computerspiel auslösen, ein unmittelbares akustisches Feedback. Bewegen sie einen Avatar, erklingen Fußschritte. Drücken sie während der Steuerung eines Fahrzeugs auf das Gaspedal, steigt die Frequenz des Motorgeräuschs. Geben sie in einem Strategiespiel Befehle, quittieren ihre Einheiten diese mit passenden Sprachsamples. Verzögerungsfreie Rückmeldungen dieser Art signalisieren den Benutzern, dass ihre Eingaben angekommen sind. Darüber hinaus vermitteln sie Informationen über den Status durchgeführter Aktionen. Bei einem gut vertonten Spiel hören wir, ob eine Aktion erfolgreich war oder fehlgeschlagen ist. Unterschiedliche akustische Signale sonifizieren somit den Ausgang der Aktionen.

Rückmeldungen über die Operationen des Systems spiegeln die Resultate interner Berechnungen und Zustandsübergänge wider. Beispiele aus der Domäne des Anwendungsprogramms sind Bestätigungen nach einem abgeschlossenen Kopier- oder Speichervorgang. Analog dazu geben Computerspiele Rückmeldungen, wenn z. B. ein neues Gebäude fertig gestellt wurde, der Cooldown einer zuvor benutzten Spezialfähigkeit abgeklungen ist, oder neue Gegnergruppen spawnen. Im Gegensatz zu dem direkten Feedback ausgeführter Aktionen stehen die Rückmeldungen der System-Operationen nicht zwangsweise in unmittelbarem Zusammenhang mit Eingaben der Spieler.

Obwohl beide Varianten des Feedbacks zur Vertonung diverser Ereignisse und Aktionen zur Anwendung kommen, spielen zwei archetypische Auslöser eine besondere Rolle: Berührungen und Kollisionen.

Berührung Die Berührung dargestellter Objekte innerhalb des virtuellen Raums durch die Benutzer stellt eine wichtige Kategorie des Feedbacks dar. Sie nimmt vor allem seit dem Aufkommen grafischer Benutzeroberflächen und der Maussteuerung eine zentrale Position in der Mensch-Computer-Interaktion ein. Die gesamte Desktop-Metaphorik greift mit ihren Papierkörben, *Files* und *Folders* auf materielle Gegenstände zurück, die im virtuellen Raum durch Operationen wie *Drag and Drop* angepackt und bewegt werden können. Selbst geschriebener Text, der in den Kommandozeilen früherer Benutzerschnittstellen sämtliche Befehle an die Maschine transportierte, mutiert in den Menüs grafischer Benutzeroberflächen als anklickbares Objekt zu etwas Berührbarem.

In Computerspielen ist der Umgang mit Objekten, die materiellen Gegenständen nachempfunden sind, noch weitaus stärker ausgeprägt als in anderen Anwendungsprogrammen. Die Manipulation dieser Objekte geschieht durch Berührungen, die einen Großteil der den Spielern zur Verfügung stehenden Aktionen ausmachen. Im Fall der Avatar-Steuerung durchläuft die Berührung eine Zwischeninstanz. Um Gegenstände der fiktiven Spielwelt fassen zu können, ist eine räumliche Nähe zwischen den Avataren und den Objekten nötig. Häufig vollzieht sich das Berühren jedoch direkter durch Zeigen und Bestätigen, vor allem bei der Maussteuerung oder der Interaktion mit einem Touchscreen. Die Mauszeiger des Computerspiels bestehen nicht nur aus Pfeilen, sondern zeigen Hände und Werkzeuge wie z. B. einen Hammer zum Bauen eines Gebäudes oder einen Schild, um eine Verteidigungsposition einzunehmen.

Ein Großteil unserer Sinneswahrnehmungen bleibt außen vor, wenn wir auf eine grafische Repräsentation mit der Maus klicken oder über die Kinect-Kamera ein Menü mit Hilfe von Gesten ansteuern. Besonders deutlich wird dieser Verlust der sinnlichen Wahrnehmungsmöglichkeiten beim Vergleich der Berührung virtueller Objekte mit der Berührung von Spielsteinen auf einem Brettspiel. Man kann ein virtuelles Objekt nicht in die Hand nehmen und wie bei dem Spielstein fühlen, ob er aus Plastik oder Holz ist. Bei der Übertragung in den virtuellen Raum wird aus dem greifbaren Gegenstand zunächst etwas Bildhaftes, das sich uns nur über den Bildschirm – sei es als einfaches Symbol oder rotierbares 3D-Modell – erschließt. Bisher sind unsere Interfaces noch nicht in der Lage, taktile Sinnesempfindungen zu stimulieren, die mit den berührten Gegenständen korrespondieren. Zwar bestätigen einige Ausgabegeräte

Berührungen mittlerweile durch Vibrationen, bleiben damit aber auf einen Bruchteil der Fähigkeiten unseres Tastsinns beschränkt.

Statt dessen sind es akustische Informationen, die nicht nur ein Feedback über den Status der Berührung, sondern auch über die materiellen Eigenschaften des berührten Gegenstands vermitteln. Sound kompensiert die fehlende Einbindung des Tastsinns. Die Unmittelbarkeit direkter akustischer Rückmeldungen verringert weiterhin die Distanz zwischen Eingabe und Ausgabe: in unserer Wahrnehmung fallen beide zusammen. Die Latenz der ausgelösten Geräusche ist so gering, dass wir das Gefühl haben, sie erklingen gleichzeitig mit dem Mausklick oder Tastendruck. In besonders verdichteter Form lässt sich diese Kompensation fehlender Haptik in Inventar-Menüs beobachten. Immer dann, wenn die Spieler in Computerspielen wie *Diablo* einen Gegenstand in der Inventarfläche platzieren, erklingt ein Geräusch, das diesen Gegenstand charakterisiert. Bei der Kettenrüstung poltert Metall, während Edelsteine von einem hochfrequenten Klimpern begleitet werden.

In Computerspielen ist das Feedback in den häufigsten Fällen keine nüchterne Rückmeldung, sondern stattdessen ein effektiv inszeniertes Spektakel. Während die Geräusche von Berührungen bereits stark überzeichnet sind, stehen die Vertonungen von Kollisionen ganz im Zeichen der Hyperrealität medial inszenierter Klangkulissen.

Kollision

In the world of video games, objects on the screen usually interact only when they run into each other.¹⁹

Sobald zwei Dinge in der Realität aufeinander prallen, äußert sich der daraus folgende Energieverlust u. a. in der Emission von Schallwellen. Sei es der schmale Absatz eines Stöckelschuhs, der mit lautem Klacken auf Asphalt trifft, oder ein fallen gelassenes Glas, das auf Stein zersplittert - Kollisionen erzeugen auffällige Schallsignale hoher Lautstärke, die Aufmerksamkeit erregen. Als spektakuläre physikalische Ereignisse markieren sie Momente der Veränderung, die insbesondere in kompetitiven Situationen tiefgreifende Konsequenzen nach sich ziehen. Füße, Hände und Schläger schmettern gegen Bälle in der Hoffnung auf einen Punktgewinn, Karosserien verformen sich im Crash zweier

¹⁹(Robinett, 2006, 679)

Fahrzeuge, und Projektile zerpfücken in die Luft geschleuderte Tontauben. Insofern ist es nicht verwunderlich, dass Kollisionsereignisse auch in Computerspielen, die in Echtzeit ablaufen, eine zentrale Rolle einnehmen. Die Herbeiführung und Vermeidung von Kollisionen stellt in Action-Spielen das zentrale Mittel dar, um übergeordnete Spielziele einer größeren Agenda zu erreichen. Sie gehören zu den konstitutiven Bestandteilen des Spielvorgangs.

Prinzipiell können sich in den simulierten Räumen von Computerspielen mehrere Dinge an demselben Ort befinden, ohne dass sie aufeinander einwirken. Bereits die Berührung zweier Objekte auf dem Bildschirm ist ein konstruiertes Ereignis, das erst aus kontinuierlichen Überprüfungen ihrer Umrisse auf gegenseitige Überschneidungen hervorgeht. Man nennt diese Problemstellung, mit der sich in der Informatik die Disziplin Algorithmische Geometrie befasst, *Kollisionserkennung*. Die Kollisionserkennung ist eine fundamentale Voraussetzung für grafische Oberflächen jeder Art. Schon bei jeder Überlagerung des Mauszeigers mit einem Icon findet im Hintergrund eine Kollisionserkennung statt, die diese Überlagerung registriert. In Hinblick auf die akustische Gestaltung schlägt die Kollisionserkennung eine Brücke zwischen Raum und Zeit. Das Resultat der räumlichen Berechnungen ist nicht nur ein positiver Wahrheitswert und ggf. der Ort der Überschneidung, sondern auch ein Zeitpunkt. Er ist zu Beginn des Spiels unbekannt und markiert den Moment, in dem eine Reaktion als Folge der Berührung, die sog. *Kollisionsantwort*, stattfindet.

Kollisionsantworten bestehen sowohl aus Operationen, die im Hintergrund ablaufen, als auch Rückmeldungen für die Spieler in Form einer Ausgabe. In den meisten Fällen sind es audiovisuelle Blöcke, die den Spielern die Kollisionsvorgänge vermitteln. Die Kollision eines Projektils mit einem Raumschiff resultiert typischerweise in der Verschmelzung beider Objekte in einer mit Rauschen und Knallen unterlegten Explosionsanimation. Weder die visuelle noch die akustische Repräsentation sind jedoch unabdingbar für die Kollisionsantwort. So können zwei Objekte ohne weiteres außerhalb des sichtbaren Bildschirmabschnitts hörbar zusammenprallen und innerhalb der sichtbaren Szenerie lautlos ineinanderrauschen.

Suggestion

Ein zweites Prinzip, das die gesamte akustische Ausgestaltung durchdringt, ist die Suggestion. Auffallende Geräusche können ebenso wie blinkende visuelle Komponenten auf etwas hinweisen, so dass sich die Spieler dann einem Objekt oder einem Ort in der virtuellen Welt zuwenden. Suggestionen zielen darauf, die Entscheidungsfindung der Spieler zu beeinflussen. Sie sind bewusst eingesetzte Fingerzeige der Entwickler, die die Aufmerksamkeit der Spieler auf bestimmte Gegenstände, Zeitpunkte, oder Orte lenken. So kann beispielsweise das Geräusch einer sich öffnenden oder schließenden Tür im Rücken der Spieler als Suggestion fungieren, die sie dazu verleitet, sich umzudrehen und nach den Ursachen zu forschen.

Die suggestive Gestaltung beginnt in der Auswahl der zu vertonenden Ereignisse und Zustände und endet in gezielt platzierten Signalen und sorgfältig austarierten Abmischungen. Im Gegensatz zum Feedback, das in jeder Form von Anwendungssoftware ein elementarer Bestandteil ist, sind suggestive Gestaltungsstrategien speziell in Computerspielen stärker vertreten. Das liegt vor allem daran, dass in ihren regelbasierten und inszenierten Umgebungen Spielsituationen vorhersehbar sind. In einem Textverarbeitungsprogramm ist es hingegen schwer möglich, die nächsten Schritte der Benutzer vorherzusehen. Es ist unklar, ob sie als nächstes Text eingeben, an der Formatierung arbeiten, oder eine Passage kommentieren. Gemeinsam mit grafischen Suggestionen wie dem Highlighting tragen akustische Suggestionen maßgeblich zu einem besseren Spielfluß bei. Sie ermöglichen die zielgerichtete Spielerführung in interaktiven Umgebungen.

Suggestionen können als subtile Akzente oder deutlicher hervortretende Signale²⁰ realisiert sein und legen notwendige oder optionale Handlungen nahe. Während Feedback jene Dinge unterstreicht, die gerade passieren oder bereits geschehen sind, zeigen Suggestionen das, was passieren wird. In *Superbrothers: Sword and Sorcery EP* erklingt z. B. beim Kampf mit dem Trigon kurz vor jeder Attacke ein Warnton, der

²⁰Die stärksten Suggestionen bestehen in direkten Manipulationen der Eingabemöglichkeiten. In *Amnesia: The Dark Descent* werden in manchen Räumen die Blicke in eine bestimmte Richtung gelenkt und die Spieler damit aufgefordert, jene Winkel der Spielwelt genauer zu untersuchen. Wenn sie in solchen Momenten mit der Maus die Blickrichtung verändern, lenkt das Programm sie wieder zurück, bis die Spieler das Gezeigte gefunden haben.

die Spieler dabei unterstützt, im richtigen Moment den Schild zu aktivieren. Der Ton ermöglicht den Spielern einen Blick in die Zukunft. Ganz ähnlich funktionieren grafische Suggestionen: In *Batman Arkham Asylum* und der *Assassin's Creed* Reihe erscheinen während der Scharmützel mit Gegnergruppen Markierungen über den Gegnern, die als nächstes einen Angriff starten. Die Spieler erhalten dadurch die Gelegenheit, bereits auf etwas zu reagieren, das erst kurze Zeit später geschieht. Sie können ausweichen oder einen Konter einleiten. Batman und Ezio Auditore da Firenze werden dadurch übermächtig und nehmen es problemlos mit gewaltigen Horden gegnerischer NPCs²¹ auf. Die Hervorhebung vermittelt den Spielern ein Wissen, das sie ansonsten nur äußert schwer und zu einem späteren Zeitpunkt aus der Betrachtung der Szene hätten ableiten können, nämlich das Ergebnis eines Algorithmus zur Kalkulation des nächsten Angreifers.

Eine besonders ausgeklügelte Form suggestiver akustischer Signale kommt in dem 2008 erschienenen Coop-Shooter *Left 4 Dead* zur Anwendung, in dem nahezu jedes musikalische Motiv spielrelevante, suggestive Zusatzinformationen transportiert. Einige dieser *audio silhouettes*, wie die Entwickler sie nennen²², verweisen auf die Anwesenheit spezieller Infizierter, die für die Spieler eine große Gefahr darstellen. Andere kündigen künftige Ereignisse wie das Eintreffen einer Horde an. In beiden Fällen verschaffen die Signale den Spielern Zeit, um sich auf die jeweilige Bedrohung einzustellen. Kommentaren der Entwickler zufolge zielen die *audio silhouettes* darauf, den Spielern auf einer unterbewußten Ebene eine größere Klarheit über die eintretenden Ereignisse zu verschaffen.

In playtests, people were often confused by certain events and elements in the game. The addition of musical cues helped distinguish these events and diminish player confusion. This was similar to the way that we experimented with distinct visual silhouettes in order to make our characters stand out in a chaotic environment.²³

²¹ *Non Player Character* (NPC), zu deutsch Nicht-Spieler-Charaktere, sind computergesteuerte Akteure.

²² Das Spiel bietet ähnlich wie die Kommentarspur auf einer DVD einen Modus, in dem die Entwickler Design-Entscheidungen kommentieren. Innerhalb der Maps sind mehrere Kommentar-Knoten verteilt, die von den Spielern aktiviert werden können. Alle folgenden Verweise auf „*Left 4 Dead Developers Commentary*“ beziehen sich auf Inhalte dieser interaktiven Kommentierung.

²³ Tobin Buttram, *Left 4 Dead Developers Commentary*

Exkurs: Highlighting Die offensichtlichste Form einer suggestiven Darstellung sind Hervorhebungen. In der grafischen Darstellung zeigen leuchtende Konturen, blinkende Animationen und weitere Formen der Farbgebung und Kontrastsetzung den Spielern die Dinge, denen sie ihre Aufmerksamkeit widmen sollten. So gilt in Adventures und Action-RPGs wie *Diablo* und *Torchlight* die häufig optional zuschaltbare Hervorhebung wichtiger Spielobjekte – im Spielerjargon als *Highlighting* bezeichnet – als Standard. Sie unterstützt die Spieler massiv in der Unterscheidung von relevanten und irrelevanten Objekten. Das Highlighting zeigt ihnen alles Benutzbare im jeweiligen Bildschirmausschnitt. Um das Auffinden von bestimmten Gegenständen stärker zu betonen, kommen gelegentlich auch Vibrationen und Klänge zum Einsatz, sobald sich Mauszeiger und Gegenstand überlagern. In einigen Fällen geben unterschiedliche Farbgebungen der Umrandungen genauere Auskunft über die Eigenschaften des jeweiligen Objekts. In *Diablo* besitzen z. B. alle blau leuchtenden Items magische Eigenschaften. Farbliche Umrandungen sind nicht auf Gegenstände beschränkt. Die *Left 4 Dead* Reihe verwendet ein Hervorhebungssystem für die Avatare der Spieler. Die Silhouetten aller Teammitglieder sind das gesamte Spiel über auch durch Hindernisse wie Wände sichtbar, was die Zusammenarbeit in dem auf Kooperation ausgelegten Shooter stark begünstigt. Weiterhin spiegeln die Farben der Silhouetten, von grün bis rot, den Gesundheitszustand des jeweiligen Avatars wider.

Einige Spiele integrieren grafische Hervorhebungen nahtlos in die Spielwelt und verzichten damit auf Kontur-Highlighting. Das Spiel *Mirror's Edge* enthält ein mit der Spielwelt verschmolzenes Navigationssystem. Oberflächentexturen, welche in die richtige Richtung weisen, färben sich beim Näherkommen des Avatars rot oder blau. Diese eingeblendeten Kontraste in der Farbgebung fördern einen schnellen, flüssigen Spielablauf. Sie ermöglichen den Spielern die schnelle Orientierung innerhalb der dreidimensionalen Spielwelt, ohne die sie die hektischen Fluchtsequenzen des Spiels kaum überstehen könnten. Ein weiteres Spiel, in dem die farbliche Kodierung für den Spielablauf eine sehr wichtige Rolle spielt, ist die Independent-Produktion *Osmos* von Hemisphere Games. Der Spieler steuert einen kreisförmigen Organismus, der durch die Absorption kleinerer Organismen des virtuellen Mikrokosmos – sogenannter Motes – wachsen muß. Alle Organismen, die der Spieler absorbieren kann, sind blau koloriert, während größere Organismen, die ihn gefährden, eine rote Färbung aufweisen. Bei jeder Größenveränderung des

Spieler-Organismus passen alle Motes ihre Farbe an, um diesem Kriterium zu genügen. Die Farbe jedes computergesteuerten Motes ist also das Ergebnis eines Größenvergleichs seines Durchmessers mit dem Durchmesser des spielergesteuerten Organismus. Diese visuelle Hilfestellung unterstützt die Spieler stark bei der Größen-Einschätzung benachbarter Motes, die eine ähnliche Größe aufweisen.

Grafische Umrandungen treten auch dann auf, wenn die Spieler eine Auswahl treffen. Ob es sich dabei um einen Menüpunkt oder Spielfiguren in Strategiespielen handelt, ist unerheblich. Ausgewählte Optionen und Einheiten werden üblicherweise mit leuchtenden Konturen hervorgehoben und der Vorgang des Auswählens durch ein akustisches Sample bestätigt. Hervorhebungen dieser Art trennen Aktives von Inaktivem, markieren Berührtes und Berührbares. Sie gehören nicht zur Darstellung der fiktiven Welt, sondern überlagern sie vielmehr als eine zusätzliche Schicht, die eigens dazu dient, dem interagierenden Spieler zu helfen. Akustische Hervorhebungen sind im Gegensatz dazu meistens subtiler.

Narrativ gerechtfertigte Suggestionen

Die Spieler akzeptieren, dass Gegenstände durch Wände sichtbar sind und akustische Signale Informationen liefern, welche die Grenzen der Fiktion überschreiten. Nichtsdestotrotz kaschieren viele Computerspiele Teile dieser Interface-Schicht, indem sie narrativ gerechtfertigt wird. In *The Witcher 2* trägt der vom Spieler gesteuerte Hexer Geralt z. B. ein Medaillon, durch dessen Aktivierung alle Gegenstände in einem gewissen Radius um ihn herum hervorgehoben werden. Darüber hinaus gibt das Medaillon in Gegenwart von Magie ein akustisches Signal von sich. Aktivieren die Spieler es in so einem Moment, finden sie Kreise der Magie, die Geralt für eine Zeit lang verschiedenen Effekten aussetzen. Das Signal markiert die Nähe zu etwas Unsichtbarem, das die Spieler mit Hilfe des Medaillons sichtbar machen.

In der *Assassin's Creed* Reihe wird nicht nur ein einzelnes Element, sondern das gesamte Interface durch die Rahmenhandlung erklärt. Sie spielt in der nahen Zukunft, wo der vom Spieler gesteuerte Desmond Miles mit Hilfe einer Maschine die genetisch in ihm gespeicherten Erinnerungen seiner mittelalterlichen Ahnen nacherlebt. Indem die Spielfigur Miles nun über ein hochentwickeltes Neuro-Interface Zugang in eine

virtuelle mittelalterliche Welt findet, verfügt sie auch über moderne eingebaute Features des Interfaces wie einen GPS-Sender. Auch alle einblendeten Statusanzeigen, die ansonsten direkt als Software-Interface der Spieler in Erscheinung treten, werden als zusätzliche Schicht des Interfaces der Rahmenhandlung gerechtfertigt.

Unabhängig davon, ob Suggestionen narrativ gerechtfertigt sind oder nicht, verdeutlichen sie ein zentrales Prinzip der Darstellung im Rahmen der Interaktion: Das Programm spricht einige Aspekte deutlicher aus als andere. Es zeigt den Spielern nicht nur die fiktive Spielwelt, sondern signalisiert und protokolliert Möglichkeiten und Resultate ihrer Eingriffe in das Spielgeschehen. In der Tat ergibt sich durch die Notwendigkeit dieser Interaktionshilfen eine zentrale Herausforderung des Designs der sinnlich wahrnehmbaren Form. Wie weit ist eine ansprechende, unberührte Form mit einer funktionalen Form vereinbar? Welche ästhetischen Entscheidungen sind der Interaktion geschuldet?

Die grafische Darstellung ist von dieser Problematik weitaus stärker betroffen als die akustische Ausgestaltung. Das Bild kann über eine beliebige Zeit unverändert bleiben, während der Ton gerade aus Veränderungen in der Zeit hervorgeht. Selektiert der Spieler z. B. in einem Echtzeitstrategiespiel eine Einheit, erhält sie eine Umrandung, so lange sie ausgewählt bleibt. Die Umrandung ist ein reines Interface-Artefakt und narrativ nicht gerechtfertigt. Die Sprachsamples hingegen sind gleichzeitig Informationsgeber und narrative Komponenten. Sie bestätigen die Aktionen des Spielers und vermitteln wichtige Ereignisse. Einheiten melden sich, wenn sie fertig ausgebildet sind oder Schaden erleiden und lenken damit die Aufmerksamkeit des Spielers auf auszuführende Aktionen. Die Verknüpfung akustischer Elemente mit Prozessen des Spiels bestimmt ihren Informationsgehalt hinsichtlich des regelbasierten Systems.

Fehlende Suggestionen als Bestandteil der Spielmechanik

Die Hervorhebung bestimmter Spielgegenstände oder -situationen durch visuelle Effekte oder Klänge erleichtert den Spielern die Identifikation von dem, was spielrelevant ist. Nicht immer ist so eine Form der Erleichterung gewünscht. In einigen Spielen, Spielmodi oder -situationen wird sie vielmehr bewußt ausgelassen. Die Suche nach benutzbaren Ge-

genständen und das rechtzeitige Erkennen handlungserfordernder Spielkonstellationen gehören dann zu den Herausforderungen des Spiels.

Besonders gravierend fällt das Auslassen der Erkennungshilfen dann auf, wenn die Spieler sich bereits an sie gewöhnt haben. Während z. B. die normalen Spielmodi in *Left 4 Dead 2* alle Mitspieler farblich konturieren, ist diese Hilfe im „Realism-Mode“ entfernt. Neben angepassten Spielregeln, die den Schwierigkeitsgrad erhöhen, wird gerade die eingeschränkte Sicht auf die Positionen der Mitspieler in diesem Modus zu einer zentralen Herausforderung. Erfordert das Spiel auch in den regulären Modi ein gemeinsames Vorgehen, ist die gegenseitige Abstimmung und Rücksichtnahme nun noch entscheidender für den Spielerfolg. Sobald ein Spieler die Gruppe der Überlebenden aus den Augen verliert und in Gefahr gerät, müssen die Mitspieler nach ihm suchen. Der Ruf nach Rettung über den laufenden Sprach-Chat verlangt eine möglichst präzise Beschreibung der eigenen Position.

Ein weiteres Beispiel für das gezielte Auslassen solcher Wahrnehmungshilfen bietet der Multiplayer-Shooter *Battlefield: Bad Company 2*, in dem zwei Armeen mit jeweils maximal 16 Spielern gegeneinander antreten. Während die Mitglieder des eigenen Teams sowohl im Hauptfenster als auch in der Minimap blau bzw. grün gekennzeichnet sind, bleibt der Aufenthalt der gegnerischen Truppen zunächst verborgen. Eine gegnerische Figur zu erkennen, die in Tarnkleidung durch das Dickicht eines Dschungels stapft, bedarf hoher Aufmerksamkeit und einer konzentrierten Abtastung der sinnlich wahrnehmbaren Form. Die visuelle Wahrnehmung ist aufgrund geringer farblicher Kontraste vornehmlich auf die Bewegungserkennung angewiesen. Gegnerische Figuren, die sich in der Nähe der eigenen Position befinden, sind gegebenenfalls auch durch ihre Fußschritte hörbar.

Sichtet jemand aus dem Team einen Feind, so kann er ihn durch das Drücken der „Tagging“-Taste (Q) markieren. Begleitet von einem Sprachsample, das die Aktion bestätigt, erscheint dann über der Figur ein roter Punkt. Auch auf der Minimap ist diese rote Markierung für alle Team-Mitglieder sichtbar. Die Position des Gegners ist auch für Mitspieler entlarvt, die sich an ganz anderen Orten der Karte befinden. Wie wichtig die Nutzung dieser Tagging-Mechanik für ein erfolgreiches Spiel ist, belegt u. a. das in den Steam-News gepostete Video "How to not be a noob.. in BFBC2"²⁴, in dem "Learn how to spot your enemies" an ers-

²⁴<https://www.youtube.com/watch?v=0p4fkH9qC4M> (19.06.2013)

ter Stelle steht. Die visuelle Hilfe zur leichteren Erkennung gegnerischer Figuren ist damit nicht ganz aus dem Spiel entfernt, sondern durch die Spieler aktivierbar. Da die Sichtung gegnerischer Figuren für den Erfolg des Teams entscheidend sein kann, wird das Tagging ein wichtiger Bestandteil des Gameplays. Durch die Implementierung dieser Mechanik tritt die gebündelte Wahrnehmung beider Teams gegeneinander an.

1.2 Spiel und Erzählung: Von Zahlen zu Zeichen

Spiele und Erzählungen sind die übergeordneten Strukturen, die Computerspielen zu Grunde liegen. Sie ordnen und motivieren den Spielablauf und das Geschehen in der fiktiven Welt. Einige Computerspiele wie *Tetris* legen weniger Wert auf die Erzählung. Spielziel und Motivation liegen darin, eine hohe Punktzahl zu erreichen. Andere wie *Heavy Rain* legen weniger Wert auf das Spiel. Sie sind statt dessen darauf ausgelegt, die Spieler eine Geschichte erleben zu lassen. Beide Strukturen treten in unterschiedlichen Mischungsverhältnissen und Gewichtungen auf. Der Begriff *Computerspiel* ist in dieser Hinsicht viel mehr ein Überbleibsel der Entwicklungsgeschichte des Mediums, als dass der Begriff die Fülle der heute darunter subsumierten Ausprägungen adäquat erfasst.

Für die Auseinandersetzung mit Spielen sind in unserer Kultur zwei Arten etabliert: das Zuschauen und die Teilnahme. Wir können ein Fußballspiel oder einen Boxkampf im Stadion verfolgen oder selbst den Rasen betreten bzw. in den Ring steigen. Je nachdem, in welchem Verhältnis wir uns mit dem Spiel befinden, nehmen wir es anders wahr. Der Umstand, dass es zwei verschiedene akzeptierte und weit verbreitete Formen der Auseinandersetzung mit Spielen gibt, betrifft auch das Computerspiel. Die Ausgaben des Computerspiels und alle darin aufgehenden akustischen Elemente können auf unterschiedliche Arten wahrgenommen werden. Um die Unterschiede beider Arten nachzuvollziehen, stelle man sich eine Gruppe von Menschen in einem Raum vor, von der die eine Hälfte spielt und die andere Hälfte beobachtet. Für die Beobachter, die über Lautsprecher und Monitore ein Computerspiel wie *Batman: Arkham City* betrachten, wird dessen Positionierung als audiovisuelles Medium nicht unplausibel sein. Sie nehmen eine Kombination aus Bild und Ton wahr, die in vielen Belangen den Konventionen

des Films entspricht. Die orchestrale Musik im Hintergrund, die Dialoge in den Zwischensequenzen und die visuellen Effekte könnten genau so gut einem Hollywood-Film entstammen. Für die Spieler ist die Einordnung des Computerspiels als audiovisuelles Medium jedoch in vielerlei Hinsicht problembehaftet.

Bild und Ton sind zwar in den meisten Computerspielen die dominanten Konstituenten der Darstellung, aber nicht die Einzigen. Der Begriff „audiovisuell“ blendet textuelle Inhalte aus, die auch in Computerspielen mit kinematographisch orientierten Darstellungen vorhanden sind. Neben rein textuellen Formen des Computerspiels wie Textadventures und MUDs enthalten z. B. auch Rollenspiele immer wieder lesbares Textmaterial, sei es das Questlog oder eine Hintergrundgeschichte. In *Batman: Arkham City* schalten die Spieler durch das Lösen von Rätseln Kurzgeschichten frei, die in einer Datenbank über Gotham City und seine Bewohner einsehbar sind. Neben dem Text ignoriert der Begriff des Audiovisuellen die Vibrationen des Force-Feedback-Controllers, den die Spieler in ihren Händen halten. Immer dann, wenn Batman einen Schlag abbekommt, rüttelt das Gerät und gibt den Spielern eine Rückmeldung über ihren Misserfolg. Es gibt also zusätzliche Ausgaben, welche die Beobachter gar nicht mitbekommen.

Der Begriff des Audiovisuellen beschreibt das Computerspiel aus der Sicht der Beobachter, die neben den Spielern sitzen. Die Beobachter nehmen nur einen Teil der Ausgabe wahr und sind nicht an der Eingabe beteiligt. Sie rezipieren einen audiovisuellen Strom, der ohne die Eingaben der Spieler gar nicht zustande käme. Als Zuschauer und Hörer erkennen sie nicht den Unterschied zwischen einer fest verschmolzenen audiovisuellen Sequenz und einer prozedural entstandenen Kombination aus akustischen und visuellen Komponenten. Die Spieler hingegen sind nicht nur Zuschauer und Hörer. Vielmehr sind sie Zuschauer und Hörer, um Spieler zu sein. Es gibt einen Zweck, der hinter ihrer Wahrnehmung der Darstellung steckt, nämlich die Interaktion mit dem Spiel. Für sie ist die sinnlich wahrnehmbare Form nicht nur eine Darstellung, sondern eine Prämisse für die Teilnahme. Sie treffen Entscheidungen als Reaktion auf das Gehörte und Gesehene. Gleichzeitig zeigt ihnen die Darstellung die Resultate dieser Entscheidungen. Die Spieler bewegen sich in einem Feld von Möglichkeiten und sind sich dessen Potenzialität bewußt.

Spielmodelle

Das Spiel ist ein dermaßen elementarer Bestandteil unserer Kultur, dass keine Definition dazu in der Lage ist, all seine Facetten zufrieden stellend abzudecken. Innerhalb der Game Studies herrscht jedoch große Einigkeit darüber, dass Spiele regelbasierte Systeme sind. Laut Jesper Juul ist ein Spiel „a rule-based system with a variable and quantifiable outcome, where different outcomes are assigned different values, the player exerts effort in order to influence the outcome, the player feels emotionally attached to the outcome, and the consequences of the activity are negotiable.“²⁵.

Die Regelhaftigkeit trägt entscheidend dazu bei, dass Spiele mit Hilfe des Computers implementierbar sind. Ebenso wie ein Programm lässt sich ein Spiel als Folge von Anweisungen betrachten, die schrittweise ausgeführt werden. Es gibt einen Startpunkt und einen Ausgang. Spielregeln steuern den Ablauf, indem sie Kontrollstrukturen wie Abzweigungen und Schleifen bereit stellen. Darüber hinaus enthalten sie Bedingungen, die mit Hilfe logischer Konjunktionen, Disjunktionen und Vergleichsoperationen definiert sind. Die strukturellen Parallelen zwischen der formalen Regelsprache des Spiels und imperativen Programmiersprachen legen es nahe, Spielanleitungen in Algorithmen zu überführen. Nicht umsonst waren Computerpioniere wie Alan Turing und John von Neumann von Spielen fasziniert. Turing veröffentlichte Fachbeiträge über Schachalgorithmen²⁶, während von Neumann mit seiner Spieltheorie die Entscheidungssituationen von Gesellschaftsspielen analysierte, in eine mathematische Form überführte und als Erklärungsansatz für wirtschaftliche Fragestellungen zu Grunde legte²⁷.

Eine Besonderheit von Computerspielen besteht darin, dass die Spielregeln direkt in das Programm integriert sind. Bei Brett- und Sportspielen hingegen sind sie extern in umfangreichen Regelbüchern festgehalten. Die Internationale Skatordnung (Stand November 2010) umfasst z. B. 65 Seiten, das Regelwerk des Deutschen Fußball Bunds für die Saison 2011/2012 ist auf 120 Seiten ausgebreitet. Die Implementierung dieser Regelsätze macht den Computer zu einer Kontrollinstanz, die sämtliche Verwaltungsaufgaben rund um das Spiel übernimmt. Sie erledigt

²⁵(Juul, 2005, 36)

²⁶<http://www.turingarchive.org/browse.php/B/7> (19.06.2013)

²⁷(von Neumann u. Morgenstern, 2007)

die Spielvorbereitung und die Auszählung, würfelt für uns und stellt künstliche Intelligenzen als Gegenspieler bereit. Als Leopold George Baron von Reiszitz im Jahr 1812 dem König ein militärisches Planspiel vorlegte, das ähnlich wie heutige Strategiespiele die Zusammensetzung unterschiedlicher Karten auf Basis von Terrain-Bausteinen vorsah, war die Spielvorbereitung mit erheblichem Aufwand verbunden. „Während der eigentliche Ablauf einer Schlacht Sache von Stunden oder Minuten war, soll der Aufbau der Spielfläche Friedrich Wilhelm III mitunter einen halben Tag gekostet haben.“²⁸ Computerspiele wie *Civilization* generieren neue Karten mit mehr Geländetypen und einem größeren Spielfeld in wenigen Sekunden. Der organisatorische Aufwand für die Benutzer ist erheblich reduziert.

Computerspiele profitieren von der Fähigkeit eines Programms, Dinge und Abläufe zu simulieren. Sie formen ausschnittshafte *Modelle*, die alle für das Spiel wesentlichen Aspekte berücksichtigen und andere ausblenden. Das Modell des Opernhauses in *Civilization V* besteht z. B. aus den Produktionskosten (200 Hämmer), den Unterhaltskosten (2 Gold pro Runde), der Technologie-Voraussetzung Akustik, einem freien Slot für einen Künstler und dem Effekt, vier Kulturpunkte pro Runde zu generieren. Weitere Aspekte wie der Programmplan oder das Renommee des Hauses sind in *Civilization V* nicht mit einbezogen, würden aber wahrscheinlich von einem als Wirtschaftssimulation konzipierten „Opern-Manager“ berücksichtigt werden. Neben Akteuren und Objekten enthalten die Modelle des Spiels Raum- und Zeitstrukturen. So simuliert *Civilization V* als Spielfeld eine Weltkarte und koppelt die fortschreitende Zeit der im Spiel durchschrittenen Jahrtausende an die Anzahl der durchgeführten Spielzüge.



Abbildung 1.1: Fertigkeit aus *Diablo 3*

Je besser die Spieler mit den Modellen des Spiels vertraut sind, desto genauer können sie die Reaktionen des Systems vorhersagen. Auch wenn viele Spiele Zufallselemente enthalten, dominiert die Determiniertheit des Spielsystems. Sie ist mit

²⁸(von Hilgers, 2000, 63)

Sicherheit ein Grund dafür, dass Computerspiele in den Industrienationen zu einer beliebten Freizeitbeschäftigung und Zuflucht geworden sind. Im Gegensatz zu einer undurchschaubaren Welt bilden die durchschaubaren Modelle virtueller Welten einen Gegenpol, der gerade durch seine regelbasierte Klarheit eine große Anziehungskraft entwickelt

Spielregeln weisen den Modellen Eigenschaften und Fähigkeiten zu und legen fest, was verboten oder erlaubt ist. Ein Streifzug durch die Regelsätze beliebter Spiele zeigt, dass keines von ihnen ohne Zahlen auskommt. Das Spiel dauert 90 Minuten. Jeder Mitspieler erhält 10 Karten. Wer eine sechs würfelt, darf noch einmal würfeln. Der Spieler mit der höchsten Punktzahl gewinnt. Ein Beispiel für die mit Zahlen durchtränkten Regeln des Computerspiels zeigt Abbildung 1.1, die eine aktive Fertigkeit des Mönch-Charakters aus *Diablo 3* beschreibt. Im Vergleich zu Brett-, Karten- und Sportspielen sind die Modelle des Computerspiels potenziell umfangreicher. Sie erstrecken sich von einfachen Positionsangaben bis hin zu komplexen physikalischen Beschreibungen der Spielwelt und ihrer Inhalte.

Zahlen sind das Fundament, auf dem das Spiel Dinge modelliert. Obwohl Zahlen in Spielen eine sehr zentrale Rolle einnehmen, sind sie während des Spielablaufs nicht unbedingt sichtbar. Wir merken uns in Ballsportarten die Punktzahl oder verwenden Hilfsmittel wie Anzeigetafeln dafür. Wir wissen aus Regelbüchern, dass wir im Schach den Bauern anfangs um ein oder zwei Felder, danach nur um ein Feld nach vorne bewegen dürfen. Auch wenn sie nicht unmittelbar sichtbar sind, legen Zahlen den Rahmen des Spiels fest und beschreiben den jeweils aktuellen Stand des Spielgeschehens. Eine Schachfigur verfügt z. B. über eine eindeutige Position auf dem 64-feldrigen Brett und ist an einen klar definierten Satz von Bewegungsregeln gebunden. Ebenso besitzt der Spieleravatar in *Quake* eine Position innerhalb eines dreidimensionalen Raums, ein Inventar, das sein verfügbares Waffenarsenal enthält, und einen Wert, der seinen Gesundheitszustand repräsentiert. Zu jedem Zeitpunkt des Spiels lässt sich die Spielsituation eindeutig durch eine Menge von Zahlenwerten beschreiben, die im Folgenden auch *Zustände* genannt werden. Der Fortlauf des Spiels vollzieht sich durch Veränderungen dieser Zustandsmenge. Dies kann zum Einen durch den neuen Wert eines bereits bestehenden Spielzustands geschehen. Zum Anderen sind Veränderungen der Zustandsmenge durch die Erzeugung und Zerstörung bestimmter Spielzustände – z. B. durch das Hinzunehmen oder Entfernen eines Spielobjekts – möglich. Im Folgenden wird die Ver-

änderung von Spielzuständen als *Spielereignis*, oder kurz *Ereignis*²⁹, bezeichnet. Ein Spielereignis kann mehrere weitere Spielereignisse zur Folge haben.

Einfach demonstrieren lässt sich diese formale Beschreibung des Spielverlaufs am Beispiel des Schachspiels. Jeder Spielzug besteht aus der Positionsveränderung einer Spielfigur. Die Werte ihrer Koordinaten ändern sich entsprechend. Steht auf dem Zielfeld eine gegnerische Figur, wird diese geschlagen. Der Zug verändert dann nicht nur den Zustand der bewegten Figur, sondern entfernt auch die geschlagene Figur vom Spielbrett und tilgt deren Zustände. Um den Verlauf einer Schachpartie als Folge einzelner Spielzüge zu beschreiben, etablierte sich eine eigene Notation. Hinter den Nummern der Spielzüge beschreibt man die Positionsveränderungen in der Form „Ursprungsposition - Zielposition“. Die Kodierung Lf1-b5 bedeutet z. B., dass der Läufer von dem Feld f1 auf das Feld b5 bewegt wird. Die folgende Box zeigt die mit dieser Hilfe festgehaltene Partie zwischen den Spielern Roesch und Willi Schlage aus dem Jahr 1910, die in Stanley Kubricks Film *2001: Odyssee im Weltraum* aufgegriffen wurde.³⁰ Dort kommen Abschnitte daraus in dem Wettstreit zwischen dem Bordcomputer HAL 9000 und Dr. Frank Poole vor.

1. e2-e4 e7-e5	9. e4xd5 Sf6xd5
2. Sg1-f3 Sb8-c6	10. Sf3xe5 Sd5-f4
3. Lf1-b5 a7-a6	11. De2-e4 Sc6xe5!
4. Lb5-a4 Sg8-f6	12. De4xa8?? Dd8-d3!
5. Dd1-e2 b7-b5	13. Lb3-d1 Lc8-h3
6. La4-b3 Lf8-e7	14. Da8xa6 Lh3xg2
7. c2-c3 0-0	15. Tf1-e1 Dd3-f3!!
8. 0-0 d7-d5	

Für eine künstliche Intelligenz wie HAL sind formale Kodierungen des Spielgeschehens wie sie die Schachnotation bietet, außerordentlich gut geeignet. Wir Menschen spielen eine Partie Schach hingegen in den sel-

²⁹Die Bezeichnungen Zustand und Ereignis sind lose an das Programmierparadigma *event driven programming* angelehnt, das insbesondere für die Modellierung von Benutzeroberflächen ausgelegt ist.

³⁰siehe http://de.wikipedia.org/wiki/Roesch_%E2%80%93_Schlage,_Hamburg_1910 (11.05.2012)

tensten Fällen, indem wir uns die Codes der Spielzüge zurufen, sondern verwenden Darstellungen des Modells.³¹

Spieldarstellungen

Als formal beschreibbare Systeme lassen sich Spiele von ihren Darstellungen entkoppeln. Diese Abstraktionsmöglichkeit hat zur Folge, dass die Spiel-Modelle zur Simulation von Spieldurchgängen ausreichen. Entwickler nutzen diesen Umstand gelegentlich zu Testzwecken, um in möglichst kurzer Zeit die Daten vieler Spieldurchläufe zu sammeln und die Spielbalance zu verfeinern. Die Entscheidungen der Spieler werden an Künstliche Intelligenzen delegiert. Gabe Newell berichtet in einem Interview von der Anwendung dieser Praktik bei der Entwicklung des Spiels *Left 4 Dead*.

We can turn off the graphics and have the game play itself and then use much more statistical methods for analyzing outcomes. We can play tens of thousands of games every night and then spend some time looking at graphs rather than watching individual players play.³²

Wenn die für den Menschen vorgesehene Darstellung und Interaktionsmöglichkeit wegfällt und das Programm nur auf dem Input künstlicher Intelligenzen operiert, dann ist das Spiel auch von den zeitlichen Beschränkungen befreit, die ihm die Wahrnehmung des Menschen auferlegt. Es braucht keine 24 Bilder pro Sekunde, um einen flüssigen Bewegungsablauf zu simulieren, keine hohen Audio-Samplerates und kein gedrosseltes Gameplay-Timing, um menschenmögliche Reaktionen zu ermöglichen. Stattdessen kann es in seinem Rhythmus entsprechend den Rechenkapazitäten der Hardware ungehemmt ablaufen. Das Warten auf den langsamsten Teilnehmer des Systems, den Benutzer, entfällt. Das Beispiel verdeutlicht, dass sowohl die Ein- als auch die Ausgabemöglichkeiten speziell auf den menschlichen Wahrnehmungsapparat abgestimmt sind. Künstliche Intelligenzen wie HAL oder die Akteure aus *Left 4 Dead* benötigen keine Darstellung des Spielgeschehens, sondern lediglich die Kenntnis der Spielzustände und des Regelsatzes. Sie

³¹ Auch bei brieflich geführten Schachpartien besitzen in der Regel beide Kontrahenten ein Schachbrett, auf dem sie den zugesandten Code umsetzen.

³² (Newell, 2008)



Abbildung 1.2: HAL vs. Dr. Frank Poole; *2001: Odyssee im Weltraum*

operieren auf den mit Hilfe von Zahlen beschriebenen Modellen des Spiels. Menschen nutzen hingegen in den seltensten Fällen numerische Darstellungen für Spielobjekte.

Alle für den Menschen aufbereiteten Spielräume, -figuren und -objekte, sei es bei Brett- oder Computerspielen, bestehen nicht nur aus Zahlen, sondern allem voran aus Zeichen. HAL spielt mit den numerischen Zustandsbeschreibungen, aber Dr. Poole benutzt das Schachbrett mit seinen Figuren. Das Schachbrett ist ein äußerst effektives Darstellungsmittel, da wir mit dessen Hilfe alle Spielfelder und die Konstellationen aller Figuren darauf überblicken. Wir unterscheiden die Figuren durch ihre Gestalt und nicht durch ihre Nummern. So ist der König eines Schachspiels durch eine Krone oder andere Insignien seines Amtes identifizierbar, während die Türme als abstrahierte Miniaturgebäude erkennbar sind. Das Brett und die Figuren sind Hilfen, die theoretisch nicht notwendig sind, um Schach zu spielen, es uns aber stark erleichtern. Spieldarstellungen bereiten die Informationen des jeweiligen Spielstatus in einer Form auf, die unseren geübten kognitiven Fähigkeiten entgegen kommt. Bei Brett- und Kartenspielen sind es in hohem Maße grafische oder figurative Symbole wie das Herz, die Dame oder der König. Bei Computerspielen ist das Zeichenrepertoire angereichert durch animierte Bilder, Klänge und Worte. Akustische Elemente sind ein konstitutiver Bestandteil der Darstellung.

Wenn die Zustände der Spielmodelle des Computerspiels durch Zahlen beschrieben werden und das Spielprinzip von den Spielern die Kenntnis der Zustände voraussetzt, muß es innerhalb des Programms eine Zuordnung von Zahlen zu Zeichen geben. Die Benutzerschnittstelle übersetzt eine Auswahl der formalen Beschreibungen des Modells in eine Darstellung, die für den menschlichen Wahrnehmungsapparat in einer Form zusammengesetzt ist, deren Interpretation wir gewohnt sind. Einerseits kommt diese spezielle Aufbereitung der zum Spielen nötigen Informationen unseren geübten kognitiven Fähigkeiten entgegen. Andererseits entwickelt die Spieldarstellung dadurch eine zusätzliche Attraktivität. Es ist zu erwarten, dass die grafische und akustische Repräsentation eines Feuers die Spieler stärker in ihren Bann zieht, als eine Zahl, die das Feuer als Spielereignis repräsentiert.

Welche Beziehungen bestehen nun zwischen Oberfläche und Tiefenebene, Frontend und Backend, dem Modell und den Zeichen? Wie verhält sich in der Metaphorik von Crawford's Interaktionsbeschreibung das „Denken“ des Computerspiels zu dem, was es „sagt“? Zunächst bleibt festzuhalten, dass die oben skizzierte strikte Trennung zwischen dem formalen Modell und den Zeichen seiner Darstellung selten außerhalb der Maschine auftritt. Für den Menschen geht der Akt der Modellierung eines Spielgegenstands Hand in Hand mit der Vorstellung, wie das Modellierte aussieht und klingt. Bereits während der Entwicklungsphase stehen die Modelle und die Zeichen zur Darstellung dieser Modelle in einem Abhängigkeitsverhältnis. Game Designer erstellen Modelle für Zeichen und Zeichen für Modelle. Die künstlerische Ausarbeitung dieser Vorstellungen ist schließlich Aufgabe von Grafik- und Sounddesignern, Komponisten, Motion Capture Artists und Autoren. Das fertig gestellte Computerspiel besitzt eine klar festgelegte Zuordnung von Aspekten des Modells zu Darstellungselementen.

Die Zuordnung von Zahlenwerten zu akustischen Elementen ist kein Prozess, der ausschließlich im Rahmen von Software auf Interesse trifft. So stellt die Transformation von Daten in akustische Signale ein Gegenstandsfeld der Sonifikation dar.³³ Das Ziel der Disziplin liegt darin, Verknüpfungen zwischen Daten und akustischen Signalen zu schaffen, die möglichst verlässliche Rückschlüsse beim Hören dieser Signale auf die referenzierten Daten zulassen. Sie sollen die Kommunikation und

³³Im angloamerikanischen Raum ist die Sonifikation ein Teilaspekt des übergeordneten Forschungsverbunds „auditory display“, der sich u. a. auch mit Sprachsynthese befasst.

Interpretation der Daten erleichtern.³⁴ Die sonifizierte Daten sind in der Regel Informationen, die für den Benutzer nicht unmittelbar sichtbar sind. Ein klassisches Beispiel ist der Geigerzähler, der die Dosis ionisierender Strahlung durch die Häufung von Tönen wahrnehmbar macht. Einparkhilfen nutzen das gleiche Prinzip, um die Nähe zu nicht sichtbaren Hindernissen akustisch abzubilden. Die Agenda der Sonifikation trifft nur zum Teil auf die Klangkulissen von Computerspielen zu, da dort in vielen Fällen die Herkunft und Bedeutung akustischer Artefakte bewußt verschleiert oder mehrdeutig gehalten wird. Gerade diese unaufgelösten Elemente tragen einen Großteil zum ästhetischen Reiz des Computerspiels bei. Darüber hinaus repräsentieren akustische Signale neben Sprache und Musik nur eine Teilmenge der Klangkulisse. Ergebnisse der Sonifikation können deshalb nur bedingt auf das Feld des Game Sounds übertragen werden. Nichtsdestotrotz bietet die Sonifikation eine wichtige Grundlage zur Analyse der funktionalen Dimension akustischer Elemente des Computerspiels.

Spieldarstellungen lassen sich austauschen, ohne dass sich am formalen System des Spiels etwas ändert. Die Figur des Königs in einem Schachspiel kann z. B. ein schlichter Klotz sein oder eine kunstvoll geschnitzte Holzfigur, deren Haupt von einer Krone geschmückt wird. Das Darstellungsspektrum reicht von abstrakten geometrischen Formen bis hin zu detaillierten Abbildungen. Was ändert sich nun an einem Schachspiel, wenn man die Figuren durch andere ersetzt? Wenn der Imperator im *Star Wars* Schach an die Stelle des Königs tritt und Sturmtruppen die Bauern ersetzen, findet ein Austausch der Symbole aus vergangenen Zeiten mit Alternativen aus unserer heutigen Kultur- und Konsumlandschaft statt. Die Spielregeln bleiben dabei jedoch unberührt. Ob König, Imperator, oder zylindrischer Klotz, die Rochade läuft immer unter den gleichen Vorzeichen ab. Dennoch stellt sich die Frage, ob und inwiefern sich die Spielerfahrung bei veränderter Darstellung und gleichbleibenden Regeln ändert. Ein Bilderpuzzle, dessen Motiv deutlich sichtbare Kontraste aufweist, ist schließlich einfacher zu lösen als eines, dessen Motiv auf der gesamten Fläche sehr ähnliche Farbtöne enthält. Im Sport herrscht diesbezüglich ein simpler Pragmatismus. So besitzt im Fußball jede Mannschaft ein Auswärtstrikot, um gegebenenfalls durch eine andere Farbgebung die Spieler unterschiedlicher Teams unterscheiden zu können. Grundsätzlich müssen die Parteien und Fi-

³⁴(Kramer u. a., 1997, 6)

guren unterscheidbar bleiben. Die Darstellung ist also nicht beliebig austauschbar.

Da die medial inszenierte Darstellung des Computerspiels deutlich aufwändiger gestaltet ist, als es bei Brettspielen überhaupt möglich ist, gewinnt die Frage nach unterschiedlichen Spieldarstellungen zusätzliches Gewicht. Modpacks, die Texturen und Soundeffekte ersetzen, sind keine Seltenheit. Die *Counter Strike* Modifikation *Cartoon Strike* ersetzt z. B. die 3D Modelle der Waffen durch bunte Spielzeugpistolen und die Schussgeräusche durch unschuldig klingendes Quietschen und Hupen. Eine ähnliche visuelle Transformation bewirken die „Pyrovision Goggles“ in *Team Fortress 2*.³⁵ Dass die Spielerfahrung nach diesem Austausch eine andere ist, als in der Originalfassung des Spiels, lässt sich kaum leugnen. Die Frage nach dem Einfluss der Darstellung auf die Spielerfahrung wird die folgenden Überlegungen begleiten. Eine erste Einordnung verspricht die Betrachtung der dargestellten Zeichen als Konstituenten einer Erzählung.

Narrative Zeichenkonfigurationen

Unter einer Erzählung verstehen wir im weitesten Sinne eine bestimmte Konfiguration von Zeichen, die ein Geschehen vermittelt. Je nachdem, ob die Erzählung mündlich oder schriftlich, in Form eines Romans, Theaterstücks, Films oder Computerspiels erfolgt, liegen unterschiedliche Zeichen und Konventionen zur Anordnung dieser Zeichen vor. Neben geschriebenem Text und gesprochener Sprache verwenden wir Bilder, Geräusche, Musik, und Gesten um Geschichten zu erzählen, Erlebnisse zu verarbeiten und Erfahrungen zu transportieren. Ein Roman erzeugt seine Handlung durch eine Folge von Wörtern, die wir vor dem Hintergrund unseres Wissens- und Erfahrungshintergrunds interpretieren. Ein Film besteht aus einer audiovisuellen Sequenz, die auf der Basis von gesprochener Sprache und bildlich-klanglicher Symbole Szenen darstellt. Das Computerspiel stützt sich als multimediales Medium einerseits auf Zeichen und Erzählweisen aus anderen Medien und entwickelt andererseits durch neue Zusammensetzungen dieser Zeichen ein eigenes Repertoire an Erzählstrategien.

³⁵siehe <http://www.youtube.com/watch?v=WUhOnX8qt3I> (24.06.2013)

Jede Spieldarstellung ist eine Konfiguration von Zeichen und damit im weitesten Sinne auch eine Erzählung, die sich interpretieren lässt. Selbst dann, wenn ein Spiel keine Handlung im Sinn einer traditionellen Erzählung aufweist und dessen Darstellung nur aus abstrakten geometrischen Formen besteht, bleiben die Spielabläufe interpretierbar. So versteht z. B. Janet Murray *Tetris* als Allegorie auf das amerikanische Leben der 1990er Jahre:

This game is a perfect enactment of the overtasked lives of Americans in the 1990s – of the constant bombardment with tasks that demand our attention and that we must somehow fit into our overcrowded schedules and clear off our desks in order to make room for the next onslaught.³⁶

Je abstrakter die dargestellten Zeichen ausfallen, desto größer scheint der Interpretationsspielraum der Spielabläufe zu werden. Interpretationen dieser Art stützen sich stärker auf die Regeln des Spiels als auf die Zeichen der Spieldarstellung. Murray bezieht ihre Interpretation nicht auf die Formen der Tetris-Bausteine oder das im Hintergrund laufende russische Volkslied Korobeiniki. Statt dessen stützt sie sich auf das Gameplay, das durch die immer schneller herabstürzenden Steine einen zunehmenden Druck auf die Spieler ausübt. Das Spiel legt durch seine Regeln eine ausdrucksstarke Konfiguration der Zeichen fest. Ian Bogost bezeichnet diese Form als *prozedurale Rhetorik*.

I call this new form procedural rhetoric, the art of persuasion through rule-based representations and interactions rather than the spoken word, writing, images, or moving pictures.³⁷

Bogost verwendet den Begriff im Zusammenhang mit politisch motivierten Computerspielen, die durch bestimmte Modelle und Regeln auf Mißstände aufmerksam machen wollen.³⁸ Die prozedurale Rhetorik ist in diesen Fällen intendiert. Im obigen Beispiel der *Tetris*-Interpretation ist sie es hingegen nicht. Der russische Programmierer Alexei Paschitnow wird bei der Konzeption des Spiels im Jahr 1984 kaum das amerikanische Leben der 1990er Jahre im Visier gehabt haben. Vielmehr ist der hohe Abstraktionsgrad der Zeichen eine Voraussetzung für die freie-

³⁶(Murray, 1997, 144)

³⁷(Bogost, 2007, ix)

³⁸siehe z. B. die Werke von Molleindustria: <http://www.molleindustria.org/>

re *Tetris*-Interpretation. Grundsätzlich darf nicht außer Acht gelassen werden, dass eine regelbasierte Repräsentation nie ohne Zeichen auskommt. Um Regeln und Modelle wahrnehmen zu können, sind Wörter, Bilder und Klänge unerlässlich. Sie tragen ihren Teil dazu bei, indem sie die Kontexte darstellen, innerhalb derer die Regeln bestimmte Umstände oder Prozesse hervorheben. Das Zusammenspiel aus Regeln und Zeichen reguliert zu einem starken Anteil den sich öffnenden interpretativen Spielraum.

Als Zeichenkonfigurationen, die aus regelbasierten Systemen hervorgehen, unterscheiden sich Spieldarstellungen in vielen Punkten von traditionellen Erzählweisen. Mündliche, schriftliche oder filmische Erzählungen richten sich an Zuhörer, Leser und Zuschauer. Die rezeptive Ausgangssituation des Computerspiels entspricht jedoch eher einem Dialog. Die Eingaben der Spieler und die Einbindung des Zufalls widersetzt sich einer präzisen Planung des Verlaufs der Zeichenkonfigurationen. Hinzu tritt eine repetitive Grundstruktur, die für traditionelle Erzählungen unüblich ist. Unter diesen Voraussetzungen fällt es sehr schwer, komplexe Geschehnisse und Erfahrungen gezielt zu vermitteln. Neben dem Spielsystem und seinen Darstellungen enthalten die meisten Computerspiele deshalb eine an das Spiel angelagerte Erzählstruktur, die für die Autoren eine Planungssicherheit schafft und auf unterschiedliche Weisen an das Spielgeschehen gekoppelt sein kann. Ihre zentralen Transportmittel sind Text und Sprache, die in traditionellen Erzählweisen eine privilegierte Position einnehmen. Darüber hinaus können aber auch alle weiteren Darstellungsmittel des Computerspiels inklusive aller akustischen Elemente im Dienste der angelagerten Erzählung stehen.

Die größte Planungssicherheit und die geringste Anbindung an das Spielgeschehen ist dann gegeben, wenn die Entscheidungen der Spieler die Schilderungen der Geschehnisse nicht beeinflussen. Lineare Geschichten, wie sie der Roman und der Film bieten, lassen sich häppchenweise in Zwischensequenzen als statisches Text- oder Videomaterial einbinden. Erfolg im Spiel führt zur Progression der Geschichte. Häufig dient die auf diese Weise integrierte Rahmenhandlung dazu, den Fortlauf des Spiels abzustecken. Sie gibt ein Setting vor und rechtfertigt die Aufteilung in Levels, Maps und Missionen sowie deren Reihenfolge. In offenen Spielumgebungen gibt sie möglicherweise eine Erklärung für die Grenzen der Welt. So sind die Spieler in *Gothic* in einem Landabschnitt gefangen, der durch eine magische Barriere abgeriegelt ist. Viele Spiele

nutzen das gleiche Prinzip und verorten die Handlung auf einer Insel oder in einem ähnlich abgeschlossenen Areal.

Deutlich komplexer und vielschichtiger sind hingegen Versuche, mit Hilfe des Computers die erzählerischen Möglichkeiten der Interaktion auszuloten. Hier entwickeln sich unterschiedliche Erzählkonzepte. Eines davon ist die *nichtlineare Erzählung*. Sie zeigt nicht nur eine Handlung, sondern bietet den Benutzern einen Handlungsspielraum an. Unterschiedliche Entscheidungen führen zu anderen Handlungsverläufen. An Stelle eines einzigen Handlungsstrangs steht eine baumartige, verzweigte Handlungsstruktur. Spiele wie *Batman Arkham City* und *Fallout 3* beschreiten einen Mittelweg, indem sie eine Haupthandlung mit mehreren optionalen Nebenquests kombinieren. Daraus ergeben sich mehrere parallel ablaufende Handlungsstränge, deren Reihenfolge nicht festgelegt ist. Die Struktur einer nichtlinearen Erzählung weist Gemeinsamkeiten mit der Struktur des Spiels auf. In beiden Fällen treffen die Spieler Entscheidungen, die unterschiedliche Ausgänge mit sich bringen. Die nichtlineare Erzählung besteht nicht nur aus den Zeichen, sondern besitzt einen regelbasierten Unterbau. Sie benötigt eine Metastruktur, die aus Regeln für ihren Aufbau und Ablauf besteht. Diese zentrale Gemeinsamkeit des Spiels und der Erzählung im Computerspiel ist eine Grundlage für Mischformen beider Strukturen. Im Mittelpunkt von ihnen stehen die Entscheidungen der Spieler und die Potenzialitäten, die sich daraus ergeben. Spiel und Narration überlagern sich potentiell. *Left 4 Dead* verwendet beispielsweise einen ereignis- und aktionsbasierten Erzählansatz, den Gabe Newell als „procedural narrative“ beschreibt.³⁹ Während der narrative Rahmen konstant bleibt – vier Überlebende kämpfen in einer Zombie-Apokalypse ums Überleben – bieten die Spieldurchläufe eine Reihe von Variationen. Alle im Rahmen des Spiels durchgeführten Aktionen ergeben dabei als Teile der Narration Sinn. Je nach Spielsituation vertiefen sich darüber hinaus ohne das Zutun der Spieler die Spielfiguren in kontextabhängigen, randomisierten Konversationen.

Je größer der Handlungsspielraum ausfällt, desto schwieriger gestaltet sich die Vermittlung eines geplanten Geschehens. Der Handlungsverlauf besteht vielmehr aus einer Aneinanderreihung diskreter Entscheidungen, einer Folge von Aktionen und Reaktionen. Das erzählerische Konzept besteht dann nicht mehr darin, den Spielern eine Geschich-

³⁹(Newell, 2008)

te zu erzählen. Stattdessen wird eine Umgebung geschaffen, in denen sie etwas erleben können. Die Erzählung wird zu einem Erlebnis. Ein Beispiel für diese erlebnisorientierte Erzählkonzeption ist das Computerspiel *The Graveyard* von dem Studio *A Tale of Tales*, in dem die Spieler eine gebrechliche alte Frau durch einen Friedhof steuern. Es gibt kein Spielziel und keine Aufgaben, die die Spieler bewältigen müssen. Sie begeben sich in einen anderen Körper an einem anderen Ort. Fiktive Archetypen eines solchen Versenkungsangebots sind das Holodeck aus *Star Trek* und die *Matrix*. Das Holodeck entwirft einen Zugang in fiktive Welten, der den gesamten Körper des Benutzers einbezieht. Bewegungsgesteuerte Controller wie die *Wii Mote*, *PS Move* und *Microsoft Kinect* gehen in diese Richtung. Die *Matrix* hingegen wird durch ein Brain-Computer-Interface (BCI) betreten. Nichtinvasive EEG-Geräte für den Heimanwender befinden sich zwar seit einige Jahren auf dem Markt⁴⁰, konnten sich jedoch bislang noch nicht bei der breiten Masse durchsetzen. Beide Vorbilder betonen die Illusion der Präsenz des Teilnehmers innerhalb fiktiver Umgebungen.

Welche Unterschiede bleiben trotz dieser Gemeinsamkeiten, die eine konzeptuelle Trennung beider Strukturen rechtfertigen? Im Zentrum einer Erzählung steht nicht die Zahl, sondern das Zeichen. Egal ob Wort oder Geste, Bild oder Klang, die Grundpfeiler einer Erzählung liegen grundsätzlich auf einer höheren Abstraktionsstufe als die Zahlen des Spiels. Um Erzählungen mit Hilfe des Computers zu realisieren, ist deutlich mehr Aufwand erforderlich als bei Spielen. Sie zielen auf sinnhafte Konfigurationen dieser Zeichen. Zahlen des Erzähl-Modells sind reine Hilfsmittel. Für das Spiel hingegen sind eher die Zeichen ein Hilfsmittel, während die Zahlen – Punktestände und Statistiken – den Kern bilden. Die Darstellung eines Spiels ist immer auch eine Konfiguration von Zeichen. Objekte, Akteure und die Spielwelt müssen den Spielern über die Benutzerschnittstelle zugänglich gemacht werden. In diesem Sinne kann jedes Computerspiel zu einem gewissen Grad als Erzählung verstanden werden.

⁴⁰Bereits 2008 stellte das australische Unternehmen Emotiv Systems ein EEG Headset auf der Game Developers Conference in San Francisco vor. Mittlerweile sind mehrere Modelle von verschiedenen Herstellern auf dem Markt, die teilweise für weniger als 100 Dollar angeboten werden. Siehe auch http://en.wikipedia.org/wiki/Comparison_of_consumer_brain-computer_interface_devices (12.08.2013)

Die meisten Titel vermitteln jedoch kaum ein Geschehen, das ohne Regelsätze und Spielziele Sinn ergibt. Diese Tatsache ist ein entscheidender Grund dafür, dass Filmumsetzungen von Spielen wie *Super Mario Bros.* aufgesetzt wirken. Diverse skurrile Monstertypen, Pilze und Sterne als Powerups sowie Röhren, durch die man versteckte Areale erreicht, ergeben in Hinblick auf die Spielmechanik Sinn. Als Konstituenten einer linear erzählten Geschichte sind sie jedoch völlig unplausible Erzählmuster, die narrativ schwer zu rechtfertigen sind.

Die stimmige Verbindung von Spieldarstellungen und einer geplanten Narration ist eine der größten Herausforderungen des Designs von Computerspielen. Sämtliche Spielobjekte, Aktionen, Ziele und Abläufe in eine konsistente Erzählung einzubetten, gestaltet sich sehr aufwändig. Die Schwierigkeiten bestehen in erster Linie darin, strukturelle Differenzen und auseinander driftende Motivationen miteinander zu versöhnen. Spiele enthalten Repetitionen, während Erzählungen üblicherweise Progressionen darstellen. Zeitschleifen im Zuge der Handlung wie in dem Film *Und täglich grüßt das Murmeltier* sind die Ausnahme. Spiele stellen eine Reihe von Anforderungen an ihre Teilnehmer. Je nach Spieltypus erfordern sie Reaktionsvermögen, Geduld oder strategisches Geschick. Nichtlineare Erzählungen sind weniger voraussetzungsreich. Sie erwarten Entscheidungen, die jeder Teilnehmer ohne vorausgegangenes Training fällen kann. Trotz der Differenzen sind Spiel- und Erzählelemente eng miteinander verbunden. So hält Juul fest:

Rules and fiction interact, compete, and complement each other. [...] fiction in video games plays an important role in making the player understand the rules of the game.⁴¹

Spielzüge und narrative Aktionen sind möglicherweise so dicht miteinander verwoben, dass sie als eine Einheit wahrgenommen werden. Die Zeichen der Oberfläche übernehmen eine semantische Doppelfunktion. Einerseits verweisen sie auf Spielzustände, d. h. bestimmte Zahlenwerte des Modells. Andererseits sind sie narrative Elemente, die in unterschiedlichen Abstraktionsgraden auf Dinge außerhalb des Spiels zeigen. Ein herabfallender Tetris-Stein verweist als Spielobjekt auf seine Geschwindigkeit, geometrische Form und Fallrichtung. Als narratives Element steht er z. B. für eine Aufgabe, die wir in unserem überfüllten Terminkalender unterbringen müssen. Die Verweise auf das Spielmodell sind determiniert. Narrative Verweise bleiben offen und können in

⁴¹(Juul, 2005, 163)

den Diskursen der Spielkultur immer wieder neue Richtungen einschlagen. Akustische Elemente befinden sich genau in diesem Spannungsfeld zwischen kodierter Information und für Interpretationen geöffneter Narration.

1.3 Lernvorgänge und Spielweisen

Was Computerspiele ausdrücken und welche Bedeutungen ihnen zukommen, ist nicht alleine nur durch die Produktion, sondern in hohem Maße durch die Rezeption, den Umgang mit ihnen, bestimmt. Die Arbeiten der Soziologen Stuart Hall und Richard Hoggart legen nahe, dass kulturelle Objekte innerhalb eines ideologischen Rahmens erstellt und anschließend in einem ideologischen Kontext wahrgenommen werden, den die Empfänger auf Basis ihrer Erfahrungen weiter formen. Die Spieler, die nicht nur Empfänger, sondern vielmehr Konversationspartner und Mitgestalter sind, bestimmen selber, welche Bedeutung ein Spiel für sie annimmt.

The meanings of games are not essential or inherent in their form (though form is a crucial determinant), even if we define form as a set of rules and constraints for gameplay, and certainly not in their extractable „stories“ (though the fictive storyworld matters in most games), but are functions of the larger grid of possibilities built by groups of developers, players, reviewers, critics, and fans in particular times and places and through specific acts of gameplay or discourse about games.⁴²

Um dieses Netz der Möglichkeiten in den Blick zu bekommen, das Steven E. Jones im Mittelpunkt der Bedeutung von Spielen verortet, ist die Betrachtung bestimmter Spielweisen und Diskurse über Spiele eine notwendige Voraussetzung. Die folgenden Ausführungen skizzieren zunächst die grundlegenden kognitiven Lernprozesse, die den Spielern bei der Beschäftigung mit einem neuen Computerspiel abverlangt werden. Im Anschluss daran werden zwei idealtypische Spielweisen vorgestellt, die jeweils unterschiedliche Auseinandersetzungen mit den Zeichen der Ausgabe und damit auch akustischen Elementen nach sich ziehen.

⁴²(Jones, 2008, 3)

Lernvorgänge

Durch die Komponente des Spiels unterscheidet sich die Rezeption des Computerspiels von der Rezeption anderer Kulturgüter. Im Gegensatz zum Lesen von Büchern und Betrachten von Filmen sind Computerspiele deutlich voraussetzungsreicher. Computerspiele sind immer auch Lernumgebungen.⁴³ Jedes Spiel erfordert eine Einarbeitungszeit, in der sich die Spieler mit seinen Grundlagen vertraut machen. Es geht darum, Regeln, Ziele und Spielabläufe zu verstehen. In Abhängigkeit mehrerer Faktoren wie der Komplexität des Spiels und dem Vertrautheitsgrad der Spieler mit bestimmten Genres und Konventionen fällt dieser Lernprozess unterschiedlich lang aus. Mau-Mau ist leichter nachzuvollziehen als Skat oder Doppelkopf. Karten- wie Computerspiele besitzen unterschiedliche Komplexitätsgrade. Strategie- und Rollenspiele neigen z. B. aufgrund ihres Umfangs dazu, eine Einarbeitungszeit von mehreren Stunden zu erfordern. *Casual Games* veranschlagen hingegen nur einen Lernaufwand von wenigen Minuten.

Auch dann, wenn die Spieler mit den Grundlagen des Spiels vertraut sind, setzt sich der Lernprozess fort. Die Königsdisziplin des Game Designs besteht schließlich darin, Spiele zu schaffen, die einfach zu erlernen und schwierig zu meistern sind.⁴⁴ Gewinnen und Scheitern in variierenden Schwierigkeitsgraden sind zentrale Aspekte des Spiels.⁴⁵ Die Auseinandersetzung mit jedem Computerspiel ist also über die gesamte Nutzungsdauer mit einem Lernprozess verbunden, der von den Grundlagen der Bedienung über das Erfassen möglicher Spielzüge, Regeln und Ziele bis hin zu übergeordneten Strategien reicht. Das folgende Abschreiten verschiedener Stufen dieses Lernprozesses verschafft einen ersten Überblick über die spezielle Rezeptionssituation des Computerspiels. Im Mittelpunkt stehen dabei die Auseinandersetzungen der Spieler mit den Ausgaben.

Der erste Schritt des Lernprozesses - das Kennenlernen der Bedienung - ist bei Computerspielen deutlich komplexer, als bei traditionellen Brett-

⁴³Eine Lanze für die positiven Effekte, die die Auseinandersetzung mit Computerspielen speziell für Jugendliche mit sich bringt, bricht James Paul Gee in (Gee, 2003).

⁴⁴Die Prägung der Design-Richtlinie „easy to learn, hard to master“ wird dem Atari-Gründer Nolan Bushnell zugeschrieben.

⁴⁵Computerspiele fallen damit in die von Espen Aarseth konzipierte Kategorie der ergodischen Kunst, die alle Voraussetzungen in sich trägt, um zwischen erfolgreichen und nicht erfolgreichen Benutzern zu trennen (Aarseth, 1997, 179).

und Kartenspielen. Zwischen dem Mitteilen von Spielentscheidungen und der Bewegung von Spielobjekten steht die Hürde der Benutzerschnittstelle. Die Spieler müssen sich mit den Eingabegeräten vertraut machen und die Zuordnung der Eingaben zu Spielaktionen nachvollziehen. Welche Eingaben welche Aktionen auslösen, erfahren die Spieler durch die vorrangig grafischen und akustischen Rückmeldungen. So erkennen sie z. B. in einer Fußballsimulation durch die dargestellten Animationen und Soundeffekte, dass der gedrückte grüne Knopf des Gamepads bei Ballbesitz in einem Pass zum Mitspieler resultiert. Die Ausgaben wiederum stellen die Ermessensgrundlage der Eingaben dar.

Um erfolgreich zu spielen, müssen sich die Spieler gut mit der Benutzerschnittstelle auskennen. Sie lernen Shortcuts auswendig und trainieren die im Spiel gefragten Aktionen. Auf diese Weise machen sie sich so gut mit der Benutzeroberfläche vertraut, dass ihre Funktionsweise ab einem gewissen Punkt keine Aufmerksamkeit mehr erfordert. Computerspiele legen einen möglichst effizienten Umgang mit der Benutzerschnittstelle nahe, da sich jeder unnötige Tastendruck potenziell negativ im Spielergebnis niederschlägt.

Sämtliche Aktionen finden im Rahmen eines Spielkontextes statt. An Stelle eines Passes könnte sich in einer bestimmten Situation z. B. auch eine Flanke oder ein Schuß anbieten. Auch dann, wenn die grundlegenden Zuordnungen erlernt sind, stellt sich den Spielern während des weiteren Spielverlaufs kontinuierlich die Frage, welche Aktionen in welchen Kontexten zu welchen Zeitpunkten nützlich sind. Die anwendbare Aktionsmenge in einem Kontext wird im Folgenden auch als Handlungsspielraum einer Spielsituation bezeichnet. Die Auslotung des Handlungsspielraums bildet den zweiten Schritt des Lernprozesses. Während die Zuordnung der Eingaben zu den Spielaktionen in der Regel auch in der Bedienungsanleitung eines Computerspiels abgedruckt ist, übersteigt die Menge der kontextabhängigen Aktionen häufig das Maß des Dokumentierbaren. Vor allem die Interaktionen der Spieler mit der Spielwelt und anderen Akteuren erzeugen zahlreiche Momente, in denen der Ausgang von Aktionen erst gelernt werden muß.

Die Spieler erwerben ihre Kenntnisse über solche komplexeren Handlungsspielräume größtenteils durch das Ausprobieren – die Verifizierung zuvor gesetzter Annahmen. Den häufigsten Ansatz bildet eine sehr grundlegende heuristische Methode, die von Brute-Force-Algorithmen ebenso genutzt wird wie von Spielern: Versuch und Irrtum. Kann man

auf diese Wolke springen oder fällt man hinunter in die brennenden Stacheln? Gibt es in dem Spiel „friendly fire“⁴⁶? Ist für das Gebüsch am Straßenrand eine Kollisionsabfrage programmiert oder fährt der Sportwagen hindurch? Um diese Fragen zu beantworten, sind Tests nötig, die das Verhalten der dargestellten und simulierten Dinge prüfen. Bei einem Jump'n'Run äußert sich dieser Differenzierungsakt in der Identifizierung von begehbaren Plattformen während in einer dreidimensionalen Umgebung die zum Erreichen der Spielziele förderliche Nutzung dargestellter Objekte zur Disposition steht. Das Testen erstreckt sich von der Analyse von Aktionsausgängen bis hin zu einem regelrechten Abtasten der virtuellen Spielwelt. So sehen sich die Spieler kontinuierlich mit der Frage konfrontiert, welche Dinge der virtuellen Welt berührbar und manipulierbar sind. In Spielen mit einer zweidimensionalen Raumstruktur lässt sich dieses Abtasten als primär visuelle Herausforderung begreifen, die alle Darstellungselemente befragt, ob sie im Vordergrund liegen oder zum Hintergrund gehören. Sehr deutlich tritt dieser Vorgang z. B. bei Grafikadventures hervor. Ein Großteil des Spiels besteht hier aus dem Abtasten des Bildes mit dem Mauscursor nach nutzbaren Objekten und möglichen Interaktionen. Die Spieler verfolgen gewissermaßen ein *Reverse Engineering* der zusammengestellten grafischen Komponenten. Sie müssen sich den Bauplan der für sie aufbereiteten Elemente zu einem gewissen Grad selber erschließen. Was gehört zum Hintergrundbild und welche Elemente liegen im Vordergrund und sind für die Kombinatorik des Spiels nutzbar? Während die grafische Darstellung den Raum des Befragbaren konstituiert, sind akustische Elemente ein wichtiger Teil der Antwort auf die durchgeführten Tests. Alles Hörbare wird darauf befragt, welche Resultate bestimmte Aktionen hervorbringen oder welche Ereignisse in der Spielwelt stattfinden.

Das Testen und Experimentieren nimmt in der Nutzung von Computerspielen so eine zentrale Stellung ein, dass einige Titel es bewußt aufgreifen und reflektieren. Der Platformer *Super Meat Boy* zeigt nach jedem abgeschlossenen Level z. B. eine Wiederholung aller Spielversuche, die die Spieler benötigt haben. Die Wiederholung zeichnet sämtliche Versuche und Momente des Scheiterns nach, enthält gleichzeitig aber auch den erfolgreichen Durchlauf. Die Thematik des Testens spielt auch in der *Portal*-Reihe eine zentrale Rolle. In *Portal* durchlaufen die Spieler mehrere Kammern der testbesessenen Künstlichen Intelligenz Glados.

⁴⁶„Friendly fire“ bezeichnet in Computerspielen die Möglichkeit, verbündeten Einheiten Schaden zuzufügen.

Zu Beginn erweckt die KI den Anschein, den Spielern beim Lösen der Physik-Puzzle behilflich zu sein, entpuppt sich jedoch später als zynischer und tödlicher Antagonist. Die Obsession der Maschine für das Testen um des Testens Willen offenbart einen latenten Neid auf die Testfähigkeit der Spieler. Es ist gerade die fehlerbehaftete Menschlichkeit der Spieler, die Glados Neugier befeuert und im Spiel kontinuierlich mit der sarkastischen Sachlichkeit der KI kontrastiert wird. In diesem Zug karikiert das Spiel Belohnungsmechanismen, die Computerspiele nach erfolgreich abgeschlossenen Herausforderungen abspulen. So lobt Glados die Spieler mit generischen, unaufgelösten Sätzen wie "Unbelievable. You, [subject name here], must be the pride of [subject hometown here]!". Portal vermittelt damit ein grundlegendes Prinzip der Beziehung zwischen Spielern und Computerspielen: Es sind nicht nur die Spieler, die ihre Programme testen, sondern auch die Programme, die die Spieler testen. Die Spieler sind gleichzeitig Versuchsleiter und Probanden.

Das Ausloten des Handlungsspielraums durch das Testen bringt eine bestimmte Form der Zeicheninterpretation mit sich. Die Spieler differenzieren zwischen Veränderbarem und Unveränderbarem, Dynamik und Stasis. Durch das Testen erschließen sie Verknüpfungen zwischen den Zeichen und den darunter liegenden Modellen der dargestellten Objekte. Einerseits gibt es Darstellungselemente, die entweder die Möglichkeit einer Zustandsveränderung suggerieren, oder einen bereits durchgeführten Änderungsversuch positiv oder negativ quittieren. Darunter fallen insbesondere alle Spielfiguren und Gegenstände, die die Spieler steuern oder mit denen sie in irgendeiner Form interagieren können. Bewegen die Spieler ihre Avatare, so ändert sich der Zustand ihrer Koordinaten. Entnehmen sie Gegenstände aus einer Kiste, verändern sie sowohl den Zustand der Kiste, als auch den des Avataren-Inventars, das nun diese Gegenstände enthält. Können die Spieler eine verschlossene Kiste nicht öffnen, bekommen sie eine Rückmeldung, die ihnen diese Zustandsmanipulation versagt. Sämtliche Darstellungselemente dieser Kategorie, im Folgenden *zustandsvariant* genannt, bieten einen Zugriff auf ausgewählte Spielzustände oder Informationen über ihren Status.

Andererseits enthält das Spiel in der Regel auch Darstellungselemente, die innerhalb bestimmter zeitlich-räumlicher Grenzen unveränderbar – *zustandsinvariant* – bleiben und nicht durch die Aktionen der Spieler beeinflusst werden können. Hierzu gehören z. B. Hintergrundbilder und Ambient-Sounds sowie unverrückbare Einrichtungsgegenstände in

dreidimensionalen Umgebungen. Im Grunde fällt darunter die Kulisse mit all ihren Requisiten, die entweder in Abschnitten oder im gesamten Rahmen der Spielhandlung nicht nutzbar sind. Es handelt sich hierbei um Darstellungselemente, die nicht auf mögliche Zustandsänderungen verweisen, oder für die der Zugriff zeitweise blockiert ist.

Während des Spielverlaufs können Verschiebungen zwischen beiden Kategorien auftreten. Ein zustandsinvariantes Darstellungselement kann zustandsvariant werden und umgekehrt. Ein typisches Beispiel hierfür sind Boss-Fights. Während des finalen Kampfs in *Batman Arkham Asylum* zieht sich z. B. der Joker immer wieder in den Hintergrund zurück und bleibt unantastbar, bis Batman die nächste Welle seiner Schergen bezwungen hat. Tritt der Joker wieder in den Ring, so sind dessen Eigenschaften wie der modellierte Gesundheitszustand ab diesem Moment veränderbar. Narrative Plausibilität erhalten solche Szenen häufig durch den Topos der Unverwundbarkeit. Zuerst müssen die Schildgeneratoren deaktiviert werden, bevor das eigentliche Ziel verwundbar wird. Es handelt sich um starre narrative Strukturen, deren strikte Kausalität sie wie Algorithmen erscheinen lässt, die die Spieler zur Problemlösung anwenden müssen.

Die Unterscheidung zwischen zustandsvarianten und zustandsinvarianten Darstellungselementen ist ein integraler Bestandteil der Auslotung des Handlungsspielraums. Um überhaupt spielen zu können, müssen die Spieler wissen, welche Spielobjekte mit den ihnen zur Verfügung stehenden Aktionen benutzbar und manipulierbar sind. Erst wenn die Spieler den Handlungsspielraum kennen, sind sie in der Lage, übergeordnete Handlungsschemata und Strategien zu verfolgen. Dann setzt eine dritte Lernphase ein, in der die Auseinandersetzung mit den Makrostrukturen des Spielsystems stattfindet. Je nach den Freiräumen, die das Computerspiel bietet und der individuellen Ausnutzung dieser Freiräume, entwickeln die Spieler unterschiedliche Spielweisen. Sie unterscheiden sich nicht nur in der Wahl von Aktionen und ihrer Ausführungszeitpunkte. Vielmehr umfassen sie bestimmte Haltungen und Interessenslagen, mit welchen die Spieler dem Programm begegnen und die ihre Auseinandersetzung mit der sinnlich wahrnehmbaren Form und damit auch sämtlichen akustischen Elementen maßgeblich prägen. Das gleiche Spiel kann von unterschiedlichen Spielern vollkommen anders wahrgenommen werden. Janet H. Murray demonstriert die unterschiedlichen Auffassungsweisen am Beispiel eines Ehepaars, das *Sim City* spielt:

For the husband, the program was a satisfyingly complex engineering problem, reinforcing his habitual sense of competence. For the wife, it was a narrative, in which the little parades and cheers of her contended townfolk were the most memorable dramatic events.⁴⁷

Im Folgenden werden zwei idealtypische Spielweisen und die mit ihnen einhergehenden Auseinandersetzungen mit der Oberfläche des Programms skizziert.

Spiele nach Zahlen

Die erste idealtypische Spielweise, welche im Folgenden als das *Spiele nach Zahlen* bezeichnet wird, ist am stärksten in kompetitiven Spielsituationen ausgeprägt. Übergeordnete Ziele sind das Gewinnen einer Partie oder das Erreichen einer möglichst hohen Punktzahl. Die Spieler richten ihre Eingaben und ihre Wahrnehmung der Ausgaben darauf aus, möglichst erfolgreich zu sein.

Die Affinität kompetitiver Computerspiele zu Zahlen zeigt sich nicht nur in der Aufzählung von Ergebnissen und Gesamtpunktzahlen, sondern in zunehmend feiner segmentierten Messungen der Spielperformance. Detaillierte Statistiken wie in *Battlefield 3* schlüsseln die Erfolge, aber auch die Niederlagen der Spieler in etlichen Teilaspekten des Spiels penibel auf. Abbildung 1.3 auf der nächsten Seite zeigt einen kleinen Ausschnitt einer Spieler-Statistik von der Plattform bf3stats.com, die diesen Umstand verdeutlicht. Im Genre des Echtzeit-Strategiespiels, das zu den wichtigsten *E-sport* Disziplinen gehört, etablieren sich sogar eigene Maßeinheiten. Der Wert für die Aktionen pro Minuten, engl. *Actions per Minute* (APM), dient als Indikator für die Fähigkeiten eines Spielers. Während Anfänger in der Regel deutlich unter 100 APM bleiben, erreichen Profis Werte von über 400 APM. *Starcraft II* ermöglicht den Zugriff auf die Messungen direkt im Beobachtermodus, der zur Kommentierung von Profispielen zur Anwendung kommt. Seit dem Patch 1.4.3 enthält das Spiel neben den APM außerdem die Einheit EPM (Effective Actions per Minute)⁴⁸, die nicht wahllos alle Klicks in der Minute zählt, sondern nur tatsächliche Befehle berücksichtigt.

⁴⁷(Murray, 1997, 88)

⁴⁸<http://us.battle.net/sc2/en/blog/4487503> (10.04.2012)

Vieles spricht dafür, dass der Modus Spielen nach Zahlen mit einem höheren Grad an Professionalität einhergeht. Je gezielter die Spieler alle spielrelevanten Parameter fokussieren und Referenzen, die darüber hinausgehen, ausblenden, desto höher sind ihre Erfolgsaussichten. Für die Oberflächenwahrnehmung ergeben sich daraus weitreichende Konsequenzen. Auf visueller Ebene treten Texturen und ihre Verweise auf etwas, das außerhalb des Spiels liegt, weit in den Hintergrund. An Stelle einer maroden Fabrikhalle, deren Interieur an die Arbeitsbedingungen des industriellen Zeitalters erinnert, sehen die kompetitiv ausgerichteten Spieler die Geometrie eines Levels. Wo narrativ orientierte Spieler eine idyllische Weide erblicken, erstrecken sich vor den Augen der nach Zahlen Spielenden eine ungünstige, möglichst zu meidende Fläche ohne Deckungsmöglichkeiten. Die Spieler, die den Zeichen folgen, tragen Rüstungen, die ihnen optisch gefallen. Spieler nach Zahlen bevorzugen Gewänder, deren Rüstungswerte und Resistenzen sorgfältig mit den Verteidigungswerten ihrer übrigen Ausrüstung abgestimmt sind. Auch die Wahrnehmung der akustischen Ausgestaltung konzentriert sich bei dieser Spielform auf funktionale Aspekte. Nur ihre Referenzen auf den Status zweckdienlicher Spielzustände finden Beachtung. Alle Klangschichten, die dieses Kriterium nicht erfüllen, werden, wenn möglich, abgeschaltet.⁴⁹ Alle Präferenzen und Sichtweisen dieser Spielform sind letztendlich auf Zahlen zurück zu führen, auf Wahrscheinlichkeiten und Zustände, deren Kenntnisse den Spielern einen größeren Erfolg versprechen. Sowohl ihre Wahrnehmung, als auch ihre Aktionsausführung folgt einer größtmöglichen Effizienz im Rahmen des regelbasierten Systems.

General stats	
Rounds played	23
Finished rounds	23
Skill	73
Wins	10
Losses	13
W/L Ratio	0.769
Combat stats	
Kills	136
Deaths	280
K/D Ratio	0.486
Headshots	26
Longest headshot	127.72m
Longst handheld HS	127.72m
Best kill streak	4
Vehicles destroyed	8
Shots	12 200
Hits	1 359
Accuracy	11.14%
Savior kills	21
Avenger kills	6
Nemesis kills	0
Best nemesis streak	3

Abbildung 1.3: Statistik aus *Battlefield 3*

⁴⁹siehe auch Kapitel 6

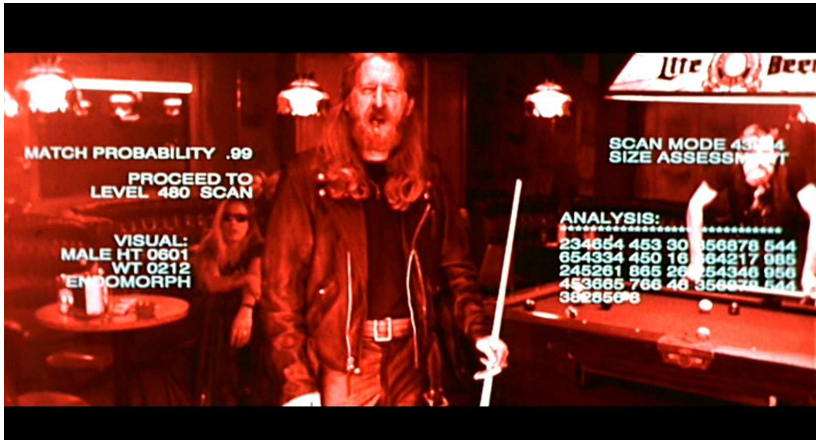


Abbildung 1.4: Umgebungsscan des Terminators

Die nach Zahlen Spielenden sind ständig auf der Suche nach dem Modell hinter den Zeichen der Oberfläche. Sie interessieren sich weniger für die Gestaltungen von Spielobjekten, als deren Eigenschaften und deren Nutzen im Rahmen des Spielsystems. Insofern besteht eine frappierende Ähnlichkeit zwischen der Oberflächenwahrnehmung der Spieler, die nach den Zahlen des Modells greifen und dem in *Terminator 2* inszenierten Umgebungsscans des T 800 (Abbildung 1.4). Bis zu einem gewissen Grad adaptieren die kompetitiv ausgerichteten Spieler die Wahrnehmungsmechanismen des Cyborgs, um in einer maschinell modellierten Welt zu bestehen.

Das kompetitive Spiel ist nicht auf Computerspiele beschränkt, die speziell dafür ausgelegt sind. Auch in Single-Player-Titeln kann die Handlungsweise kompetitiven Vorgaben folgen, die die Spieler sich selbst auferlegen. Ein prominentes Beispiel hierfür ist der „Speedrun“. Das Ziel dieser sehr fordernden Spielform besteht darin, das gesamte Spiel oder einen Abschnitt daraus so schnell wie möglich durchzuspielen. Was zählt, ist der Wettkampf um die beste Zeit durch ein virtuoses Spiel. Auf Internet-Plattformen wie Youtube und Community-Seiten veröffentlichen die Spieler anschließend ihre erfolgreichsten Versuche. Wenn die Spieler auf diese Weise von der intendierten Nutzung ihres Aktionssatzes abweichen, entstehen häufig Verzerrungen in der Form.

Zwischensequenzen werden weggeklickt, Dialoge übersprungen, Gegner ignoriert und Gegenstände liegen gelassen.

Der Medienforscher David Myers beschreibt die Schwerpunktsetzung des kompetitiven Spiels als Zeichenprozess, der die spielrelevanten Werte in den Mittelpunkt rückt und durch den die Oberfläche einen fundamentalen Wandel durchläuft.

Computer Games are meaning-making machines that, during repeated and recursive play, devalue real-world values of signs and symbols and, during that process, substitute values more relevant to the structure of the game system and, most particularly, more relevant to the identification and construction of the most desired game objects.⁵⁰

Für kompetitive Spieler, die darauf aus sind, zu gewinnen, ist die Beschreibung dieser Zeichenprozesse sehr zutreffend. Anzulasten bleibt Myers allerdings, dass er die Zeichenprozesse als inheräntes Phänomen der Form des Computerspiels postuliert und mehrere mögliche Nutzungs- und Wahrnehmungsmodi außer Acht lässt. Wie die nachfolgende Spielweise, das *Spielen nach Zeichen*, zeigen wird, müssen die Zeichen in der Wahrnehmung der Spieler diese Transformation nicht durchlaufen. Sie bieten vielmehr überlagernde Verweise, die je nach der Interessenslage der Spieler verfolgt oder vernachlässigt werden.

Spielen nach Zeichen

Das Spielen nach Zeichen orientiert sich nicht in erster Linie an dem Spiel und seiner Gewinn- und Verlustrechnung. Statt dessen richtet es sich nach der Erzählung und den narrativen Verweisen der dargestellten Zeichen. Es steht der Interpretation und dem Verfolgen einer Geschichte näher als dem Wettstreit.

Während das Spiel nach Zahlen Verknüpfungen zwischen den Zeichen und den durch Zahlen beschriebenen Zuständen des Modells sucht, sind diese Verknüpfungen für das Spiel nach Zeichen zweitrangig. Zustandsvariante und -invariante sowie spielrelevante und -irrelevante Inhalte

⁵⁰(Myers, 2010, 158)

des Computerspiels sind gleich gewichtet. Alle Ausgaben sind potentiell sinnstiftende Verweise. Das Erlebnis tritt an die Stelle der Herausforderung. Vor dem Hintergrund dieser von den Regeln des Spiels befreiten Auseinandersetzung mit den Angeboten des Computerspiels werden auch Interaktionsmöglichkeiten interessant, die nicht an das Spielsystem gekoppelt sind.

Vom First Person Shooter⁵¹ bis zum Rollenspiel enthalten Computerspiele in vielen Genres optionale Interaktionselemente, die keinen Einfluss auf den Spielausgang haben. Optionale Welt-Interaktionen dieser Art verändern in der programminternen Modellierung des Programms keine Zustände, die im Kontext des Spielsystems Sinn ergeben. Die anvisierten Zustände sind vielmehr einfache Schalter, die Oberflächeneffekte wie Animationen, Sounds und Sprachsamples triggern. In der Logik eines First Person Shooters ist es nicht zweckdienlich, Hau-den-Lukas auf dem Jahrmarkt zu spielen, ein Waschbecken zu benutzen, oder eine Jukebox anzuschalten. Weder die kompetitiv ausgerichteten Spieler noch Künstliche Intelligenzen nutzen diese Möglichkeiten. Für das Spielen nach Zeichen sind diese Handlungsmöglichkeiten hingegen wichtige Zusatzangebote. Sie schaffen einen größeren Handlungsspielraum, in dem die Aktionen der Spieler sich in suggerierten oder selbst erdachten Inszenierungen entfalten.

Ebenso wie das Spielen nach Zahlen in Computerspielen mit einem hohen erzählerischen Anteil möglich ist, kann das Spielen nach Zeichen in kompetitiven Spielen verfolgt werden. Ein Beispiel dafür, wie stark sich selbst verhältnismäßig abstrakt gehaltene Spiele für das Spiel nach Zeichen eignen, bietet die folgende Konversation aus der dritten Episode der ersten Staffel der Serie *The Wire*. D'Angelo Barksdale, der als Cousin des Drogenbosses Avon Barksdale die Distribution des Stoffs in einem Viertel der Stadt beaufsichtigt, erklärt zweien seiner Unterhändlern das Spiel der Könige.

D'Angelo: Check it, it's simple. See this? This is the kingpin. And he the man. You get the other dude's king, you got the game. But he's trying to get your king, too, so you gotta protect it. Now the king, he move one space any direction he damn choose, cos he's the king. Like this, all right? But he ain't got no hustle. The rest of these motherfuckers on

⁵¹Bereits der 1996 erschienene Shooter *Duke Nukem 3D* enthielt diverse Interaktionsmöglichkeiten, die im Rahmen des Spielsystems keine Funktion erfüllten.

the team got his back. And they run so deep, he really ain't gotta do shit.

Bodie: Like your uncle.

D'Angelo: Like my uncle. You see this? This is the queen. She smart and she fierce. She moves any way she want, as far as she want. And she is the "go get shit done" piece.

Wallace: Remind me of Stringer.

D'Angelo: And this over here is the castle. It's like the stash. It moves like this and like this.

Wallace: Dog, stash don't move, man.

D'Angelo: Think. How many times we move the stash house this week? Every time we move the stash, we got to move a little muscle with it to protect it.

Bodie: True. You right. What about them little bald-headed bitches right there?

D'Angelo: These right here? These are the pawns. They like the soldiers. They move like this, one space forward only, except when they fight. Then it's like... Or like this. They the front lines. They be out in the field.

Wallace: So how do you get to be the king?

D'Angelo: It ain't like that. The king stay the king, all right? Everything stays who he is, except for the pawns. If a pawn make it all the way down to the other dude's side, he get to be queen. And like I said, the queen ain't no bitch. She got all the moves.

Bodie: All right, so, if I make it to the other end, I win?

D'Angelo: If you catch the other dude's king and trap it, then you win.

Bodie: But if I make it to the end.. I'm top dog.

D'Angelo: It ain't like that. The pawns, man, in the game, they get capped quick. They be out the game early.

Bodie: Unless they're some smart-ass pawns.

D'Angelo stellt eine Analogie zwischen dem Spiel und dem Drogengeschäft her und erklärt mit ihrer Hilfe die Regeln des Spiels. Wallace und Bodie erkennen in den als Soldaten gekennzeichneten Bauern ihre eigene Rolle wieder und projizieren ihre Wünsche auf die Ebene des Spiels. Sie fragen sich, wie man zum König aufsteigt oder zumindest „top dog“ wird. Die Metaphorik der Spielfiguren spielt für sie in diesem Moment eine größere Rolle als die eigentlichen Regeln. Sie sind stärker an dem Wohlergehen der Bauern und den Möglichkeiten ihres Aufstiegs innerhalb der Organisation interessiert als an allem andern. Auch das Spiel wird unter diesen Prämissen zu einer Erzählung. Es erzählt von der Hierarchie des Drogenrings, den unterschiedlichen Rollen seiner Mitglieder und den ungeschriebenen Regeln des Geschäfts. Derjenige an der Spitze bleibt an der Spitze, und diejenigen in den unteren Positionen werden schnell geschlagen.

Auf ähnliche Weise wie Wallace und Bodie können auch Computerspieler den narrativen und metaphorischen Angeboten der Darstellungen begegnen. Der Späher in *Civilization* ist dann nicht nur eine Einheit, welche die verborgenen Felder des Spielfelds aufdeckt und pro Runde zwei Felder weit bewegt werden darf. Möglicherweise sehen die Spieler in ihm einen Wanderer und Abenteurer, der eine gefährliche Reise durch unerforschte Länder antritt. Vielleicht identifizieren sie sich mit ihm so stark, dass sie in Hinblick auf das übergeordnete Spielziel schlechte Entscheidungen treffen, nur um ihn im Spiel zu behalten.

Das Spiel nach Zahlen und das Spiel nach Zeichen ergeben sich aus den in Abschnitt 1.2 analysierten semiotischen Eigenschaften der Spieldarstellungen. Die Kenntnis der Modelle hinter den Zeichen ist essentiell für den Erfolg im Spiel. Die Zeichen hingegen schaffen Raum für Erzählungen, die in vielen Fällen durch an das Spiel angelagerte Geschichten konkretisiert werden. Beide Spielweisen treten potentiell in unterschiedlichen Mischungsverhältnissen auf. Die Spieler können je nach Interesse und Spielsituation die Schwerpunkte verlagern.

Emergentes Gameplay

Die verschiedenen Herangehensweisen, mit denen die Spieler den Angeboten des Computerspiels begegnen können, erzeugen eine Reibungs-

fläche zwischen den Intentionen der Entwickler und dem tatsächlichen Umgang durch die Nutzer. Das typische Verhalten in bestimmten Spielsituationen wird häufig in Form von angeleiteten Spieleinführungen – dem *Tutorial* oder *How to play* – an die Spieler herangetragen. In manchen Fällen empfehlen Einführungstexte explizit eine bestimmte Spielweise. Die Independent-Produktion *Amnesia: The Dark Descent* lädt beim ersten Start eine Nachricht der Entwickler, die den Spielern ans Herz legt, nicht zu spielen, um zu gewinnen. In anderen Fällen legen die Game Designer durch die Regelsätze implizit fest, wie die Spieler die ihnen zur Verfügung stehenden Aktionen anwenden sollten. Je nachdem, wie die Regeln konzipiert sind, werden bestimmte Handlungsschemata durch das Spielsystem belohnt und andere bestraft. Ein Beispiel hierfür ist der Zwang zum Schleichen, langsamen Vorgehen oder zur Flucht durch eine unbezwingbare gegnerische Übermacht. Ob sich die Spieler darauf einlassen oder im Rahmen der gegebenen Möglichkeiten nach ihren eigenen Regeln spielen, bleibt jedoch ihnen überlassen. In vielen Fällen gerät das Experimentieren mit den Möglichkeiten des Computerspiels zu einer eigenständig motivierten Form der Auseinandersetzung, die neue Spielweisen hervorbringt. Innerhalb der Game Studies spricht man in diesen Fällen, die von dem Vorhergesehenen abweichen, von nicht-intentionalem *emergentem Gameplay*⁵².

Die Erfahrung, dass Software auch auf andere Weise verwendet wird, als ursprünglich vorgesehen, mußte Microsoft im Jahr 2008 mit seinem Musikgenerationsprogramm *Songsmith* machen. Auf Basis einer eingesungenen Melodie und eines ausgewählten Stils erstellt das Programm ein instrumentales Hintergrundarrangement. Das offizielle Werbevideo⁵³ zeigt eine singende Familie, die mit Hilfe des Programms eigene Stücke entwirft. Große Bekanntheit erlangte *Songsmith* jedoch nicht durch seinen vorgesehenen Zweck als Kompositionshilfe für Laienmusiker, sondern als Werkzeug zur Neuinterpretation älterer Hits. Benutzer extrahierten die Gesangsstimmen aus populären Aufnahmen, speisten sie als neuen Input in *Songsmith* ein, und stellten die Ergebnisse bei *youtube* online.⁵⁴ Aus dem Metal-Klassiker *Ace of Spades* von Motörhead wird so ein seichter Country-Song und Eminem wirkt mit

⁵²siehe (Juul, 2005, 75)

⁵³<http://www.youtube.com/watch?v=3oGFogwcx-E> (28.09.2012)

⁵⁴Mittlerweile ist dieses Internet-Phänomen durch den juristischen Druck der Rechteinhaber weitgehend verdrängt worden. Viele der beliebtesten *Songsmith*-Versionen mit mehreren hunderttausend Views wurden aus *youtube* entfernt.

einer Polka-Begleitung einem Kommentar zufolge wie ein „Volksmusikant aus Kasachstan“.

Ebenso wie Benutzer Programme wie *Songsmith* unterlaufen, experimentieren die Spieler mit den komplexen Systemen, die Computerspiele bereitstellen. In einigen Fällen beugen sie sich den Vorgaben des Spiels. In anderen Fällen weichen sie massiv von dem Umgang ab, der den Entwicklern vorschwebte. Im Gegensatz zu Anwendungsprogrammen wie *Songsmith* ist der kreative Umgang mit Computerspielen weitgehend akzeptiert.⁵⁵ In Anlehnung an den Slogan eines Grafikkartenherstellers nennt die Webseite tvtropes.org dieses Phänomen „Not the way it is meant to be played“:

Some people will take a Video Game and do everything you can possibly imagine with it... except playing it as they were supposed to. [...] Over time, many of these subversions get acknowledged, expected, encouraged and finally assimilated into new games, becoming part of standard gameplay.⁵⁶

Ergänzend zu den beiden vorgestellten Umgangsmöglichkeiten mit den Zeichen, die entweder auf die dahinter liegenden Modelle zielen, oder der Metaphorik der Zeichen folgen, entstehen durch das emergente Gameplay weitere Umgangsformen mit ihnen. So lassen sich z. B. in einem Shooter musikalische Erfahrungen hervorbringen, indem Rhythmen geschossen werden und die Spieler nach einem gezielten Arrangement der von Aktionen getriggerten Samples trachten. Ein weiteres Beispiel für die unkonventionelleren Umgangsmöglichkeiten bietet der Trance-Vibrator des PS2-Spiels *Rez*, auf dessen Verwendbarkeit zur körperlichen Stimulation innerhalb der Netzgemeinschaft mehrfach hingewiesen wird.

That's why I was so excited by *Rez's* trance vibrator, since it seems to have no other purpose than to act as a masturbatory aid. Its shape is pretty nice, it can slip easily under your skirt or in your panties, it comes with a protective "glove" which you can wash, and it emits a regular pulsating

⁵⁵Eine Antwort auf die Formen emergenten Gameplays ist die Entstehung des Sandbox-Genres. Programme wie *Garry's Mod* zielen nicht auf die Bereitstellung eines komplexen Spielsystems oder einer packenden Geschichte, sondern bieten statt dessen freie Experimentierstätten.

⁵⁶<http://tvtropes.org/pmwiki/pmwiki.php/Main/NotTheWayItIsMeantToBePlayed> (30.05.2013)

rhythm that gets ever more intense and thrilling the deeper you go into the game. Damn, by the end I was writhing on the floor! Synesthesia indeed.⁵⁷

Je nachdem, wie das Angebot aussieht, liegen bestimmte Spielweisen näher als andere.

⁵⁷<http://www.gamegirladvance.com/2002/10/sex-in-games-rezvibrator.html#000141/> (30.05.2013)

2 Akustische Bausteine im Wandel der Zeit

Wie kein anderes Ausdrucksmittel ist das Computerspiel von technologisch bedingten Veränderungen geprägt, die sich nicht zuletzt in der Gestalt der Ausgaben - dem, was man über den Bildschirm sieht, die Lautsprecher hört und das Force Feedback Gamepad spürt - niederschlagen. Mit jeder neuen Generation leistungsfähigerer Hardware-Komponenten und weiterentwickelter Software-Tools steigt die Meßlatte des Darstellbaren. In der Tat scheint in der nunmehr über 40-jährigen Geschichte des Computerspiels diese kontinuierliche qualitative Veränderung des Dargestellten eine ihrer auffälligsten Konstanten zu sein. So ließ sich die seit den 1970er Jahren im Bereich stationärer Computer vollzogene Transformation von grob aufgelösten Pixelfiguren und piepsenden Klängen bis hin zu detaillierten 3D-Modellen und hochauflösendem Sound im ersten Jahrzehnt des 21. Jahrhunderts in einer rapide beschleunigten zweiten Iteration auf mobilen Endgeräten beobachten.

In Hinblick auf die akustische Gestaltung spiegeln sich die deutlichsten Transformationen dieser Entwicklung in der Form des Klangmaterials wider. Bevor bestimmte akustische Inhalte wie Musik und Sprache in das Computerspiel integriert werden können, müssen die technischen Voraussetzungen dafür erfüllt sein. Kenntnisse der Darstellungsmöglichkeiten und -limitierungen der technischen Basis sind unentbehrlich, um deren Einfluss auf die Gestaltung angemessen zu berücksichtigen. Jedes entwickelte Programm bewegt sich innerhalb der Grenzen seiner zugrunde liegenden technisch determinierten Ausdrucksmöglichkeiten und den damit verbundenen Paradigmen. Ein erstes Ziel besteht deshalb darin, in einem chronologischen Abriss verschiedener Entwicklungsstufen die wichtigsten technischen Umbrüche zu skizzieren, die neue oder veränderte Formen der akustischen Gestaltung mit sich brachten. Im Mittelpunkt der Betrachtung steht das Klangmaterial, welches im Rahmen der Architektur-Metapher die Bausteine konstituiert. Ein

akustischer Baustein bezeichnet im Folgenden ein für die Ausgabe definiertes Schallereignis - sei es in Form einer digitalisierten Aufnahme oder eines Stücks Programmcode.

Die Fülle verschiedener Plattformen gebietet eine exemplarische Darstellung. Die im Folgenden vorgenommene historische Konturierung der Sound-Technologien des Computerspiels von 1971 bis 2011 konzentriert sich deshalb auf zwei technisch bedingte, für die akustische Gestaltung besonders folgenreiche Umbrüche: den Wechsel vom Hardware-Design zum programmierbaren Sound in Software und den Paradigmenwechsel der Klangerzeugung von der Synthese zum Sampling. Die Analyse dieser Umbrüche soll die Einordnung ermöglichen, welche akustischen Gestaltungsformen den technischen Grundbedingungen geschuldet, und welche als freie ästhetische Entscheidungen innerhalb der jeweils gegebenen technischen Mittel und Möglichkeiten zu verstehen sind. In diesem Zusammenhang stellt sich die Frage, welche gestalterischen Kompensationsstrategien in den jeweils abgeschrittenen Entwicklungsstufen als Antwort auf technische Limitierungen zum Ausdruck kommen.

Ein weiterer Aspekt, der im Zuge des Abschreitens dieser technischen Rahmenbedingungen der akustischen Gestaltung in den Blick genommen werden soll, ist die Auswirkung eines zunehmend komplexeren Hörangebots auf die Spielerfahrung. Offensichtlich gibt es eine große Diskrepanz zwischen der Wahrnehmung der jeweils aktuellen Darstellungsformen und zurückliegenden Ausprägungen, die plötzlich von geringerer Qualität erscheinen.

2.1 Frühe Formen der Vertonung

The computer even had an audio output: while the program ran, a speaker underneath the console would make a sort of music, like a poorly tuned electric organ whose notes would vibrate with a fuzzy, ethereal din. The chords on this "organ" would change, depending on what data the machine was reading at any given microsecond; after you were familiar with the tones, you could actually HEAR what part of your program the computer was working on.¹

¹(Levy, 1984, 15)

Die ersten Vorkommnisse akustischer Elemente als Ausgaben des Computers waren geprägt von der Nähe zur Maschine. Das galt sowohl für den oben von Steven Levy beschriebenen TX-0, der ab Mitte der 1950er Jahre am MIT zum Einsatz kam, als auch für die ersten Spielautomaten.

Sound der Spielhallen

Während die ersten Formen von dem, was wir heute Computerspiel nennen, bereits eine Dekade früher als vereinzelte Experimente in Forschungseinrichtungen und Labors ihren Ursprung fanden, entstanden die ersten Spiele mit akustischen Inhalten im Zuge der Kommerzialisierung des Computerspiels Anfang der 1970er Jahre. Maßgeblichen Anteil an der Popularität der neuen Unterhaltungsmaschinen hatten Nolan Bushnell und sein Partner Ted Dabney, die 1971 mit *Computer Space*, einer Weiterentwicklung des 1962 am MIT entworfenen Spiels *Space War!*, den ersten münzbetriebenen Computerspielautomaten konstruierten. Ein Jahr später gründeten sie die Firma Atari und veröffentlichten mit *Pong* ein überaus erfolgreiches Spiel, dessen Wurzeln ebenfalls weiter zurückreichten. Bereits im Jahr 1958 konstruierte der Physiker William Higinbotham die Simulation *Tennis for two*, die auf einem Oszillator die Seitenansicht eines über zwei Potentiometer steuerbaren Tennismatches skizzierte.² Nachdem *Tennis for two* für die Besucher des Brookhaven National Laboratory am Tag der offenen Tür einen Höhepunkt darstellte, eroberte *Pong* Anfang der 1970er Jahre mit der gleichen Multiplayer-Mechanik die Spielhallen im Sturm.

Mit der Positionierung des Computerspiels als Vergnügungsgerät der Penny Arcades trat die neue Spielmaschine in eine Reihe zu anderen münzbetriebenen Attraktionen des Nachtlebens öffentlicher Räume wie Billardtischen, elektromechanischen Spielen und Pinballautomaten. Im Gegensatz zu jenen, die auf materiellen Spielgegenständen wie Bällen und Kugeln, ihren jeweiligen physikalischen Eigenschaften sowie elektromechanischen Prinzipien basieren, operiert das Computerspiel in seinem virtuellen Raum mit keinerlei materiellen Dingen mehr. Es gibt keinen Ball, der über Holz und Metall gleitet, und keine auf Federn oder Gummibändern basierenden Komponenten wie Bumper und Slingshots, die ihn bei Berührung fortschleudern könnten. Statt dessen

²(Pias, 2010, 13)

entfaltet sich das gesamte, mit hohem technischem Aufwand realisierte Geschehen auf der flachen Scheibe eines Bildschirms oder Fernsehers, der das Spiel als *Video*-Spiel von den übrigen Geräten abgrenzt. Natürliche Schallquellen durch Spielgegenstände sind damit ausgeschlossen. Sowohl das visuelle Erscheinungsbild und die Geräusche der dargestellten Dinge, als auch deren Verhalten und deren Umwelt sind nicht unmittelbar an physische und physikalische Vorgaben der Außenwelt gebunden, sondern im Rahmen ihrer technisch vollzogenen Modellierung frei setzbar.

Bereits elektromechanische Spiele und Pinballautomaten enthielten nicht nur die Geräusche ihrer Mechanik, sondern verfügten über künstlich erzeugte Klänge, die von Ingenieuren mit Hilfe elektrischer Schaltkreise in die Automaten integriert wurden.³ Die Adaptierung dieses Sound-Designs auf Computerspiele war neben der Unterbringung der neuen Technik in mannshohen Gehäusen die logische Konsequenz, um die Erwartungen der Spieler an einen Arcade-Automaten zu erfüllen. Nicht nur das Image des Computerspiels als unbedarftes Unterhaltungsgerät, das Jugendliche in zwielichtigen Spelunken und Spielhallen der Verwahrlosung aussetzt, sondern auch die Grundzüge früher klanglicher Gestaltung sind auf die Einschreibung in diese Tradition zurückzuführen.

Da die ersten in Serie produzierten Mikroprozessoren wie der Intel 4004 und Intel 8080 für die junge Videospieleindustrie nicht erschwinglich waren, griff man für die technische Realisierung der Automaten auf Schaltkreise der Transistor-Transistor-Logik (TTL) zurück, die per Hand zusammengelötet wurden. Die Anordnung und Verbindung der TTL-Chips repräsentierte dabei sämtliche Aspekte des Spiels. Sowohl die für den Spieler wahrnehmbare Darstellung, als auch deren interne Modellierung waren vollständig in Hardware abgebildet. Schraubte man einen Automaten auf, so konnte ein versierter Ingenieur mit bloßem Auge aus den Schaltkreisen der Platine die Konstruktionsweise des Spiels rekonstruieren. Eine frühe Form des Kopierschutzes bestand darin, die Identifikationsnummern der integrierten Schaltkreise abzuschmirgeln und die Oberflächen mit silberner Farbe zu bemalen.⁴

Die Einbindung akustischer Elemente gestaltete sich innerhalb dieses reinen Hardware-Designs ebenso aufwändig wie die Gestaltung aller

³siehe (Bagnall, 2011, 41)

⁴<http://www.computerspacefan.com/Ted2.htm> (30.05.2013)

anderen Spielaspekte. Üblicherweise verwendete man klanggenerierende und -formende Einzelkomponenten wie Oszillatoren und Filter, die in ihrer jeweiligen Verschaltung individuelle Lösungen bildeten. Bereits *Computer Space* erzeugte auf diese Weise eine Klangkulisse aus schrillen Tönen und pulsierendem Rauschen. Die Technik der Klangerzeugung bestand dabei weiterhin auf analogen Syntheseverfahren, die bereits bei älteren Spiel- und Pinballautomaten zum Einsatz kamen. Während eine Welle der Digitalisierung in den 1960er Jahren analoge Computer durch digitale Pendanten ersetzte, blieben Synthesizer für musikalische Zwecke bis in die 1980er Jahre Ausnahmen, die sich weiterhin analoger Technik bedienten.⁵

Nachdem sich *Computer Space* nur zögerlich verkaufte, wurde der 1972 eingeführte Automat *Pong* ein durchschlagender Erfolg. *Pong* war dermaßen gewinnbringend, dass es heute als Grundstein der Videospieldustrie gilt. Noch im Konsolendesign des 1977 erschienenen Atari 2600 ist dessen nachhallender Einfluss unverkennbar. Die Konsole war offensichtlich dafür ausgelegt, Spiele wie *Pong* zu ermöglichen.⁶ Am Beispiel dieses Spiels sollen im Folgenden die Grundmuster zwei unterschiedlicher Vertonungsformen früher Computerspiele ausgemacht werden.

Abstrakte Repräsentation

Die akustische Gestaltung erfolgte bei der Entwicklung von *Pong* erst zu einem Zeitpunkt, als alle anderen Aspekte des Spiels fertig verschaltet waren. Da der verantwortliche Ingenieur Allan Alcorn bereits viele TTL-Chips für die Spiellogik verwendet hatte und eine möglichst geringe Anzahl verbauter Komponenten für eine günstige Produktion der Automaten unerlässlich war, entschloß er sich für eine Notlösung. Ohne die Integration zusätzlicher Schaltkreise für die Klangerzeugung fand er eine Möglichkeit, mit Hilfe der bereits verbauten Komponenten den Sound des Spiels zu produzieren. So kam es, dass das "namensgebende, onomatopoetische Pop-Art-«Pong»" im Grunde nichts anderes war als "ein extrem verstärktes Knacken im vertikalen Zeilenzähler des ebenfalls auf TTL-Chips aufgebauten Geräts".⁷ Genau genommen ist jenes kurze Knacken gleich in drei verschiedenen Varianten mit jeweils un-

⁵(Fuller, 2008, 22)

⁶(Montfort u. Bogost, 2009, 11)

⁷(Pias, 2010, 113)

terschiedlichen Tonhöhen vorhanden. Sie treten bei der Kollision des Balls mit den horizontalen und vertikalen Bildschirmrändern sowie den von zwei Spielern steuerbaren Schlägern auf (siehe Abbildung 2.1).

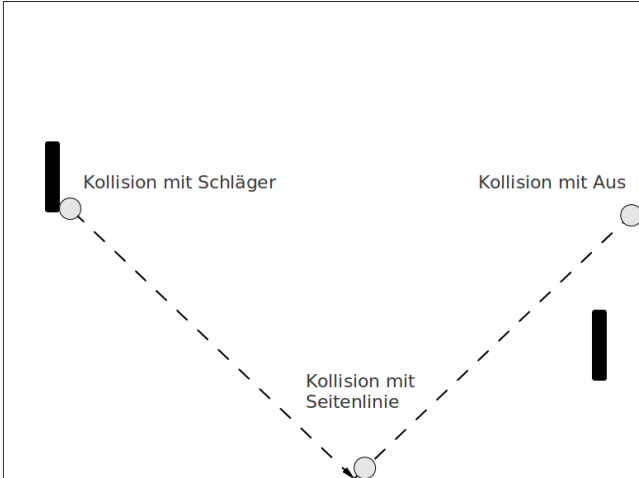


Abbildung 2.1: Kollisionsereignisse in *Pong*

Das erste kompensatorische Grundmuster früher Vertonungen des Computerspiels ist die Abstraktion. Was uns heute als Knacken erscheint, galt damals offensichtlich als passable Simulation der Kollisionsgeräusche eines abprallenden Balls. Trotz der schlichten Behelfslösung führte ein Werbeflyer aus dem Jahr 1972 den Punkt „Realistic Sounds of Ball Bouncing“⁸ als einen der Vorzüge des Geräts auf. Diese in der Marketingrhetorik von Spielepublishern bis heute⁹ gepflegte Positionierung des Dargestellten als originalgetreue Abbildung von Etwas kann zum Einen als künstlich erzeugtes Gütekriterium verstanden werden, das dem Spiel bescheinigt, eine glaubhafte Illusion zu erzeugen. Zum Anderen wirft sie aber auch Fragen bezüglich des Zusammenhangs zwischen der Repräsentation und dem Repräsentiertem auf. Inwiefern kann *Pong* als Abbildung von *Ping Pong* verstanden werden?

⁸http://flyers.arcade-museum.com/flyers_video/atari/11013001.jpg (17.10.2012)

⁹Ein Beitrag auf gameinformer.com preist z. B. das 2011 erschienene Spiel *Battlefield 3* als „most realistic sounding war experience in gaming“ (Hanson, 2011)

Geoff King und Tanya Krzywinska beschreiben den vermeintlichen Realismus der Darstellung in Computerspielen als den Grad, zu dem Repräsentationen des Spiels mit der realen Welt, Vorstellungen von Realität und anderen Repräsentationen korrelieren.¹⁰ Entscheidend für diesen Abgleichungsprozess ist nicht nur die Realerfahrung. Vielmehr sind zu einem großen Anteil unsere „Wahrnehmungsformen [...] durch Errungenschaften der Imagination und Kunst geprägt“.¹¹ Etwas als realistisch zu empfinden ist dementsprechend eng mit der Erwartungshaltung verknüpft, die durch den Vergleich des Gesehenen und Gehörten mit bereits Bekanntem entweder erfüllt oder nicht erfüllt wird. Die Summe dieser herangezogenen Vergleichswerte wird für akustische Repräsentationen im Folgenden als Klangvorstellungen bezeichnet.

Zu den Quellen der Klangvorstellungen, mit denen damalige Spieler den *Pong*-Automaten begegnet sein könnten, gehört neben den Konventionen der Arcades vor allem das Tischtennispiel. Wenn in *Pong* der Ball nacheinander von einem Schläger und einer horizontalen Seitenbegrenzung abprallt, so erklingen zwei unterschiedliche Töne. Auf diese Weise erfasst das Spiel einen Kernaspekt realer Tischtennisgeräusche, nämlich den Frequenzunterschied zwischen der Ballkollision mit dem Tischtennisschläger und der Kollision mit der Holzplatte. Auch wenn diese minimalistische Form synthetischer Klangmodellierung zahlreiche weitere Aspekte der realen Geräusche nicht fassen kann, ist eine gewisse Ähnlichkeit unverkennbar.

An diesem Beispiel lässt sich ein zentrales Prinzip erkennen, das die qualitative Dimension der Darstellung früher Computerspiele durchdringt. Sowohl die visuellen, als auch die auditiven Repräsentationsmöglichkeiten erlauben nur eine rudimentäre, auf Abstraktion beruhende Abbildung realer oder fiktiver Gegenstände. Insbesondere ist die Nachbildung von Klängen meist nur annäherungsweise möglich, indem die auffälligsten Eigenschaften ihrer Schallsignale so gut es geht simuliert werden.¹² In der Reduktion dieser künstlichen Geräusche auf ihre essentiellen Merkmale liegt aber nicht nur ein qualitativer Makel, sondern auch ein sinnlicher Reiz. Gerade wegen ihrer Schlichtheit und des Auslassens von Details ermöglichen sie den Spielern die Auseinandersetzung mit ihrer offenkundigen Metaphorik. Im Sinne dieses Mo-

¹⁰(King u. Krzywinska, 2006, 143)

¹¹(Welsch, 1998, 205)

¹²Eine Ausnahme bilden höchstens Klänge, die ohnehin synthetischen Ursprungs sind, z. B. Soundeffekte des Science-Fiction-Films.

dellierungsvorgangs ist die synthetische Klangerzeugung nicht nur ein mathematisch-physikalischer, sondern auch ein kreativer und künstlerischer Akt.

Ein weiterer bestimmender Faktor dieser Form des Sounddesigns besteht darin, dass die Technik der Klangsynthese gegenüber den abzubildenden Sounds nicht neutral ist. Es besteht eine große Abhängigkeit zwischen den klangerzeugenden und -formenden Komponenten und den damit simulierbaren Sounds. Einige Geräusche wie Wind lassen sich leichter rekonstruieren¹³, während die Klangsynthese komplexer Schallsignale wie z. B. menschlicher Sprache mit unverhältnismäßig großem Aufwand verbunden bleibt. Hier liegt ein grundlegendes Problem früher Klangerzeugung, das insbesondere auch für die programmierbaren Sound Generatoren der 1980er Jahre bestehen bleibt.

Musikalische Abstraktion

Während in *Pong* die Kollisionsvertونungen des Balls mit den Schlägern und den horizontalen Seitenbeschränkungen, von denen er abprallt, auf ihre reduzierte Art an die Geräuschvorlagen des Tischtennispiels angelehnt sind, greift bei der Kollision mit den vertikalen Seitenbeschränkungen ein anderes Schema. Auch für diesen dritten Fall, nämlich dass der Ball ins Aus geht, gab es ursprünglich konkrete Vorstellungen von Seiten des Atari-Vorstands für die Vertonung. Bereits während der Entwicklung rieten Nolan Bushnell und Ted Dabney dem zuständigen Ingenieur Allan Alcorn, das frenetische Jubeln und Zischen sowie die Buhrufe einer Menschenmasse bei einem Punktgewinn bzw. -verlust zu simulieren.¹⁴ Alcorn ersann jedoch eine pragmatischere Lösung.

I said, "Screw it, I don't know how to make any of those sounds. I don't have enough parts anyhow."¹⁵

Im Zuge seiner oben beschriebenen Notlösung wies er dem Ereignis den bereits bei den Abprallereignissen eingesetzten Klang zu, wobei

¹³Mit Hilfe eines Rauschgenerators und eines Filters lassen sich bereits glaubhafte Ergebnisse erzielen.

¹⁴siehe Alcorn-Interview in (Kent, 2001, 42)

¹⁵ebd.

er dessen Frequenz wiederum anpasste. Der dadurch entstandene dritte Ton des Spiels weist keine Verbindung zu den von Bushnell und Dabnell vorgeschlagenen Geräuschen auf. Statt dessen erschließt sich seine Bedeutung in Relation zu den anderen beiden Tönen. Während der Frequenzunterschied zwischen den beiden Abprallereignissen etwa demjenigen einer Oktave entspricht, liegt der dritte Ton einen Halbton höher. Bleibt der Ball im Spiel ergibt sich dadurch eine hörbare Kontinuität, eine Verkettung von konsonanten Tönen gleicher Länge. Landet er hingegen im Aus durch die Berührung der vertikalen Grenzen, so erklingt ein Ton, der deutlich abweicht und auf struktureller Ebene einen Bruch erzeugt. Konsonanz belohnt den erfolgreichen Ballwechsel, während Dissonanz dessen Abbruch als Fehlschlag kommentiert. Einige der in Umlauf gebrachten *Pong-Systeme* nutzten ein ähnliches Sonifikationsprinzip, indem sich statt dem Parameter der Frequenz die Länge des dritten Tons von den beiden anderen unterschied.

Mit dieser nüchternen Bestimmung einer Abweichung besitzt *Pong* die gleiche Funktionalität der Töne, wie man sie heute an Supermarktkassen wiederfindet. Die Abweichung eines Standardtons beim Scannen der Artikel tritt dort vorrangig bei altersbeschränkten Rauschmitteln wie Zigaretten und alkoholhaltigen Getränken auf. Ebenso wie die Abbildung verschiedener Lebensmittelkategorien auf verschiedene Töne die entsprechenden Paragraphen des Jugendschutzgesetzes hörbar in das Kassensystem integriert, bewerten die Kollisionstöne in *Pong* die Ergebnisse des Spiels.

Auch die beiden Abprallgeräusche erhalten durch die abstrakte, kontextabhängige Gestaltung des dritten Tons eine andere Qualität. Sie sind gleichzeitig eine vage, stark reduzierte Abbildung der Ping Pong Geräusche und Teil jenes hörbaren Codes, der den Ballwechsel an den entscheidenden Wendepunkten kommentiert. Der Sound verweist dadurch stark auf die innere Struktur, die Regeln und Gesetze des Spiels und weniger auf etwas, das außerhalb des Spiels liegt.

Das zweite kompensatorische Muster früher Computerspielvertonungen bildet also die musikalische Abstraktion. Sie stellt für das frühe Computerspiel eine häufig eingesetzte Alternative zu der annäherungsweise Abbildung eines Schallsignals dar. Ein weiteres typisches Beispiel ihrer Verwendung ist die Vertonung von Sprung- und Fallgeräuschen. Das Springen einer Spielfigur wird dabei mit einem Glissando in einen höheren Frequenzbereich vertont, während das Fallgeräusch auf ähnli-

che Weise in tiefere Frequenzbereiche gleitet. Im Zentrum dieser Vertonungsstrategie steht eine Analogie zwischen der Klangstruktur und der Ereignisstruktur, in diesem Fall eine Bewegung nach oben bzw. unten.

2.2 Programmierter Sound

In der zweiten Hälfte der 1970er Jahre wird das aus Kostengründen verfolgte Hardwaredesign des Computerspiels obsolet. Das Aufkommen preisgünstiger Mikroprozessoren ermöglicht es den Herstellern, Computerspiele als Software zu entwickeln. In Kombination mit Soundchips, speziellen Hardwarekomponenten für die synthetische Klangerzeugung, entstehen damit neue Ausgangsbedingungen für die akustische Gestaltung.

Mikroprozessoren und Soundchips

Das Zusammenlöten von Schaltkreisen für die akustische Gestaltung der Arcade-Automaten war ein aufwändiger Prozess, der aber immer nach einem ähnlichen Schema ablief. Ein Ingenieur verschaltete mehrere Basiskomponenten wie Oszillatoren und Filter miteinander, um bestimmte Klänge zu erzeugen. Diese Schaltkreise wurden dann wiederum mit anderen Bereichen des Spiels wie der Kollisionserkennung verbunden, die den Abspielmoment der Soundeffekte bestimmten. Um den Arbeitsaufwand zu verringern, der bei jedem Automaten durch das anfängliche Aufsetzen der Schaltkreise für die Klangerzeugung anfiel, fassten Elektronikhersteller bald mehrere Basiskomponenten in sogenannten *Soundchips* zusammen. Während sich an den grundlegenden Prinzipien der auf Synthese basierenden Klangerzeugung nicht viel änderte, ergab sich daraus ein organisatorischer Wandel. Die Entwickler des Spiels wählten nun die zu integrierenden Klänge aus, indem sie den Soundchip programmierten.

Als erster Hersteller brachte Atari 1977 mit dem *Video Computer System* (VCS, später Atari 2600) eine mikroprozessorbasierte Spielkonsole auf den Markt, die einen speziell entwickelten Chip für die Grafik- und

Soundausgabe enthielt.¹⁶ Das *Television Interface Adapter* (TIA) gilt damit als der erste programmierbare Soundchip. Bei einigen frühen Formen dieser Soundchips wie dem Texas Instruments SN76477, der u. a. bei *Space Invaders* zum Einsatz kam, geschah die Programmierung noch nicht losgelöst von der Ebene der Hardware.

In 1978, Texas Instruments had just released a sound-effects chip that could be programmed to generate a variety of sound effects. Of course, the programming was done primarily by wiring capacitors and resistors of certain values onto certain pins.¹⁷

Mit der Verbreitung von Mikroprozessoren wurde die Steuerung der Chips schließlich auch durch Software möglich. Auf direktestem Weg erfolgte die Programmierung in der Maschinensprache Assembler, die den Entwicklern direkten Zugriff auf die Register des jeweiligen Chips gewährt und Befehle für die Manipulation ihrer Inhalte bietet. Ein Zugeständnis an die Entwickler bestand darin, dass sie nicht mit binären Codes hantieren mussten, sondern durch Befehle in Form von Buchstabenkombinationen und hexadezimalen Zahlen mit leichter lesbaren Zeichen operierten. Die frühen Heimcomputer verfügten zudem über die Programmiersprache BASIC, die allen Anwendern die Programmierung der Maschinen ermöglichen sollte.

Der Siegeszug mikroprozessorbasierter Geräte für den Privatanwender begann im Jahr 1977 mit der Markteinführung der ersten Heimcomputer wie dem *Commodore PET* und dem *Apple II*. Steve Wozniak, der Konstrukteur des *Apple II*, war zuvor an der Entwicklung der Pong-Variante *Breakout* beteiligt, einem der letzten Videospiele, das komplett in Hardware realisiert war. Mit seiner neuen Maschine rekonstruierte er das mühsam erarbeitete Hardware-Design innerhalb von kürzester Zeit.

I started typing away in Basic and, within half an hour, I not only had my Pong game working, but I had done about 50 or so variations of colors and speeds and sizes and where the score was and all that stuff. I had changed so

¹⁶Dittbrenner listet den den Chip in seiner Arbeit über Soundchips als frühesten Vertreter der Programmierbaren Sound-Generatoren, abgekürzt PSG. (Dittbrenner, 2007, 9).

¹⁷(Crawford, 2003, Kapitel 13)

many things around and put in little features that would just take forever to do in hardware. [...] So I said to him [Steve Jobs], „Now that games are software, it's going to be a whole different world for games.“¹⁸

Mit den Heimcomputern der 1970er Jahre gab es zum ersten Mal programmierbare Maschinen, die außerhalb finanzkräftiger Institutionen wie Universitäten und Forschungseinrichtungen zugänglich waren. Ebenso wie Steve Russel gemeinsam mit Alan Kotok, Peter Samson und Dan Edwards 1962 am MIT mit ihrem Programm *Space War!* die Möglichkeiten eines in Echtzeit gesteuerten Raumschiff-Shoot-Em-Ups auslotete, erlebte Steve Wozniak knapp 15 Jahre später erneut die Kraft der Symbolik eines Software-Designs. Die Programmierung des großen Hits der vorausgegangenen Generation wurde zur Fingerübung.

Als Software kehrten Computerspiele technisch gesehen zurück zu ihren experimentellen Ursprüngen in Laboratorien und Universitäten. Die Befreiung des Entwicklungsprozesses von der materiellen Ebene der Hardware hob die Spieleproduktion auf eine neue Abstraktionsstufe. Anstatt die entworfenen Schaltungen mit dem LötKolben umsetzen zu müssen, standen nun weitaus effektivere Mittel zur Verfügung, um Spielaspekte innerhalb kürzester Zeit zu manipulieren.

Ebenso wie Steve Wozniak die Farben, Geschwindigkeiten und Positionen grafisch dargestellter Objekte variieren konnte, ermöglichte die Kombination aus Mikroprozessoren und Soundchips die schnelle Programmierung akustischer Elemente.

Soundchips förderten zum Einen eine gewisse Konsistenz der in ihrem Rahmen erstellten Artefakte, die den markanten Sound dieser speziell für die Videospieleindustrie konstruierten Synthesizer miteinander teilten. Zum Anderen ermöglichte ihre Programmierung die Realisierung komplexerer Klangstrukturen.

¹⁸Steve Wozniak, zitiert nach (Livingston, 2007, 50)

Sound in Software - Combat Pong

Was durch die Soundchips möglich wurde, soll im Folgenden am Beispiel des TIA-Chips, der von Joe Decuir und Jay Miner speziell für den Atari 2600 entwickelten Darstellungshardware¹⁹, demonstriert werden.

STIMME 1			
<i>Parameter</i>	<i>Register</i>	<i>Bits</i>	<i>Werte</i>
Volume	AUDV0	4 Bit	16
Frequency	AUDF0	5 Bit	32
Control	AUDC0	4 Bit	16
STIMME 2			
<i>Parameter</i>	<i>Register</i>	<i>Bits</i>	<i>Werte</i>
Volume	AUDV1	4 Bit	16
Frequency	AUDF1	5 Bit	32
Control	AUDC1	4 Bit	16

Tabelle 2.1: Sound-Register des TIA Chips

Er bot zwei identisch aufgebaute Stimmen, die mit jeweils drei Registern (AUDV0, AUDF0 und AUDC0 sowie AUDV1, AUDF1 und AUDC1) angesprochen werden konnten. Der vierte Buchstabe jeden Registers steht dabei als Kürzel für die Parameter Lautstärke (Volume, V), Frequenz (Frequency, F) und Kontrolle über die Wellenform (Control, C). Während die Kontroll- und Lautstärke-Register jeweils 4 Bit umfassen und damit 16 verschiedene Werte annehmen können, erlaubt das Frequenz-Register mit einer Größe von 5 Bit die Ansteuerung 32 verschiedener Tonhöhen.²⁰ Die Werte der Register entsprachen nicht den Maßeinheiten der von ihnen bezeichneten Parameter, sondern bildeten losgelöste Codes. Hinzu kam, dass z. B. beim Frequenzregister ein höherer Wert einen tieferen Ton ergab als ein niedriger Wert.

Folgendes Listing zeigt ein Beispiel für die Programmierung dieser Register. Das Code-Fragment stammt aus einer kommentierten Disassemblierung des Spiels Combat²¹, das von der Einführung der Konsole 1977 bis ins Jahr 1982 dem Atari 2600 beilag. Es kommt in dem Spielmodus

¹⁹vgl. (Montfort u. Bogost, 2009, 13)

²⁰vgl. (Slocum, 2003)

²¹(Dodgson u. a., Z. 1017-1023)

Combat-Pong zur Ausführung, in dem die Projektile zweier sich bekriegender Panzer von den Wänden abprallen und dabei jenen im Listing programmierten Pong-Sound erzeugen.

```
; Pong sound .
LDA #$04
STA AUDC0,X
LDA #$07
STA AUDV0,X
LDA BounceCount ,X
STA AUDF0,X RTS
```

Mit Hilfe des Akkumulatorregisters A erhält das Kontrollregister den fixen Wert 4 und das Lautstärkeregister den Wert 7. Die Bestimmung des Wertes für das Frequenzregister erfolgt durch die Variable *BounceCount*. An anderen Stellen innerhalb des Programms wird ihr der hexadezimale Startwert 1F zugewiesen und der Variableninhalt bei jeder Kollision um eins dekrementiert. Je häufiger ein Projektil abprallt, desto höher steigt der bei einer Kollision erzeugte Ton. Das Beispiel zeigt damit die einfache Verknüpfungsmöglichkeit von akustischen Parametern und anderen Aspekten des Spielgeschehens. Die doppelte Nutzung der Variable *BounceCount* – als Frequenzwerte der Kollisionstöne und Kollisionsanzahl – spiegelt unmittelbar hörbar das Zählen der erkannten Kollisionen wider. Da die meisten Spiele des Atari 2600 von einer einzigen Person entwickelt wurden, konnte die akustische Gestaltung prinzipiell ohne Absprache mit beliebigen Bereichen des Programms verschaltet werden. Im Gegensatz zum früheren Hardware-Design ist der Aufwand solch einer Verschaltung deutlich geringer.

Als frühester Vertreter seiner Art bot der Soundchip des Atari 2600 eine sehr übersichtliche Anzahl an Features. Die innerhalb des folgenden Jahrzehnts entwickelten Modelle enthielten deutlich umfangreichere Möglichkeiten der synthetischen Klangerzeugung.²²

²²Eine empfehlenswerte Übersicht der technischen Spezifikationen einer Reihe von einflussreichen Soundchips bietet die Arbeit von Nils Dittbrenner (Dittbrenner, 2007).

Integration von Musik

Die auffälligste Veränderung der akustischen Architektur im Zuge der Einführung programmierbarer Soundgeneratoren und Mikroprozessoren war die Integration von Musik. Aufgrund des hohen Aufwands und der zusätzlichen Kosten wurde sie in früheren Automaten nicht realisiert. Es war schlichtweg zu kompliziert, auf Basis des Hardware-Designs Musik zu integrieren. *Space Invaders* markiert mit seiner chromatisch absteigenden Basslinie den Wendepunkt.

Außerhalb des Videospiele-Kontextes wurde die Möglichkeit computergenerierter Musik weitaus früher in Betracht gezogen. Die früheste bekannte Aufnahme entstand 1951 auf einem Ferranti Mark I der Universität Manchester. Wenige Wochen zuvor soll bereits Australiens erster Computer CSIRAC Musik wiedergegeben haben.²³ Bevor Ivan Sutherland mit Sketchpad 1963 den Prototypen von Grafikbearbeitungsprogrammen entwickelte, schuf Peter Samson in der Wiege der Hackerkultur am MIT 1960 das Kompositionstool Harmony Compiler. Erst Ende der 1970er Jahre trat Musik auch in Computerspielen auf.

Wie die Programmierung von Musik funktionierte, soll im Folgenden am Beispiel des Sound Interface Device Chip (SID) demonstriert werden. Der 1981 unter der Kennung MOS 6581 konzipierte Soundchip des Commodore 64 gehört zu den prominentesten Soundchips der PSG-Ära. Anteil daran hatte nicht nur die weite Verbreitung des C 64, dem bis heute der Status als weltweit meistverkaufter Heimcomputer zukommt, sondern vor allem die klanglichen Möglichkeiten des Chips selbst. Obwohl auch er auf drei unabhängige Stimmen beschränkt war, ließen sich auf Basis von vier verschiedenen Wellenformen (Dreieck, Sägezahn, Rechteck und Weißes Rauschen), zuschaltbaren Filtern, digitalen Hüllkurven und weiteren Modulationsfeatures eine breite Palette von Klängen erzeugen.

Die große Popularität und der charakteristische Klang des Soundchips trugen dazu bei, dass er auch außerhalb des C 64 zum Einsatz kam. Ab 1990 produzierte und vertrieb die schwedische Firma Elektron mit der SIDStation einen Synthesizer, dessen Herzstück der MOS 6581 bildet. Populäre Anwender des Instruments sind u. a. Machinae Supremacy, Madonna und Trent Reznor. Der verantwortliche Ingenieur Bob Yan-

²³(Fildes, 2008)

nes, der später die auf Synthesizer spezialisierte Firma Ensoniq gründete, hatte von Beginn an die Konstruktion eines musikauglichen Chips im Sinn.

I thought the sound chips on the market, including those in the Atari computers, were primitive and obviously had been designed by people who knew nothing about music.²⁴

Die Schwerpunktsetzung auf Musik schlägt sich auch in der Bedienungsanleitung des C64²⁵ nieder, die Benutzer in die Programmierung der Maschine mit BASIC einweiht. Aus heutiger Perspektive mag es befremdlich erscheinen, in der Anleitung des Computers detaillierte Informationen über dessen Programmierbarkeit vorzufinden. Anfang der 1980er Jahre steckte die Software-Industrie jedoch erst in ihren Kinderschuhen, und jeder Benutzer eines Computers war ein potentieller Programmierer. Neben diesem Beispiel enthält das Kapitel der C64-Bedienungsanleitung zahlreiche Tips für die Ansteuerung des Chips, von der WellenformEinstellung über die Hüllkurve bis hin zum schließenden Abschnitt "Wir programmieren ein Lied auf dem Commodore 64".

Betrachtet man den in Abbildung 2.2 dargestellten Code, fällt auf, dass analog zu dem Code-Beispiel aus *Combat Pong* ein Großteil des Programms aus dem Schreiben von Werten in die Register des Soundchips besteht. In BASIC erledigt diese Aufgabe das Kommando *Poke*. Nachdem ausgehend von der Basisadresse des SID-Chips alle für das Beispiel benötigten Registeradressen definiert werden, beginnt die Zuordnung von Werten in diese Speichersegmente. Ebenso wie beim TIA-Chip besteht die digitale Repräsentation der erstellten Klänge aus den Wertemengen, die den Registern der Soundchips mit ihren jeweiligen klangerzeugenden Funktionen zugeordnet werden. Sie entspricht insofern nicht den Klängen selber, sondern bezeichnet die Parameter für den Soundchip als eine Art klangerzeugendes Instrument. Trotz der Exaktheit und Vollständigkeit dieser Maschinenpartitur, stimmte der hörbare Output auf unterschiedlichen Rechnern nicht unbedingt miteinander überein.²⁶

²⁴(Bagnall, 2011, 372)

²⁵(Commodore, 84-85)

²⁶Ein Grund dafür war die unterschiedliche Zeilenwiederholffrequenz der Normen NTSC und PAL. Siehe (Dittbrenner, 2007, 84)

```

10 REM TONLEITER //Programmname
20 SI=54272:FL=SI:FH=SI+1: //Definition der
   W=SI+4:A=SI+5:H=SI+6: //Registeradressen
   L=SI+24
30 POKE L,15 //Volle Lautstärke
40 POKE A,9 //Anschlag
50 READ X:READ Y //Hi-Byte und Lo-Byte
 //der Frequenz aus den
 //Data-Zeilen 130,140 lesen
60 IF Y=-1 THEN //Wenn das Programm die
   POKE W,0:END // -1 am Ende findet,
 //soll es abschalten
70 POKE FH,X:POKE FL,Y //Hi-Byte und Lo-Byte in
 //Frequenzregister POKEn
80 Poke W,17 //Wellenform, Generator
 //einschalten
90 FORT=1TO100:NEXT //Tondauer
100 POKE W,0 //Generator ausschalten
110 FORT=1TO50:NEXT //Kurze Pause zum Ausklingen
120 GOTO40 //Nächster Ton
130 DATA17, 103, 19, 137, //Diese Zahlenpaare stellen die
   21, 237, 23, 59, 26, //Töne der C-Dur Tonleiter dar,
   20, 29, 69, 32, 219 //immer abwechselnd ein Hi-Byte,
   34, 207 //ein Lo-Byte
140 DATA -1,-1 //Diese (als Frequenz sinnlosen)
 //Daten signalisieren dem Programm
 //in Zeile 60, dass die Tonleiter
 //zu Ende ist.

```

Abbildung 2.2: SID-Programmierbeispiel (Commodore, 85)

Man beachte insbesondere die Repräsentation der C-Dur Tonleiter in der Data-Zeile 130 als Aggregation von Daten.²⁷ Mit der Beauftragung einzelner Personen für die Erstellung solcher Datensätze für die Darstellungselemente beginnt in der Spieleentwicklung die Formierung von Teams. Die Audio-Verantwortlichen zeichneten sich in der Regel nicht nur gleichzeitig für Geräuschkulisse und Komposition, sondern auch für die Programmierung rund um den akustischen Part verantwortlich. Nicht selten schrieben sie ihre eigenen Tools, um Aspekte der Komposition und die Integration der erstellten Datensätze zu vereinfachen.

Zu den Soundeffekten, die massiv auf musikalische Kodierungen zurückgriffen, gesellten sich nun Melodien. Während der Soundchip des Atari 2600 aufgrund seines geringen Tonumfangs nur bedingt zur Programmierung musikalischer Strukturen brauchbar war, lösten die PSGs der 1980er Jahre hingegen einen regelrechten Boom der Computerspielmusik aus. Waren es zunächst nur kurze musikalische Fragmente, die besondere Momente oder Zeitfenster markierten, nutzten bereits ab Anfang der 1980er Jahre viele Spiele eine kontinuierliche Musikuntermalung. Eine Vorreiterrolle übernimmt in dieser Hinsicht der von Shigeru Miyamoto entwickelte und 1981 erschienene Arcade-Automat *Donkey Kong*. Die lückenlose musikalische Untermalung ist hier eng mit zentralen Spielzuständen und -ereignissen verzahnt. Eines dieser Ereignisse ist die Kollision der später als Mario getauften Figur Jumpman mit Fässern aus den Händen des Kidnapper-Gorillas. *Donkey Kong* etabliert damit Musik als Teil der Kollisionsantwort jenes verhängnisvollen Zusammenpralls, der den Tod der Spielfigur nach sich zieht. Noch heute ist dieses dramaturgische Element des Scheiterns stark in Gebrauch. Jedesmal, wenn in *Civilization 5* eine Einheit des Spielers das Zeitliche segnet, erklingt eine solche Todesmelodie. Auch in Rollenspielen wie *Fallout 3* ist die musikalische Kennzeichnung des Ablebens der Spielfigur sehr gebräuchlich.

²⁷ Auch wenn die technischen Bedingungen für die Integration von Musik damit gegeben sind, stehen die Gründe für die Integration zur Disposition. Musik ist kein traditioneller Bestandteil von Spieldarstellungen, sondern trat erst durch den Computer mit ihnen in Verbindung. Eine Erklärung wäre die Verbindung zum Film, der mit den gleichen Ausgabegeräten operiert und in dem die musikalische Untermalung eine Konvention darstellt. Weitere Gründe sind in der Verbindung von Musik und Spektakel zu suchen. So stehen die frühen musikalischen Gestaltungen der in Spielhallen aufgestellten Automaten sicher auch in der Tradition von Jahrmak- und Zirkusmusik.

Ein Grund für die Durchsetzung von Musik im Computerspiel der 1980er Jahre liegt sicher darin, dass sich die Konstrukteure der Soundchips wie Bob Yannes zunehmend an den Entwicklungen der Synthesizerindustrie orientierten. Die Soundchips wurden komplexer und enthielten später auch neuere Formen der Klangerzeugung wie die *FM-Synthese*. Ein weiterer Grund war die Arbeitsteilung innerhalb der Spieleentwicklung, die sich im Laufe der 1980er Jahre etablierte. Es gab nun Personen, die ausschließlich für die Erstellung des gesamten Audio-Spektrums verantwortlich waren. Komponisten wie Chris Hülsbeck, Martin Galway und Rob Hubbard loteten die Grenzen der Soundchips aus, erweiterten durch programmier- und kompositionstechnische Feinheiten ihre Ausdrucksmöglichkeiten und etablierten damit eine eigene Musikästhetik, die heute in der Chiptune-Szene fortlebt.

In der heutigen Wahrnehmung werden die Chips unmittelbar mit Musik assoziiert. Dass man mit ihnen die gesamte Klanglandschaft der Spiele erzeugte, bleibt weitgehend unbeachtet. Nach wie vor blieben hier jedoch die in den vorherigen Abschnitten beschriebenen Kompensationsformen synthetischen Sounddesigns intakt. Die Schrittgeräusche in Donkey Kong waren ein rhythmisches Muster zweier Töne, das beim Betreten der Leitern an Geschwindigkeit einbüßte. Herabfallende Gegenstände wurden durch einen fallenden Frequenzgang sonifiziert.

2.3 Reale Klänge in virtuellen Welten

Mit dem Aufkommen von 16-Bit Plattformen Ende der 1980er Jahre hielt gleichzeitig auch eine neue Technik der Klangerzeugung Einzug in Computer- und Konsolensysteme. Das als *Sampling* bezeichnete Verfahren ermöglichte die Integration zuvor digitalisierter Tonaufnahmen. Aufgezeichneter Schall wurde ein neuer Baustein des Computerspiels. Prinzipiell unterstützten auch schon frühere Soundchips wie der SID das Sampling, aufgrund beschränkter Rechen- und Speicherkapazität war es jedoch nicht praktikabel anwendbar. Während einer Übergangsphase zu Beginn der 1990er Jahre kamen zunächst nur kurze Aufnahmen für Soundeffekte in niedriger Auflösung zum Einsatz. Die musikalische Gestaltung basierte weiterhin auf Verfahren der Klangersynthese wie der Frequenzmodulation. In dieser Zwischenphase, in der Computerspiele auf das neu eingeführte MIDI-Format zurückgriffen,

resultierte die Musikwiedergabe auf unterschiedlichen Soundkarten in stark abweichenden Ergebnissen. Bis Mitte der 1990er Jahre ersetzte das Sampling die Klangsynthese in allen Bereichen nahezu vollständig. Dedizierte Soundchips kamen während der 1990er Jahre nur noch in portablen Spielkonsolen wie dem *Game Boy* zur Anwendung, während die Klangerzeugung stationärer Computer und Spielsysteme einen fundamentalen Wandel durchlief. Die Verbreitung der CD-ROM ermöglichte erstmals die Nutzung von Audiotracks und knüpfte eine Bande zwischen Musik- und Videospieldindustrie. Kompressionsverfahren wie MP3 begünstigten die Integration größerer Datenmengen, bevor mit der Einführung der DVD auch die Vertonung von Dialogen zur Norm wurde.

Mit der Verbreitung des Samplings setzte sich eine neue Repräsentationsform für akustische Inhalte durch, die sowohl aus technischer als auch ästhetischer Perspektive weitreichende Änderungen mit sich brachte.

Sampling

Bei der Überführung einer Aufnahme in eine digitale Repräsentation werden die analogen Signale in regelmäßigen Intervallen abgetastet und als diskrete Werte, den sogenannten *Samples*, gespeichert. Zwei Faktoren beeinflussen dementsprechend die Güte der digitalen Repräsentation. Die *Samplingrate* bezeichnet die Anzahl der Abtastungen pro Sekunde. Als zweiter Parameter bestimmt die *Bittiefe*, wie viel Speicherplatz für die einzelnen Samples aufgewandt wird. Übliche Werte der Bittiefe reichen von 4 bis 24 Bits, während die verwendeten Samplingraten zwischen 4.000 bis 96.000 Samples pro Sekunde pendeln. Es wird angenommen, dass ab 16 Bit pro Sample und 44.100 Samples pro Sekunde das menschliche Ohr den Unterschied zwischen der analogen und digitalen Repräsentation einer Aufnahme nicht mehr wahrnehmen kann.²⁸

Die technischen Anforderungen des Samplings unterscheiden sich drastisch von denjenigen der Klangsynthese. Als zentrale Komponente der

²⁸(Boer, 2002, 17). Für eine konzise technische Darstellung sei an dieser Stelle auf das Kapitel „Theory and Principles of Digital Sound and Acoustics“ in (Boer, 2002, 13-23) verwiesen.

Soundhardware fungiert nun ein Digital-Analog-Wandler, der die binären Zahlenketten der digitalen Repräsentation in analoge Signale überführt. Sämtliche klassischen Synthesizerkomponenten der PSGs sind für diese Form der Klangerzeugung nicht mehr erforderlich. Eine zentrale Hardware-Abhängigkeit besteht darin, dass die Verwendung von Samples unmittelbar an den verfügbaren Speicherplatz gebunden ist. Ohne jegliche Komprimierungsverfahren benötigt eine gesampelte Minute mit einer Bittiefe von 16 Bit und einer Samplingrate von 44.100 Samples bereits mehr als 5 MB. Tabelle 2.2 vermittelt am Beispiel der Ego-Shooter von idSoftware, wie drastisch der Speicherbedarf akustischer Daten seit Anfang der 1990er Jahre zugenommen hat. *Doom 1* und *2* enthielten neben den angezeigten Audio-Dateien alle Soundeffekte zusätzlich in einer synthetischen Variante für den PC-Speaker. Soundkarten waren zur Erscheinungszeit kein integraler Bestandteil des PCs. In Ergänzung zu den Soundeffekten nutzten beide *Doom-Titel* Midi-Dateien für die Musikwiedergabe, während die Musik in *Quake 1* und *2* über Audio-CDs eingespeist wurde. Beachtet man zusätzlich, dass idSoftware seit *Doom 3* auch Kompressionsformate wie Ogg Vorbis verwendet, erhält man einen ersten Eindruck von dem gewaltigen Anstieg des integrierten Audio-Materials. Der Übergang von synthetisierten Geräuschen zu hochauflösenden Samples geschah nicht über Nacht, sondern war ein Prozess, der sich über Jahre erstreckte. Die Schemenhaftigkeit der synthetischen Geräusche machte zunächst einer neuen Unschärfe Platz, die durch niedrige Samplingraten und geringe Bittiefen bedingt wurde. Mit steigenden Speicherkapazitäten gewannen die verwendeten Sounds zunehmend an Schärfe und Klarheit.

Eine Konsequenz dieser veränderten Anforderungen ist der Abstieg der Soundkarte als eigenständige Hardware-Komponente. Nahezu jedes Mainboard enthält heute grundlegende Soundfunktionalitäten, die von den meisten Anwendern offensichtlich als ausreichend erachtet werden. Gerade die Unterschiede zwischen heutiger Grafik- und Soundhardware verdeutlichen den Kern des Paradigmenwechsels im Zuge der Sampling-Durchsetzung. Computergrafik basiert weiterhin auf Berechnung, während in der Domäne des Game Sounds die Wiedergabe als wichtigste Funktion fungiert. Tatsächlich ist die Rechenpower der Grafikkarte mittlerweile so gewaltig, dass damit auch außerhalb des Spielkontexts rechenintensive Aufgaben wie numerische Simulationen²⁹ bearbeitet werden.

²⁹(www.scai.fraunhofer.de, 2011)

Titel	Jahr	Samplerate	Bitrate	#Samples	Größe (MB)
Doom	1993	11.025 Hz	8 Bit	67	0,673
Doom 2	1994	11.025 Hz	8 Bit	107	1,19
Quake	1996	11.025 Hz	16 Bit	325	3,59
Quake 2	1997	22.050 Hz	16 Bit	553	37,9
Doom 3	2004	22.050 Hz	24 Bit	4.103	342

Tabelle 2.2: Zunahme gesampelter Audiodaten am Beispiel idSoftware

Die technischen Voraussetzungen des Samplings umfassen neben dem D/A-Wandler den gesamten Aufnahmeapparat, angefangen bei Mikrofonen über Vorverstärker bis hin zum Analog-Digital-Wandler. Die spezielle Expertise der synthetischen Klangerzeugung für das Erschaffen von Sound durch programmierte Anweisungen ist nicht mehr erforderlich. Stattdessen tritt das praktische Aufnahmewissen des Filmsounddesigners in den Vordergrund, demzufolge man zerberstende Knochen z. B. hervorragend durch das Zerbrechen trockener Spaghetti simulieren kann. Da Computerspiele nun auf das gleiche Klangmaterial zurückgriffen wie traditionelle vertonte Medien, entstanden Allianzen zu anderen Zweigen der Unterhaltungsindustrie, die ebenfalls mit gesampeltem Sound operieren. Sowohl die Feld- oder Studioaufnahme, als auch die Nachbearbeitung der Aufnahmen gehörte nun zum Vertonungsprozess.

Sämtliche Sounds der PSG-Ära wurden berechnet, ihr Ursprung lag in den Wertzuweisungen an die Register der Soundchips. Samples liegen hingegen als externe akustische Elemente vor. Sie sind statische Datenmengen, die an den entsprechenden Stellen im Spiel abgerufen werden. An die Stelle der Erzeugung während der Laufzeit trat die Wiedergabe. Konnten vorher einzelne Parameter der synthetisierten Sounds mit beliebigen anderen Zuständen des Spiels verschaltet werden (siehe 2.2 auf Seite 69), bildeten nun die Samples die kleinsten Einheiten eines weitaus größeren Rasters. Wurden zuvor alle Belange der Klangerzeugung im Code abgewickelt, traten nun primär zeitkritische Lade- und Verwaltungsoperationen großer Datenmengen in den Vordergrund.

Um den softwaretechnischen Herausforderungen des Samplings zu begegnen, entstanden Audio-APIs³⁰ wie die 1995 eingeführte Bibliothek *DirectSound* und das fünf Jahre später veröffentlichte *OpenAL*. Sie stel-

³⁰Die Abkürzung API steht für „Application Programming Interface“. Dabei handelt es sich um Programmcode für wiederkehrende Aufgaben. Anstatt z. B. in jedem

len Funktionen für das Audio-Rendering und der einfachen Integration gesampelter Daten bereit. Auf einer höheren Abstraktionsstufe als diese APIs steht speziell für die Vertonung des Computerspiels entwickelte Middleware.

Middleware is software that connects game developers with the hardware (Xbox 360, PS3, PC, etc.) they use in development. Just as Pro Tools lets you generate sound from a computer, middleware lets users link sounds to game objects, such as animations (firing a gun, running), scripted events (a column falling across a road, a ship going to lightspeed) or areas (inside a church, at the foot of a cliff). Where a programmer was once required for all this integration, this is no longer the case.³¹

Die Middleware richtet sich nicht nur an Programmierer, sondern auch an Sounddesigner. Üblicherweise enthält sie eine Reihe von Werkzeugen für beide Rollen – ein Verwaltungs- und Integrationstool für die Sounddesigner und eine API für die Programmierer. Zu den frühesten weit etablierten Middleware-Paketen gehört das *Miles Sound System*, dessen erste Version im Jahr 1991 erschien.³² In den vergangenen Jahren fanden die Systeme *FMOD*³³ von Fireflight Technologies und *Wwise*³⁴ von Audiokinetic Inc. weite Verbreitung.

Integration von Sprache

Mit der Durchsetzung des Samplings geht eine gewisse Gleichberechtigung akustischer Elemente einher. Während auf Basis der Klangsynthese einige Wellenformen je nach Beschaffenheit des Synthesizers einfacher realisierbar sind als andere, ist die Technologie des Samplings seinen Inhalten gegenüber neutral. Alles, was Schall emittiert und aufgenommen werden kann, lässt sich gleichermaßen in eine samplebasierte Repräsentation überführen. Als Folge davon finden Geräusche und Klänge Eingang in das Computerspiel, die vorher nur sehr schwer

Programm mehrere Code-Blöcke für die Initialisierung der Soundkarte schreiben zu müssen, stellen Audio-APIs dafür vorgefertigte Befehle bereit.

³¹(Brandon, 2007)

³²<http://www.radgametools.com/miles.htm>

³³<http://www.fmod.org/>

³⁴<http://www.audiokinetic.com/en/products/208-wwise>

nachgebildet werden konnten. Eine Kategorie davon ist gesprochene Sprache.

Bereits während der Soundchip-Ära gab es einige Versuche, Sprache in Spiele zu integrieren, die sich jedoch aufgrund der bescheidenen Klangqualität und stark eingeschränkten Möglichkeiten nicht durchsetzten. Ein erwähnenswerter Arcade-Titel, der sich Methoden der Sprachsynthese bediente, war das 1980 erschienene Shoot-Em-Up *Berzerk*. Ausgehend von einem Vokabular von 30 Worten neckte eine roboterhafte Stimme die Spieler während des Attract-Modes mit den Worten „Coins detected in pocket“ und vermerkte „The humanoid must not escape“, wenn die Spielfigur einen Raum verließ, in dem alle Roboter beseitigt waren. Ein weiteres Beispiel ist das Sprachsynthese-Modul *Intellivoice*, das 1982 für die Konsole *Intellivision* von Matell auf den Markt kam und bereits ein Jahr später, mit lediglich fünf Spielen, die den Adapter unterstützten, wieder eingestellt wurde.

Neben der Sprachsynthese gab es weitere Bemühungen, um Sprache in Spiele zu integrieren. Dabei handelte es sich um zusätzliches Audio-Material, das nicht direkt in die Software integriert war, sondern auf zusätzlichen Datenträgern beilag und gleichzeitig mit dem Spiel abgepielt werden sollte.

The technical limitations of the early home PC systems such as the ZX Spectrum and the Commodore 64, and early consoles such as the NES, resulted in some creative ways to employ voice in order to both market and bring a cinematic edge to games. Usually these attempts used the b-side of the tape, or a supplemental tape that contained a music or a theme song, or even full blown, slickly-produced „interactive“ content.³⁵

Audio Director Rob Bridgett nennt als Beispiel für diese zusätzlichen Audio-Inhalte Automata UKs Spiel *Deus Ex Machina* aus dem Jahr 1984, dem neben dem eigentlichen Spiel eine parallel abzuspielende Kassette mit Musik, Soundeffekten und gesprochener Erzählung beilag. Einige Computerspiele dieser Zeit enthielten an bestimmten Punkten des Spielgeschehens Anweisungen für die Spieler, die beiliegende Kassette zu starten, zu stoppen oder an eine angegebene Position zu spulen. Die Spieler übernahmen damit genau jene Steuerungsfunktio-

³⁵(Bridgett, 2010, 39)

nen der Soundwiedergabe, die bei digitalisiertem Sound dem Programm zufallen.

Durch das Sampling wurde es in den 1990er Jahren schließlich möglich, echte Sprecher auch für Computerspiele ins Studio zu bitten und ihre Aufnahmen direkt in die Software zu integrieren. Die Sprachsynthese verschwand nahezu vollständig aus dem Computerspiel, da sie bis heute nicht an die Ausdrucksfähigkeit professioneller Sprecher heranreicht.³⁶ Bis zur Vertonung ganzer Dialoge sollten jedoch einige Jahre ins Land gehen, da der benötigte Speicherplatz zunächst nicht zur Verfügung stand. Zunächst kamen kurze Sprachsamples zum Einsatz, die den Spielern Rückmeldungen über Spielzustände gaben. In *Wolfenstein 3D* hörten die Spieler kurze Schreie und das Jaulen von Hunden als Kollisionsantworten. Mitte der 1990er Jahre folgten Sprüche der Spielfiguren, von denen vor allem Strategiespiele ausgiebig Gebrauch machen. Die Truppen in *Warcraft Orcs and Humans* quittierten die Befehle der Spieler mit sprachlichen Rückmeldungen, die ihre Ergebnisse bekunden. In *Jagged Alliance* (1994) spiegeln die Kommentare der Söldner ihre innere Disposition und verweisen damit auf Spielzustände und Ereignisse. Dies geschieht nicht nur über das Gesagte, sondern auch die Artikulation des Gesagten. Auch Spieler ohne Russisch Kenntnisse verstehen die spielrelevanten Bedeutungen der Äußerungen der Kultfigur Ivan Dolvich, der im ersten Teil ausschließlich Rückmeldungen in seiner Heimatsprache gibt.

Mittlerweile sprechen nicht nur Spielfiguren, sondern auch die Spieler selber. Neben der Kommunikation mittels „Voice over IP“ dient die Sprache der Spieler damit auch als Steuerelement. Vereinzelte Titel bieten eine Spracherkennung an, die gesprochene Befehle in Spielhandlungen übersetzt. Dabei kommen sowohl Wörter als auch Geräusche zum Einsatz. Auf dem *Nintendo DS*, das auf der Frontseite über ein eingebautes Mikrofon verfügt, gehört auch das Pusten in einigen Spielpassagen zu den Eingabemöglichkeiten. Es wird im Spiel eingesetzt, um Gegenstände wie Ballons zu bewegen. Wenn die Spieler in das Mikrofon pusten, registriert die Software das eingehende Rauschen und mißt dessen Intensität. Die unterschiedlich hohen Werte werden anschließend in Spielhandlungen übersetzt. Im Gegensatz zur Sprachsteuerung, die das Gesagte auf bestimmte Schlüsselwörter durchmustert, ist beim Pusten die gemessene Lautstärke entscheidend. Es ist durchaus denkbar, dass

³⁶ Abgesehen davon haben sich die Stimmen bekannter Sprecher – seien es Schauspieler oder Fußball-Kommentatoren – als werbewirksames Mittel erwiesen.

in Zukunft beide Dimensionen der akustischen Eingangssignale auch gleichzeitig analysiert werden und damit eine Vielzahl weiterer Steuerungsmöglichkeiten eröffnen. In *Second Life* gibt es bereits seit einigen Jahren ein Mapping von sprachlicher Artikulation auf die Gestik des Avatars.³⁷ Je lauter und schneller der Benutzer in das Mikrofon spricht, desto auffälliger gestikuliert sein Avatar und ergänzt das Gesagte durch generische nonverbale Kommunikationsweisen.

Der Fall Doom

Die Frage nach dem Wandel der Darstellungsqualität des Computerspiels ist nicht nur in Hinblick auf dessen Ästhetik interessant. Sie gewinnt dadurch an Relevanz, dass die qualitative Dimension des Dargestellten in kulturpolitischen Entscheidungen wie der Indizierung eines Computerspiels als Beurteilungskriterium für dessen Wirkung herangezogen wird. In Deutschland erlangte diesbezüglich vor allem der First Person Shooter *Doom* eine zweifelhafte Prominenz. Es dauerte nicht lange, bis die Bundesprüfstelle für jugendgefährdende Schriften (BPjS, heute BPjM³⁸) auf das 1993 von dem noch jungen Entwicklungsstudio idSoftware erstellten Spiels aufmerksam wurde. Einige Monate nach dem Erscheinen des Spiels unterzog es die Bundesprüfstelle einer Untersuchung, in deren Folge *Doom* in Deutschland indiziert wurde. Im Begründungsschreiben spielte die Darstellungsqualität neben einer Reihe von anderen Kriterien eine wichtige Rolle. So wurde nicht zuletzt die glaubwürdige Inszenierung der sinnlich wahrnehmbaren Form als ein Grund für dessen jugendgefährdendes Potential aufgelistet.

In der Indizierungsentscheidung Nr. 4637 (V) vom 25.5.1994 führte das Gremium der Bundesprüfstelle aus, wesentlicher Inhalt des Spiels sei die bedenkenlose, realistisch inszenierte Tötung der Gegner, darunter u.a. menschlich gestaltete Gegner. Ein erfolgreiches Durchspielen werde einzig durch

³⁷<http://community.secondlife.com/t5/English-Knowledge-Base/Using-gestures-and-animations/ta-p/700069> (17.10.2012)

Das Mapping in *Second Life* berücksichtigt zur Auswahl der Gesten nur die Eingangslautstärken des Gesagten. Komplexere Synchronisationen der Gestik und Mimik von Avataren mit den sprachlichen Eingaben ihrer Benutzer sind ein aktuelles Forschungsfeld der Mensch-Computer-Interaktion.

³⁸Aus den „Schriften“ wurden nach der Verabschiedung eines neuen Jugendschutzgesetzes im Jahr 2003 „Medien“.

die Liquidation zahlloser Gegner gewährleistet; die Tötungshandlungen würden mit blutig zerfetzten gegnerischen Körpern aufwändig dargestellt und akustisch untermalt (Geräusch der Waffe, Todesschreie).³⁹

Im Zuge eines Revisionsverfahrens wurde die Indizierung im Jahr 2011 aufgehoben und das Spiel von der Liste jugendgefährdender Medien gestrichen. Im Anschluss daran erteilte die USK eine Freigabe für Jugendliche ab 16 Jahren. In der Mitteilung über die Revision schrieb die Bundesprüfstelle:

Auf Kinder und Jugendliche übten die grafischen Darstellungen und Audioinhalte des Spiels kaum Anziehungskraft aus, da sie überwiegend technisch ausgefeilte, realitätsnahe und qualitativ hochwertige Inhalte gewöhnt seien. Sowohl die Rahmenhandlung wie auch die Darstellungsform wiesen keinen Realitätsbezug auf.⁴⁰

[..] es handle sich bei den Darstellungen in „Doom“ aus heutiger Sicht eher um symbolhafte als um realistische Darstellungen.⁴¹

Die Ausführungen in der BPjM-Entscheidung legen ein faszinierendes Zeugnis ab von den rasanten Entwicklungen der qualitativen Dimension des Computerspiels und den damit verbundenen Wahrnehmungsverschiebungen. Innerhalb von 17 Jahren wird aus einer realistisch empfundenen Inszenierung eine symbolhafte Darstellung. Unsere Einschätzung älterer Spiele ist heute eine vollkommen andere, als sie es zur Erscheinungszeit der jeweiligen Titel gewesen ist. Je länger das Erscheinungsdatum eines Computerspiels zurückliegt, desto schwerer erscheint es uns, die damaligen Reaktionen darauf nachvollziehen zu können. *Doom* ist dabei kein Einzelfall, sondern steht in der Tradition skandalträchtiger Spiele, die ihren großen Bekanntheitsgrad nicht zuletzt den in den öffentlichen Diskurs getragenen Kontroversen über Gewalt im Computerspiel verdanken. Als Prototyp dieser vieldebattierten Spiele gilt das 1976 erschienene *Death Race*. Es sorgte für Schlagzeilen, da das Ziel des während der Entwicklung unter dem Namen „Pedestrian“ gehandelten Spiels in dem Überfahren von „Gremlins“ liegt, die viele Zeitgenossen

³⁹(BPjM, 2011, 12)

⁴⁰(BPjM, 2011, 13)

⁴¹(BPjM, 2011, 15)

für Fußgänger hielten. Wer das Arcade-Spiel heute zu sehen und hören bekommt, wird kaum glauben, dass es in den USA einen Sturm der Entrüstung auslöste und nach anhaltenden Protesten vom Markt genommen werden mußte.⁴² Für unsere heutigen Begriffe sind die dargestellten Objekte – Strichfiguren und Kreuze – abstrakte Abbildungen, die jede Menge Fantasie erfordern, um gewalttätige Handlungen aus ihnen zu schließen.

Wie lässt es sich erklären, dass Darstellungen, die vor einigen Jahren als qualitativ hochwertig und überzeugend galten, heute als abstrakt und symbolisch eingestuft werden? Ist alleine die Gewöhnung an neue Standards für die Wirkung der sinnlich wahrnehmbaren Form ausschlaggebend? Einen Startpunkt zur Beantwortung dieser Frage bildet eine genauere Differenzierung der im Computerspiel dargestellten Entitäten. Sie lassen sich auf einem Spektrum verorten, dessen entgegengesetzte Pole die Abstraktion und die möglichst „realitätsgetreue“ Abbildung bilden.

„To *abstract* something is to simplify it, reducing it to a few essentials and basic forms instead of trying to reproduce it. *Representation*, which seeks to create resemblances and reproduce something, is the polar opposite of abstraction (and is sometimes conflated with *realism*).“⁴³

Retrospektiv lässt sich die Geschichte der sinnlich wahrnehmbaren Form des Computerspiels als Verschiebung auf einer Skala von Abstraktionsgraden begreifen. Zusätzlich zu der prinzipiellen Möglichkeit, Spielgegenstände in abstrakter Form darzustellen, besitzt das Videospiele im Laufe seiner Geschichte zu jeder Zeit eine technisch vorgegebene Grenze, die den höchstmöglichen Repräsentationsgrad markiert. Je weiter sich diese Grenze des technisch Machbaren durch neue Hard- und Softwareentwicklungen in Richtung höherer Auflösungen und Detailgrade verschob, desto mehr Optionen gab es hinsichtlich der Gestaltung.

Mit dem Übergang zum Sampling ging ein größerer Sprung auf dieser Abstraktionsskala einher. Die Begrenzung der Soundchips bzgl. der Geräuscherzeugung aus den jeweils verfügbaren klangmodellierenden Komponenten wurde aufgehoben. Gab es bei der Klangsynthese noch Geräusche, die durch die Architektur der Soundchips leichter als andere

⁴²(Montfort u. Bogost, 2009, 125)

⁴³(Wolf, 2003, 48)

konstruiert werden konnten, herrscht nun eine gewisse Gleichberechtigung akustischer Elemente. Unabhängig von der Komplexität seiner Wellenform kann prinzipiell jeder Sound nach der Überführung in eine digitale Repräsentation in das Spiel integriert werden. Für viele Spielereignisse führte die Einführung des Samplings deshalb zu Veränderungen ihrer akustischen Antworten, die vorher aufgrund von technischen Limitierungen nicht realisierbar waren. Bestand der Sound der Aktion „Springen“ während der PSG-Ära noch aus einer Art Frequenz-Glissando, sind es nun Atemgeräusche, die in Spielen wie *Quake* den Sprung begleiten.⁴⁴

Diese Erschließung neuer Repräsentationsweisen wird ein Grund dafür gewesen sein, warum *Doom* 1994 auf die Bundesprüfstelle so „realistisch“ wirkte. Die Integration aufgenommener Geräusche trug dazu ebenso seinen Teil bei wie die grafische Umsetzung dreidimensionaler Räume und die First-Person-Perspektive. *Doom* sorgte aufgrund seiner Popularität dafür, dass der First Person Shooter zu einer der dominanten Formen des Computerspiels aufstieg. Von der grafischen und akustischen Gestaltung bis hin zur Spielphysik und den Bewegungen des Avatars verbesserten die nachfolgenden Titel des Genres nahezu alle Aspekte des Spielgerüsts. Das Streben nach fotorealistischen Darstellungen und hochauflösendem Sound durchbrach in Begleitung von stets leistungsfähigeren Technologien immer wieder die Grenzen des bis dato Repräsentierbaren.

Das Phänomen der Wahrnehmungsverschiebung, in dessen Zuge wir Darstellungen zu anderen Zeiten in Hinblick auf ihre Qualität und ihren suggerierten Realismus unterschiedlich bewerten, betrifft nicht nur das Computerspiel. Auch bei Gemälden, Filmen und Tonaufnahmen treten in Abhängigkeit der jeweils aktuellen Repräsentationstechniken diesbezüglich andere Einschätzungen auf. Ein Beispiel dafür ist der High Fidelity Standard. Ursprünglich in den 1960er Jahren als DIN Norm 45500 festgehalten, wurde die Definition des Qualitätsstandards 1996 durch die Norm DIN EN 61305 ersetzt, die „[...] im Hi-Fi-Markt keine qualitätsbestimmende Rolle mehr ein[nimmt], da die meisten Aufzeichnungsgeräte, Tonträger und Wiedergabegeräte diese Anforderungen bei weitem übertreffen.“⁴⁵. Die Besonderheit dieses Phänomens im Kontext des Computerspiels liegt jedoch in der rapiden Beschleunigung verän-

⁴⁴Die zusätzliche Vertonung des Aufpralls auf dem Boden legt zudem eine feinere Segmentierung der Ereignisse nahe.

⁴⁵http://de.wikipedia.org/wiki/High_Fidelity (17.10.2012)

derter Wahrnehmungen des Dargestellten. An Stelle von Jahrtausenden oder Jahrzehnten ändern sich in unserer hochtechnisierten Informationsgesellschaft die Repräsentationsweisen so schnell, dass die ständige Anpassung von Qualitätsstandards für uns geradezu selbstverständlich ist. Vor allem innerhalb der ersten 30 Jahre des Computerspiels waren die Spieler an Veränderungen des Dargestellten gewöhnt. Bereits nach einem Zeitraum von ein bis zwei Jahren wurde ein Titel, dessen Darstellungsqualität sie zum Erscheinungsdatum bestaunten, wegen „schlechter Grafik“ diskreditiert. Seit der Jahrtausendwende scheinen sich die Zyklen dieser Obsoleszenz qualitativer Standards in langsameren Schritten zu vollziehen. Einerseits sind die Repräsentationsmöglichkeiten so weit gediehen, dass sie an filmische Standards nahezu heranreichen. Andererseits verfolgen die wenigsten Computerspiele einen Darstellungsstil, der realistisch wirken soll.

Nichtsdestotrotz bleibt auch für die Zukunft die kontinuierliche Erweiterung des Repräsentierbaren im Rahmen des Computerspiels zu erwarten. Auch wenn die Qualität des Films erreicht wird, stehen die Matrix und das Holodeck als weitere Zielpunkte bereit.

Trotz der Indizierung in Deutschland hatte *Doom* großen Einfluss auf die Entwicklung der Spielindustrie. Die Unterstützung von Mods und die spätere Veröffentlichung des Quellcodes führten zu einer stärkeren Einbindung der Spieler-Community und benutzergenerierter Inhalte.⁴⁶ Auch was die Entwicklung und Vermarktung von Computerspielen angeht, beschritt idSoftware neue Wege. Die Lizenzierung der *Doom*-Engine und der in den Folgejahren weiterentwickelten Engines ermöglichte es den Lizenznehmern, Spiele auf Basis eines bereits bestehenden technologischen Fundaments aufzubauen. Die Auseinandersetzung mit den Low-Level-Spezifika einer hochkomplexen Engine-Architektur entfiel, was die Entwicklungszeit enorm beschleunigte. Die Verbreitung als Shareware, welche die kostenlose Herausgabe eines Teils des Spiels vorsah, während die meisten Inhalte käuflich erworben werden mussten, etablierte ein bis heute intaktes Verkaufsmodell. An Stelle von Levels sind es heute lediglich virtuelle Schwerter, Rüstungen oder sonstige Güter, die kostenpflichtig sind. Im Grunde genommen lässt sich das heutige „Free to play“-Modell als konsequente Erweiterung des *Doom*-Verkaufsmodells verstehen. Anstatt die freie Version auf die ersten Level

⁴⁶(Dyer-Witthoford u. de Peuter, 2009, 24f)

zu beschränken, sind nun in viel feineren Segmentierungen Inhalte mit weiteren Kosten versehen.

Grenzen des Sampling

Der Paradigmenwechsel von der Synthese zum Sampling wird in der Riege der Sounddesigner nicht ausschließlich als Fortschritt bewertet. Obwohl sich die Mehrheit mit diesem Umstand offenbar sehr gut arrangiert, da die Produktionsbedingungen in weiten Zügen denjenigen aus anderen Kontexten wie der Audio-Produktion im Film entspricht, melden sich kritische Stimmen zu Wort. Der Sounddesigner Andy Farnell kritisiert das entstandene Daten-Modell als Affirmation eines rückständigen wirtschaftlichen Denkens, das aus der Musik- und Filmindustrie auf die Computerspielbranche übertragen wurde:

The data model sits nicely with the old school philosophy from the film and music business which is principally concerned with assets and ownership. The methodology is simple, collect as much data as possible in finished form and assemble it like a jigsaw puzzle, using a hammer where necessary, into a product. At each stage a particular recording can be ascribed ownership by a game object or event.⁴⁷

Er fordert statt dessen die Weiterentwicklung eines prozeduralen Modells, das analog zur Computergrafik auf der Berechnung basiert und den collagen-artigen Charakter gesampleten Sounds im Spiel hinter sich lässt:

As mentioned above, a recorded sound always plays precisely the same way, whereas procedural sound may be interactive with continuous real-time parameters being implied. [...] Synthesized crowds can burst into applause or shouting; complex weather-systems where the wind speed affects the sound of rainfall; rain that sounds different when falling on roofs or into water; realistic footsteps that automatically adapt to player speed, ground texture, and incline – the dynamic possibilities are practically endless.⁴⁸

⁴⁷(Farnell, 2007, 16)

⁴⁸(Farnell, 2010, 321f)

Farnells Vision sieht die enge Verknüpfung zwischen einer fortgeschrittenen synthetischen Klangerzeugung und den physikalischen Modellen der Spielwelt vor. Das auf diese Weise in Echtzeit berechnete akustische Material böte eine deutlich größere Vielfalt und einen höheren Detailgrad als statische Samples. In der Tat zielen viele der aktuell gebräuchlichen Praktiken zur Einbindung von Samples gerade darauf, ihre Fixiertheit zu kaschieren. So ist es für die Vertonungen vieler Ereignisse und Zustände üblich, an Stelle von einem Sample eine Gruppe von Samples einzurichten. Innerhalb der Gruppe legen Gewichtungen fest, mit welcher Wahrscheinlichkeit die enthaltenen Samples beim Eintreten der Aktion abgespielt werden. Durch Zufallswerte ermittelte Abweichungen der Lautstärken und Tonhöhen sorgen für weitere Variationen. Der folgende Skript-Abschnitt aus *Left 4 Dead* zeigt ein Beispiel solcher Klanggruppen.⁴⁹

```
// _____
// Hunter Zombie
// _____
// random vocalizations
"HunterZombie.Voice"
{
  "channel"      "CHAN_VOICE"
  "volume"      "1"
  "soundlevel"  "SNDLVL_85dB"
  "pitch"       "95, 105"

  "GameData"
  {
    "Priority"   "Interesting"
  }

  "rndwave"
  {
    "wave"      "player/hunter/voice/idle/Hunter_Stalk_01.wav"
    "wave"      "player/hunter/voice/idle/Hunter_Stalk_04.wav"
    "wave"      "player/hunter/voice/idle/Hunter_Stalk_05.wav"
  }
}
```

Der Eintrag „rndwave“ verweist auf drei Samples, von denen beim Eintreten des Ereignisses „HunterZombie.Voice“⁵⁰ jeweils eines ausgewählt wird. Die zwei Werte des Parameters „pitch“ bezeichnen eine Unter- und eine Obergrenze. Innerhalb dieses Intervalls wird ein Zufallswert ermittelt. Alle weiteren Geräusche, die Variationen aufweisen – von

⁴⁹left4dead/scripts/game_sounds_infected_special.txt:1248-1270

⁵⁰Der Hunter gehört in *Left 4 Dead* zu den „besonderen Infizierten“. Wenn er sich in der Nähe der Überlebenden befindet, gibt er in unregelmäßigen Abständen Laute von sich, welche sie vor ihm warnen. Die Zeitpunkte dieser unfreiwilligen Artikulationen werden vom Spielsystem bestimmt.

Schritt- und Schuss- bis hin zu Umgebungsgeräuschen – sind nach dem gleichen Prinzip vertont (siehe auch 4.2 auf Seite 160).

Im Gegensatz zu diesen Hilfsmitteln böten Formeln, die auf Basis physikalischer Parameter der Spielwelt Klänge akustische Ausgaben berechneten, eine elegantere Lösung. Mehrere Gründe sind denkbar, weshalb sich dieser Ansatz bislang nicht durchsetzte. Zunächst einmal ist er in erster Linie für Computerspiele prädestiniert, die überhaupt komplexe physikalische Simulationen wie Wettersysteme beinhalten. Dazu gehören allen voran Computerspiele, die dreidimensionale Umgebungen mit einem hohen Glaubhaftigkeitsanspruch entwerfen. In abstrakt gehaltenen Spielen, die nicht auf eine realistisch wirkende Darstellung abzielen, sind die Vorteile prozeduralen Sounds überschaubar. Geringerer Speicherplatzverbrauch und höhere CPU-Anforderungen halten sich die Waage. Möglicherweise sind Variationen der Soundeffekte nicht erwünscht.

Ein weiterer Grund für die Dominanz des Sample-Paradigmas sind die vorherrschenden Produktionsmethoden. Es gibt umfangreiche Sample-Bibliotheken, aber kaum adäquate Formel-Bibliotheken für vollständig synthetisch generierte Geräusche. Während Sounddesigner routiniert im Umgang mit Samples sind, ist die intime Kenntnis mathematisch-physikalischer Berechnungen zur synthetischen Klangerzeugung weniger weit verbreitet. Diese Art der Auseinandersetzung mit akustischen Elementen ist zwar in industriellen Anwendungsgebieten wie dem Sound Design der Automobilindustrie, sowie vereinzelt in Klanginstallationen vorzufinden, aber kaum im Rahmen der gängigen Medienproduktion.

Trotz der Vorbehalte sind erste Ansätze einer in Echtzeit berechneten akustischen Szenerie innerhalb der derzeitigen Produktionspraxis vorhanden. So ermöglichen z. B. die *Environmental Audio Extensions* (EAX) von Creative Labs die Simulation von umgebungsbedingtem Nachhall (siehe auch S. 155). Obwohl die Parameter des Halls nicht aus der jeweiligen Raumgeometrie ermittelt werden, sondern als zuschaltbare, durch numerische Parameter anpassbare Effekte bereit stehen⁵¹, leisten sie immerhin eine Transformation der Samples in Echtzeit. Einen weiteren Ansatz bildet die *Sound Seed* Technologie der Audio-Middleware *Wwise*. *Sound Seed* bezeichnet eine im Aufbau befindliche Bibliothek von Soundgeneratoren. Bislang enthält sie die Module *Sound*

⁵¹(Boer, 2002, 507)

Seed Air und *Sound Seed Impact*, welche mit Wind- und Kollisionsgeräuschen wie dem Aufeinanderprallen von Schwertern zwei typische Klangkategorien des Computerspiels abdecken.⁵²

2.4 Vergangenes, Gebliebenes und Wiederkehrendes

Trotz des rasanten technischen Fortschritts und den damit einhergehenden Veränderungen von Produktion und Rezeption ist die Entwicklung der akustischen Gestaltung nicht nur durch Umbrüche gekennzeichnet. Veränderungen sind ebenso zu beobachten wie Kontinuitäten und das Wiederaufgreifen älterer Gestaltungsformen, die eine Zeit lang aus der Mode geraten waren. Synthetisches Klangmaterial und bewusst repetitiv gehaltene Vertonungen treten heute auf der Vertonungsebene ebenso in Erscheinung, wie auf Ebene des Game Designs phasenweise erloschene Genres wie das Grafikadventure, der Platformer und der Side Scrolling Shooter.

Vertonungen abstrakter Ereignisse und Zustände

Zu vielen Spielereignissen gibt es Vorbilder plausibler akustischer Antworten aus spielexternen Kontexten, seien es in der Realität hörbare Geräusche oder deren medial etablierte, hyperreale Pendanten. Wie in den vorhergehenden Abschnitten dargelegt wurde, resultierte die Einführung des Samplings in der Ersetzung der vorher synthetisch modellierten Sounds durch aufgenommene Sounds. Das Geräusch einer sich öffnenden Tür wurde dementsprechend nicht mehr mit Hilfe des Synthesizers konstruiert, sondern aufgenommen und als Sample integriert. Einige der typischerweise vertonten Ereignisse treten jedoch nur in der Domäne des Spiels auf. Sie besitzen keine akustischen Vorbilder aus anderen Kontexten, die überhaupt aufgenommen werden könnten. Dazu gehören in erster Linie die Vertonungen abstrakter, repetitiv auftretender Spieloperationen wie die Vergabe von Punkten oder das Einsammeln von Gegenständen. Auch wenn sich die zugrunde liegende Technik der Klangerzeugung geändert hat, blieb die Gestaltung dieser

⁵²<http://www.audiokinetic.com/en/products/210-soundseed> (18.10.2012)

speziellen Klangkategorie bis heute nahezu unverändert und reicht in ihren Grundzügen bis zu dem Sounddesign elektromechanischer Spielautomaten zurück. In der Regel handelt es sich um kurze synthetische Sounds mit mittigen und hohen Frequenzanteilen.

Besonders prägnant treten diese Relikte des synthetisierten Sounds in der Übergangsphase des Samplings Anfang der 1990er Jahre zu Tage. In *Wolfenstein 3D*, dem ersten Spiel von idSoftware, das Samples verwendete, ist der Kontrast deutlich zu hören. Die Geräusche von Türen, Schüssen, und Schreien sind niedrig aufgelöste Aufnahmen, die in dieser Form mit den damaligen Soundchips nicht ohne weiteres synthetisch modelliert werden konnten. Sammeln die Spieler jedoch Gegenstände ein, erklingen synthetische Soundeffekte, die genau so gut aus einem Titel der 1991 bis 1992 entwickelten *Commander Keen* Reihe hätten stammen können.

Selbst in Spielen, die großen Wert auf eine realistisch anmutende Darstellung legen, sind solche Klänge vermehrt zu hören, auch wenn sie dezenter ausfallen als ihre Ahnen aus der Frühzeit des Computerspiels. Jedes Mal, wenn die Spieler in *Dead Space* einen Gegenstand einsammeln – sei es Geld, Munition, oder ein Logeintrag – erklingt das gleiche synthetische Sample. In *Battlefield Bad Company 2* folgt jeder Punkteinblendung ein hochfrequentes Klicken. Diese dezenten Klänge übernehmen eine wichtige Rolle. Sie signalisieren den Spielern Erfolg, indem sie die Anhäufung guter Aktionen oder Besitztümer sonifizieren, die für die Bewältigung des Spiels zuträglich sind. Sie erfüllen wichtige Funktionen hinsichtlich der Spieler motivation.

Trotz des Paradigmenwechsels der Klangerzeugung enthalten Computerspiele für bestimmte abstrakte Ereignisse wie die Punktevergabe oder die Verwendung von Menüs also immer noch synthetische Geräusche, auch wenn sie technisch gesehen als Samples vorliegen.

„Hadouken“ - Kultsounds und -samples

Ein zentrales Charakteristikum der akustischen Gestaltung des Computerspiels ist das Hervortreten einiger ständig wiederholter Klänge. Aufgrund der repetitiven Struktur des Spiels ist es unvermeidbar, wenn bestimmte Ereignisse und Zustände häufig auftreten. Sofern ihre Vertonungen nicht aus mehreren Sample-Variationen bestehen, sind die

mit ihnen verknüpften akustischen Ausgaben entsprechend ebenso häufig zu hören. In der Frühzeit des Computerspiels war es aufgrund der beschränkten Ressourcen nicht ökonomisch, einem Ereignis mehr als einen Klang zuzuordnen. Später ging man dazu über, mehrere mögliche akustische Ausgaben für bestimmte Ereignisse zu definieren (siehe S. 88). Trotz dieser Techniken zur Schaffung akustischer Variationen sind repetitive Klänge heute nach wie vor weit verbreitet. Sie besitzen die Eigenschaft, dass sie sich in das Gedächtnis der Spielergemeinschaft einstanzen. Diese Einprägsamkeit machen sich Sounddesigner gezielt zu Nutze. Die Abwägung, ob ein Ereignis durch einen bestimmten Sound oder mehrere Varianten vertont wird, ist heute eine der typischen Herausforderungen der akustischen Gestaltung des Computerspiels. Bewußt repetitiv gelassene Sounds und Samples können als eine Art Markenzeichen gezielt dazu eingesetzt werden, Wiedererkennungswerte zu schaffen. Insbesondere bei Sequels stellen sie als bekannte Elemente von Beginn an eine hohe Vertrautheit mit dem Spiel her. Das Sounddesign von *Diablo 3* nutzt diese Strategie auf massive Weise, indem es viele Soundeffekte nahezu identisch von dem 12 Jahre älteren Vorgänger übernimmt. Ein weiteres Beispiel ist die Star Power und Star Man Music der *Super Mario* Spiele, die der Serie über zwei Jahrzehnte erhalten geblieben ist.

Innerhalb der Spielergemeinschaft genießen viele Sounds und Samples einen Kultstatus und dienen als Ausgangspunkte kultureller Produktion. Bands wie *Hadouken!* benennen sich nach ihnen⁵³, Fans spielen mit ihrer Metaphorik, und die Netzgemeinschaft dokumentiert die daraus emergierenden Formen in Portalen wie *knowyourmemes.com*. Auch wenn die Computerspielounds der 1970er und 1980er Jahre nicht weniger repetitiv und einprägsam waren als diejenigen der nachfolgenden Dekaden, ergaben sich erst mit dem Umschwung aufs Sampling die Voraussetzungen für deren einfache Weiterverarbeitung durch die User. Anders als der programmierte Sound kann das Sample dem Programm leicht entnommen und modifiziert werden. Vor allem Sprachsamples des Computerspiels sind mit Bedeutungen aufgeladen, die in spieleexternen Kontexten referenziert, weiterentwickelt oder gebrochen werden.

⁵³ *Hadouken!* ist eine britische New-Rave-Band. Das Sprachsample „Hadouken“ begleitet einen Special Move in der *Street Fighter* Reihe, bei dem die Spielfigur ihre Handballen aneinander presst und einen Energieball in Richtung des Gegners schleudert.

Durch Plattformen wie *youtube* treten Samples aus Computerspielen in eine Art kreative Arena der Netzgemeinschaft, in der es darum geht, vielfältige Verknüpfungen mit ihnen aufzubauen. Beobachten lässt sich der Formenreichtum solcher Weiterverarbeitungen am Beispiel des Samples „Pills here“, das als Internet-Phänomen bereits Eingang in die Datenbanken der Netzgemeinschaft fand:

In the video game *Left 4 Dead*, pain pills are one of the available HP boosting items. When a character sees them, he or she exclaims, “Pills here.”⁵⁴

Eins der bekanntesten *youtube*-Videos, das mit diesem Sample operiert, stellt Louis, einen der vier Überlebenden des Spiels, als pillenabhängig dar.⁵⁵ Etliche weitere Videos greifen diese Thematik spielerisch auf, wobei die Beiträge vom kurzen Zusammenschnitt einer Ibuprofen-Dose und dem Sample bis hin zu mehrminütigen Machinimas und Kollagen reichen.⁵⁶ Darunter befinden sich ein Dutzend Songs, die größtenteils auf Basis der Spielsamples erstellt wurden und mit Videos aus *Left 4 Dead* oder in *Garry's Mod* bearbeiteten Materialien daraus unterlegt sind.

Wenn Kultsounds auf diese Weise dem Computerspiel entnommen und transformiert werden, lösen sie sich potentiell von den Prozessen, mit denen sie im Spiel verbunden sind. In manchen Fällen bilden aber gerade die Bedeutungen und Assoziationen dieser Prozesse den Mittelpunkt der Weiterverarbeitung. So markiert das Sample „Wololo“⁵⁷ aus *Age of Empires* einen plötzlichen, von außen oktroyierten Sinneswandel. Im Spiel besitzen die Priester die Eigenschaft, gegnerische Einheiten zur eigenen Fraktion zu bekehren. Die betroffenen Einheiten signalisieren dann durch die Anpassung der Farbe ihre veränderte Zugehörigkeit. Einige Comicstrips und retuschierte Bilder der Netzgemeinschaft spielen mit dieser farblichen Oberflächentransformation, indem sie Personen in roter Kleidung nach dem Zauberwort „Wololo“ in blauer Kleidung zeigen. Im Sprachgebrauch des Gamers steht die Nachahmung des Sam-

⁵⁴<http://knowyourmeme.com/memes/grabbin-pills> (19.03.2012)

⁵⁵<http://www.youtube.com/watch?v=wOBnrAtIhRo> (19.10.2012)

⁵⁶siehe z. B. http://www.youtube.com/watch?v=llnmd5_sC64 (19.10.2012)

⁵⁷Siehe <http://knowyourmeme.com/memes/wololo> (19.03.2012): *Wololo* is the sound made by the Priest unit in the 1997 game *Age of Empires*. It is used as an interruptor, causing the color of something to change colors between red and blue.

ples allgemeiner als schalkhaftes Zeichen dafür, Jemanden oder Etwas unter seine Kontrolle bringen zu wollen.

Retro-Inszenierungen und Demakes

Innerhalb der vergangenen Jahre gewann das Darstellungsspektrum des Computerspiels deutlich an Vielfalt. Die Gestaltung richtet sich nicht nur nach den technisch aktuellen Verfahren, sondern leistet zunehmend Rückgriffe auf technisch obsolete Darstellungsweisen. Computerspiele mit „Klötzchengrafik“ und Chiptunes erschienen zur gleichen Zeit wie Computerspiele mit hochauflösenden HD-Grafiken und Klangkulissen in Dolby Surround. Längst überholte und noch vor einiger Zeit als qualitativ minderwertig bewertete Darstellungen wurden nun gezielt als ästhetische Mittel eingesetzt. Die Umdeutung früherer Darstellungsformen als gleichberechtigte Gestaltungsmittel, deren ursprünglich aus technischen Limitierungen entstandene Form nun keinen qualitativen Makel mehr bedeutete, sondern einen ästhetischen Stil, ist ein weiteres Indiz für die Maturität des Computerspiels. Es greift frühere technische Entwicklungen als Gestaltungsmöglichkeiten auf, die mit Hilfe aktueller Technologie simuliert werden.

Retro-Inszenierungen⁵⁸, die mit solchen Rückgriffen operieren, sind in erster Linie in Independent Games vorzufinden, die seit dem Aufkommen von Vertriebsplattformen wie *Steam* einem Massenpublikum zugänglich sind. Anstatt den opulenten Darstellungen größerer Computerspielproduktionen nachzueifern, besinnen sich die Entwickler von Independent Games auf reduziertere Darstellungsformate.

Eine spezielle Form der Rückbesinnung auf vergangene Plattformen stellen *Demakes* dar. Dabei handelt es sich um die Umsetzungen und Neuinterpretationen aktueller Computerspiele mit den Mitteln vergangener Epochen, die von Hobby-Programmierern, Fans, und Retro-

⁵⁸Auch wenn der Begriff „Retro“ (lat. zurück, rückwärts, nach hinten) aufgrund seiner derzeitigen inflationären Verwendung etwas sehr Aktuelles zu beschreiben scheint, verweist er auf ein sehr altes Phänomen. Rückgriffe auf zurückliegende Kulturprodukte sind keine Besonderheit unserer Zeit, sondern in historischen Zusammenhängen immer wieder zu beobachten. Bezeichnend für unsere Kultur ist jedoch die erhöhte Geschwindigkeit, in deren Fahrtwind das Heutige zum Gestrigen wird und das Gestrige in Rückbezügen wiederum das Heutige beeinflusst. Eine Annäherung an dieses Phänomen im Kontext aktueller Technologien leistet der Sammelband (Böhn, 2010).

Enthusiasten angefertigt werden.⁵⁹ In seiner avanciertesten Form simuliert das Demake nicht nur die charakteristischen Eigenschaften einer obsoleten Plattform, sondern ist auf dessen Basis programmiert und darauf ausführbar. Neben der Berücksichtigung klassischer Gameplay-Mechaniken zielt auch die Gestaltung von Bild und Ton auf eine möglichst originalgetreue Nachbildung der Konventionen eines je nach Projekt grob oder fein abgesteckten Zeitraums der Computerspielgeschichte. So entwickelte der Demake-Programmierer Eric Ruth z. B. eine Hommage an *Left 4 Dead*, die sich stilistisch an den Gepflogenheiten des NES im Jahr 1986 orientiert.⁶⁰ Die komplexe Level- und Gameplay-Architektur des First Person Shooters wird darin auf ein schlichtes Shoot-Em-Up in einer scrollbaren 2D-Umgebung reduziert, während Musik und Soundeffekte mit den Mitteln der Klangsynthese realisiert sind. Wie es zur damaligen Zeit üblich war, verzichtet das Demake auf eine Ambient-Kulisse und exponiert stattdessen kontinuierlich fortlaufende Hintergrundmusik als dominante Konstituente des akustischen Gefüges. Das Beispiel zeigt, dass sich Demakes nicht nur mit den technischen Limitierungen, sondern auch mit den künstlerischen Entscheidungen und Konventionen einer Plattform auseinandersetzen.

Textuelle Relikte

Obwohl das Computerspiel heute über alle Voraussetzungen verfügt, um gesprochene Sprache zu verarbeiten oder Spielereignisse und -zustände mit Samples zu vertonen, blieb geschriebener Text eine wichtige Darstellungskomponente. In der Tat enthalten Computerspiele mitunter derart hohe textuelle Anteile, dass grafische und akustische Ausgaben hinter ihnen zurücktreten. Anders als in der Entwicklung des Films, in der beim Übergang von der Stummfilm- in die Tonfilmära textuelle Beschreibungen bis auf wenige Ausnahmen⁶¹ in der Tonspur aufgingen, trat in der Entwicklung des Computerspiels auf vielen Ebenen Text hinzu oder blieb bestehen. Während sich in Arcade-Games und frühen Konsolen die Ausgabe von Text auf dem Bildschirm äußerst mühsam gestaltete, traten auf Universitäts- und Heimcomputern mit Textad-

⁵⁹ vgl. (Görig, 2009) und „Bootleg Demakes Competition“: <http://forums.tigsources.com/index.php?board=22.0> (10.04.2012)

⁶⁰ <http://ericruthgames.com/> (10.04.2012)

⁶¹ z. B. in Prologen und Epilogen, bei eingeblendeten Zeit- und Ortsangaben oder zur Darstellung schreibender und lesender Charaktere

ventures und *Multi User Dungeons* (MUDs) ab Ende der 1970er Jahre zwei textbasierte Genres hinzu. Der Einzug von multimedialen Darstellungsweisen in diese Genres legte den Grundstein für Grafikadventures und *Massively Multiplayer Online Games* (MMOGs).

In der Übergangsphase entfielen textuelle Beschreibungen der Spielwelt. Grafische und akustische Darstellungen traten an ihre Stelle. Mit der steigenden Verfügbarkeit von Speicherplatz begann die Sprachvertonung der Spielfiguren. Trotz dieses Umbruchs verschwand der Text selbst in Dialogen nicht vollständig. Häufig werden vertonte Dialoge gleichzeitig auch textuell dargestellt.⁶² In Computerspielen, die mit einer First Person Perspektive operieren, bleiben die Vertonungen dieser Antworten gewöhnlich aus. Die Spieler klicken auf geschriebenen Text und die NPCs antworten mit gesprochener Sprache. In den meisten Rollenspielen steht den Spielern eine unbeschränkte Zeit für das Antworten zur Verfügung. Um die Folgen ihrer Entscheidungen abzuwägen, analysieren sie die Antwortmöglichkeiten sehr genau. Die textuelle Darstellung des zuvor Gesagten unterstützt sie dabei. Weiterhin enthalten Rollenspiele Questlogs, die den Spielern in textueller Form Gedächtnisstützen über geführte Gespräche und erteilte Aufträge bereit stellen. Generell entfaltet Text eine geringere Verbindlichkeit als vertonte Sprache, der die Spieler ihre Aufmerksamkeit widmen müssen. Dargestellte Texte lassen sich darüber hinaus schneller durchmustern als gesprochene Sprache. Sie können überflogen werden und führen schneller zu den gewünschten Informationen, als wenn der gesamte Text vorgelesen wird.

In MMOGs stellt der Text-Chat trotz der integrierten VOIP-Lösungen immer noch die vorherrschende Kommunikationsform dar. Ein entscheidender Grund dafür ist die Neutralität geschriebenen Textes, der die wahre Identität seines Verfassers nicht preis gibt. Gerade dann, wenn die Stimmen der Spieler von den erwarteten Stimmen ihrer Avatare abweichen, wahrt der Text die Fiktion. Software-Lösungen, die Stimmen in Echtzeit transformieren, um sie den Avataren anzupassen, befinden sich zwar in Entwicklung, konnten sich bislang jedoch nicht durchsetzen.

⁶²Viele unterschiedliche Strategien sind hier in Gebrauch. Einige Computerspiele zeigen die kompletten Antworten. Andere Spiele wie *Mass Effect* stellen textuell lediglich stichwortartige Tendenzen der Antworten. Erst nach der Auswahl einer Antwort bekommen die Spieler zu hören, welche konkrete Antwort mit der Auswahl verbunden ist.



Abbildung 2.3: Texteinblendungen in *Diablo 3*

Neben der Aufzeichnung und Repräsentation von Gesagtem kommt Text auch dann als beschreibendes Mittel zur Anwendung, wenn die grafische und akustische Darstellung ihre Grenzen erreicht. In Computerspielen, welche die Spielwelt aus einer isometrischen Perspektive zeigen, sind fallen gelassene Gegenstände z. B. schwer erkennbar und werden deshalb mit einblendbaren Textbeschreibungen versehen. Abbildung 2.3 zeigt ein Beispiel aus *Diablo 3*. Alleine durch ihre grafische Darstellung wären die nach dem Kampf übrig gebliebenen Gegenstände nicht identifizierbar. Die Texteinblendungen erleichtern die Entscheidung, welche Gegenstände die Spieler einsammeln und welche sie liegen lassen. Farbliche Kodierungen markieren dabei unterschiedliche Klassen von Gegenständen. So erhalten Items, die magische Eigenschaften aufweisen, beispielsweise eine blaue Schrift.

Einen dritten Fall, in dem geschriebener Text unverzichtbar ist, stellt die Beschreibung von Spielaktionen dar. Ebenso wie die Regeln des Spiels im Programm implementiert sind, enthält die Darstellung die Spielanleitungen. Benutzbare Gegenstände und durchführbare Spielaktionen sind mit textuellen Beschreibungen versehen, die ihre Wirkungen und Effekte offenlegen. Manchmal werden diese Beschreibungen durch „Tooltips“ eingeblendet. Manchmal werden sie dauerhaft neben einem

Gegenstand oder dem Button einer Aktion angezeigt. Wie hoch der Anteil dieser Textpartien ausfällt, ist von dem jeweiligen Genre und der Komplexität des Spiels abhängig. Rollen- und Strategiespiele neigen zur einer höheren Textdichte als Action-Spiele.

Herzklopfen

Das Herz ist als Symbol des Lebens fest in der Darstellung des Computerspiels verankert. In Arcade-Spielen tauchte es als grafisches Element häufig auf, um die Anzahl der Leben des Spielers darzustellen, wobei ein Leben im Grunde einem Spielversuch entsprach. Die auf diese Weise eingeführte Beschränkung der Spielzeit war ein wichtiger Aspekt, um die Rentabilität der münzbetriebenen Automaten zu garantieren. Mit der Etablierung von Konsolen- und PC-Spielen für den Heimgebrauch setzte sich eine andere Konzeption des virtuellen Lebens durch. An die Stelle mehrerer Leben, von denen man welche verlieren, aber auch neue dazugewinnen konnte, trat nun ein einziges Leben gekoppelt mit einem Modell der Gesundheit. Die Gesundheit der Spielfigur wird dabei durch den Wert auf einer Vitalitätsskala simuliert. In der grafischen Darstellung machte das Herzsymbold im Lauf der Jahre zunehmend dem roten Kreuz als international etabliertem Schutz- und Kennzeichen medizinischer Versorgung Platz. Mit dem Verschwinden des grafischen Herzsymbols und dem Einzug der Vitalitätsskala tritt auf der akustischen Ebene das Geräusch des Herzklopfens als Warnhinweis hinzu.

Ein Computerspiel, das diesen Übergang deutlich demonstriert, ist das 1986 für die SNES-Konsole erschienene *Legend of Zelda*. Immer dann, wenn der Gesundheitswert unter eine gewisse Schwelle sinkt, gibt es eine akustische Rückmeldung über den bedrohlichen Zustand der Spielfigur. Die Lebensenergieleiste in *Legend of Zelda* ist aus mehreren Herzsymbolen zusammengesetzt. Erleidet die Spielfigur Schaden, verschwindet die Hälfte eines Herzens. Sobald die Leiste auf ein halbes Herz herabsinkt und die Spielfigur somit in akuter Gefahr vor dem Spieltod schwebt, erklingt in regelmäßigen Intervallen ein synthetisches Piepsen. Ohne das grafische Symbol wäre das Piepsen nur schwer als Herzklopfen erkennbar. Die mangelnde Ähnlichkeit des synthetischen Sounds zu bekannten Geräuschen aus anderen Kontexten eröffnet einen weit größeren Interpretationsspielraum. Das Bild beeinflusst die Auslegung des Tons. In heutigen Spielen, die das Herzklopfen als Warnung einblenden, ist eine

visuelle Stütze dieser Art nicht mehr nötig. Die Aufnahme ist direkt als Herzklopfen identifizierbar, weil sie sich stark an den akustischen Konventionen des Films orientiert.

3 Zeitliche Konstruktionsprinzipien

Das Zusammensetzen der akustischen Bausteine lässt sich als Bauvorgang bezeichnen. Immer, wenn das Computerspiel in Abhängigkeit der Eingaben seine Ausgaben konstruiert, folgt es den Vorgaben des in Ausführung befindlichen Bauplans. Dieses Prinzip betrifft nicht nur die akustische Dimension, sondern sämtliche Ausgaben des Computerspiels. Stärker als jedes andere Medium existiert es nur in seiner Prozesshaftigkeit. Selbst, wenn die Ausgaben statisch erscheinen und die Spieler für einen Moment auf Eingaben verzichten, durchläuft das Programm stoisch seine Anweisungen und Schleifen, um diese scheinbare Stasis beizubehalten. Sobald das Programm seine Ausführung stoppt, fällt alles Gebaute abrupt in sich zusammen.¹

Die zeitlichen Strukturen des Bauplans bestimmen maßgeblich die Form der akustischen Architektur des Spiels. Als prozedurales interaktives Medium erfordert das Computerspiel andere Einbindungsstrategien akustischer Elemente als traditionelle Medien. Das Arrangement von Soundeffekten, Musik und Sprache ist nicht fixiert wie die Tonspur des Films oder die Aufnahme eines Hörspiels. Es ist zeitlich variabel. Würde man den Ton von mehreren Spieldurchgängen mitschneiden, erhielte man immer wieder unterschiedliche Klangergebnisse. Während die Bausteine dieser Mitschnitte jeweils identisch sind, unterscheiden sich die Zeitpunkte und Kombinationen, in denen sie zu hören sind. Der Mix entsteht erst während der Laufzeit in Abhängigkeit der getroffenen Entscheidungen, Spieldurchgänge und Zufallsprozesse. Diese zeitliche Ungewissheit bedeutet sowohl für das Audio-Design eines Spiels, als auch für die kritische Auseinandersetzung damit eine große Herausforderung. Denn um die Ausdrucksfähigkeit akustischer Elemente in den variablen

¹Die einzige Ausnahme sind Spielstände oder Zustände in persistenten virtuellen Welten, die bestimmte Lücken des Bauplans mit Werten füllen und beim nächsten Aufbau des immer wieder Zerfallenden einen Teil rekonstruieren.

Strömen der Ausgabe in den Blick zu nehmen, reichen Mitschnitte alleine nicht aus. Stattdessen muss der Bauplan mit seinen Regeln für die Zusammensetzung akustischer Elemente in den Fokus rücken.

Das Ziel dieses Kapitels besteht darin, die wichtigsten Strategien für die zeitliche Einbindung akustischer Elemente zu identifizieren und ihre Funktionen im Rahmen der Interaktion zu erörtern. Ausgehend von der Betrachtung verschiedener Zeitstrukturen, die innerhalb des Computerspiels auftreten, sowie der Organisation von Bauplänen für akustische Ausgaben, werden Funktionen akustischer Arrangements in Hinblick auf ihre zeitliche Dimension untersucht. Ein abschließender Schwerpunkt liegt auf unterschiedlichen Musikkonzeptionen als Antwort auf die zeitliche Ungewißheit des Spiels.

3.1 Zeitstrukturen des Computerspiels

Akustische Elemente werden von uns in ihrer gesamten Länge erst in der Zeit wahrnehmbar. Zeit ist ein existenzieller Parameter jedes akustischen Gefüges und die akustische Architektur des Computerspiels bildet keine Ausnahme. Im Computerspiel ist das Gefüge der akustischen Ausgestaltung nicht autark, sondern befindet sich vom Takt der CPU bis zur zeitlichen Konzeption des Spiels in einem Geflecht weiterer Zeitstrukturen. Sowohl Hard- und Software, als auch das Spiel schaffen zeitliche Möglichkeitsbedingungen und Anforderungen der akustischen Gestaltung. Ein erster Schritt besteht deshalb darin, die wichtigsten Zeit-Formen dieses Geflechts in ihren auftretenden Varianten abzuschreiten. Im Anschluss daran folgt eine Beobachtung und Diskussion von ereignisbasierten Integrationsstrategien akustischer Bausteine. Sie loten insbesondere die Verzahnungsmöglichkeiten der Zeitstruktur des Spiels mit akustischen Strukturen aus.

Laufzeit und Echtzeit

Den zeitlichen Rahmen jedes Programms und damit auch jedes Computerspiels bildet die *Laufzeit*. Sie bezeichnet die Zeitspanne, innerhalb derer das Programm ausgeführt wird. Anwendungssoftware besitzt die

Eigenschaft, dass sie durch den Einsatz von Schleifen und Wartezeiten künstlich am Leben gehalten wird. Die Laufzeit des Programms ist hier nicht von einer Problemstellung abhängig, die möglichst schnell gelöst werden soll, sondern vom Wunsch der Benutzer, das Programm für eine gewisse Zeitspanne zu verwenden. Die Länge der Laufzeit ist ungewiß, da die Benutzer das Programm zu beliebigen Zeitpunkten starten und beenden. Wie auf Seite 31 am Beispiel der *Left 4 Dead* Entwicklung demonstriert wurde, könnte ein Programm schneller ablaufen, wenn die Benutzer durch Künstliche Intelligenzen ersetzt würden. Die zeitliche Gestaltung der Darstellung ist also speziell für den menschlichen Wahrnehmungsapparat konzipiert. Viele Computerspiele erinnern direkt nach dem Start an diesen Umstand, indem sie mit dem Hinweis „Press any key to continue“ eines der traditionsreichsten Eingeständnisse an die beschränkte Verarbeitungsgeschwindigkeit des Menschen im Kontext der Mensch-Computer-Interaktion aufgreifen. Seit den 1970er Jahren nutzen Programmierer den Textbaustein in Kombination mit einer Pausierung der maschinellen Verarbeitung. In Konsolenanwendungen raste der ausgegebene Text so schnell über den Bildschirm, dass ihn niemand mehr lesen konnte.² Durch das zwischenzeitliche Pausieren wurden die Ausgaben in mehrere bildschirmfüllende Segmente unterteilt, die auf Tastendruck nacheinander angesteuert werden konnten.

Innerhalb des übergeordneten zeitlichen Rahmens der Laufzeit bietet sich nun Platz für die Ausgestaltung anderer zeitlicher Konstrukte wie Musik, Sound und Animation, sowie deren jeweilige Synchronisation. Das Programm agiert gewissermaßen als Dirigent, der über Start, Verlauf und Ende aller Einsätze mit taktgenauer Präzision verfügen kann. An einigen zuvor festgelegten Punkten wartet es jedoch auf Eingaben, welche die akustische Komposition beeinflussen. Hier offenbart sich eine Tiefenkomplexität, die auf eine entscheidende Differenz zum Film verweist. Während die Wiedergabe eines Films aus einem einzigen, kontinuierlichen Abspielvorgang besteht, bringt die Ausführung eines Computerspiels diverse kleine, kaum voneinander unterscheidbare Abspielvorgänge mit sich. Während der Laufzeit greift das Programm auf

²Die rasante Geschwindigkeit, mit der Programme während der Laufzeit ihre Anweisungen durchlaufen, erzeugt auch Probleme für die Untersuchung von Code. Aufbauend auf Kittlers Diagnose, dass Menschen Code zwar lesen und schreiben, aufgrund der schichtweisen Struktur von Programmiersprachen aber nicht fassen können (Kittler, 1997, 148), erkennt Mackenzie in der Ausführungsgeschwindigkeit eine zeitliche Unzugänglichkeit von Code (Mackenzie, 2006, 24f).

Datensätze von Klängen und Bildern zu, die gemäß des ausgeführten Bauplans zu einem kohärenten Ganzen zusammengesetzt werden.

Die zusammenaddierte Laufzeit aller Spielsitzungen gilt als Maßstab für die *Spielzeit*. Sie mißt die gesamte Zeit, in der das Programm genutzt wurde und ist ebenfalls ungewiß. Es ist möglich, dass manche Spieler ein Computerspiel in 20 Stunden bewältigen, während andere dafür 40 Stunden benötigen. Durch das Spielen mehrerer Partien steigt die Gesamtstundenzahl, welche Spieler einem Titel widmen, mitunter auf mehrere hundert bis tausend Stunden. Darüber hinaus besitzen persistente Online-Universen wie *World of Warcraft* kein festgelegtes Spielende. Hier ist das Spiel erst dann vorbei, wenn die Server abgeschaltet werden. Auch, wenn die Spieler ausgeloggt sind, bleiben die serverseitig gespeicherten Spielzustände erhalten. In Singleplayer-Spielen ist der laufende Abspann ebenfalls kein sicheres Indiz für das Spielende. Titel wie *Diablo* errichten eine zyklische Zeitstruktur, indem sie das Spielende als neuen Anfangspunkt setzen und die Spieler zum erneuten Durchspielen mit härteren Gegnern und wertvolleren Gegenständen ermutigen.³

Eine weitere technische Zeitform, die im Zusammenhang mit Computerspielen auftritt, ist die *Echtzeit*. Sie bezeichnet im weitesten Sinne die Anforderung an ein Programm, innerhalb einer festgelegten Zeitspanne auf bestimmte Ereignisse reagieren zu können.

Crucially, hard and soft real-time systems are subject to a „real-time constraint.“ That is, they need to respond, in a forced duration, to actions predefined as events. The measure of real time, in computer systems, is their reaction to the live; it is their liveness – their quick acknowledgment of and response to our actions.⁴

Jede Anwendungssoftware, und damit auch jedes Computerspiel, besitzt Echtzeit-Anforderungen, die das Feedback der Eingaben betreffen. In der Domäne des Spiels nimmt der Echtzeit-Begriff neben dieser Anforderung des Antwortens innerhalb bestimmter Zeiträumen jedoch noch eine weitere Bedeutung an. Er bezeichnet Computerspiele, in de-

³In den vergangenen Jahren nutzten viele Spiele dafür Bezeichnungen wie „Story Plus“. Die erzählte Geschichte wird beim erneuten Durchspielen jedoch in der Regel nicht erweitert. Statt dessen steigt der Schwierigkeitsgrad an.

⁴(Chun, 2011, 68)

nen die simulierte Zeit kontinuierlich weiterläuft. Echtzeitbasierte Spiele bilden damit einen Gegenpol zu rundenbasierten Spielen. In einem Echtzeitstrategiespiel müssen die Spieler unter Zeitdruck genauestens abwägen, wann sie welche Aktion ausführen. Es gibt keine Pausen, in denen das Spiel still steht. Es verlangt den Spielern eine Kombination aus strategischem Geschick und Reaktionsfähigkeit ab. Das rundenbasierte Strategiespiel verschafft ihnen hingegen beliebig viel Zeit, um die Konstellationen des Spielfelds zu überblicken und ihre Züge mit Bedacht zu wählen.

In Echtzeitbasierten Spielen steht nicht nur das Programm in der Pflicht, innerhalb bestimmter zeitlicher Fristen auf Ereignisse reagieren zu müssen. Umgekehrt stellen die Spiele auch Anforderungen an die Spieler, indem sie in bestimmter Zeit Aktionen auszuführen haben. Spiele wie *Super Meat Boy* und *Street Fighter IV* loten die Grenzen dieser Reaktionsfähigkeit aus. Sie verlangen ein präzises Timing der Eingaben, das durch ständige Repetition trainiert werden muß. Die Spieler sind in diesem Fall sehr stark in die zeitlichen Vorgaben des Spiels eingespannt.

Zeitkritische und -entspannte Spielmomente

Für das Spiel als regelbasiertes System nimmt Zeit unterschiedliche Bedeutungen an: Je nach Regelsatz ist sie für den Spielablauf kritisch oder unerheblich. Abseits von Turnierregeln, die zeitliche Beschränkungen vorsehen, kann ein Schachzug beispielsweise beliebig viel Zeit in Anspruch nehmen. Eine brieflich geführte Partie zieht sich womöglich über Jahre hin. Für ein Fußballspiel sind hingegen 90 Minuten vorgesehen, auch wenn die Auslegung der durch Spielpausen und Nachspielzeiten beeinträchtigten Länge im Ermessen des Schiedsrichtergespanss liegt.

Zeitspannen und Zeitpunkte erhalten hier jeweils eine andere Gewichtung, eine andere Bedeutung im Kontext des Spiels. Der Zeitpunkt, zu dem ein Schachzug ausgeführt wird, ist für den Zug an sich nicht entscheidend. Fußball hingegen ist ein *zeitkritisches* Spiel. Sowohl der Zeitpunkt eines Spielereignisses, als auch die Dauer eines Spielzustands sind potentiell spielentscheidend. Der Pass zum Mitspieler muss genau so wie der Schuss aufs Tor im richtigen Moment erfolgen. Das *Timing*, die möglichst präzise abgestimmte zeitliche Ausführung von Spielaktionen, bestimmt zu einem bedeutenden Anteil den Wert und die Qualität

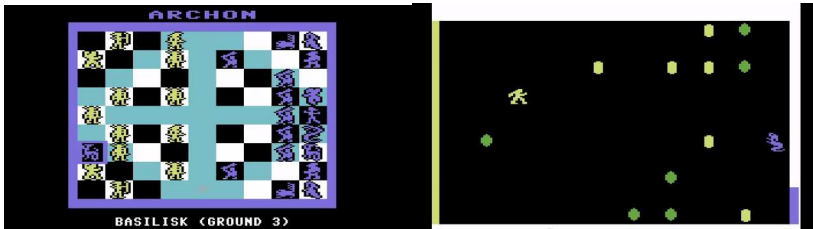


Abbildung 3.1: *Archon* (C64, 1984) - Planungsansicht und Echtzeit-Kampf

eines Spielzuges. Darüber hinaus ist der zeitliche Ablauf in den Regeln verankert. Wenn der Torwart einer Fußballmannschaft in einem Match zu lange mit dem Abstoß wartet, bekommt er eine Verwarnung wegen Zeitspiels.

Auch in Computerspielen lassen sich Spielformen unterscheiden, die zeitkritische Entscheidungen erfordern oder es nicht tun. Ein notwendiges, aber nicht hinreichendes Kriterium für die Implementierung zeitkritischer Spielsituationen ist ein Ablauf in Echtzeit. Darunter fallen z. B. sämtliche Sportsimulationen sowie nahezu alle Titel, die unter der Genrebezeichnung „Action“ subsumiert werden. In anderen Genres bleiben die Zeitpunkte, zu denen die Spieler Entscheidungen treffen, unbedeutend. Im Grafikadventure bewegt sich die Spielfigur vielmehr in Zeitschleifen. So lange ein Rätsel nicht gelöst wird, antworten NPCs mit identischen Sätzen, wiederholen sich Hintergrundanimationen und Musikstücke. Eine weitere bedeutende Klasse dieses Gegenpols bilden rundenbasierte Spiele, die eine in Züge gerasterte Zeiteinteilung vorgeben. Diese in der Tradition klassischer Brettspiele wie Schach stehende Zeitverwaltung tritt überwiegend in Strategiespielen wie der *Civilization*-Reihe auf.

Zeitenspannte Momente geben den Spielern die Gelegenheit, ihrer Fantasie freien Lauf zu lassen. Im Gegensatz zu einem Film, der kontinuierlich weiterläuft, können die Spieler stehen bleiben und in Ruhe die derzeitigen Konstellationen, Objekte und Räume betrachten. In dieser Hinsicht kombinieren zeitenspannte Spiele die Rezeptionsmöglichkeit des Buchs, im eigenen Tempo lesen und sich die beschriebenen Inhalte vorstellen zu können, mit der Direktheit repräsentativer Darstellungen, wie sie der Film bietet.

In nicht-digitalen Sport- und Gesellschaftsspielen ist die Vermengung unterschiedlicher Spielprinzipien und Zeitstrukturen sehr selten. Ausnahmen wie das von dem Aktionskünstler Iepe Rubingh ins Leben gerufene Schachboxen⁵ gelten eher als Kuriositäten. Im Computerspiel ist diese Hybridbildung hingegen sehr geläufig. Einige Mischformen wie das Echtzeitstrategiespiel sind etablierte Genres geworden. Bereits das 1984 für den C64 erschienene Spiel *Archon* kombiniert Elemente des Schachspiels mit denjenigen gängiger Action-Shooter. Betritt eine Spielfigur das Feld einer gegnerischen Einheit, erscheint ein neuer Bildausschnitt, in dem sie in einem Echtzeit-Kampf gegeneinander antreten (Abbildung 3.1). Die Spiele der *Total War* Reihe setzen dieses Prinzip fort, indem die Verwaltung des Herrschaftsbereichs in einem rundenbasierten Modus abläuft und die Schlachten in Echtzeit vonstatten gehen. Im Grunde genommen handelt es sich um zwei miteinander verknüpfte Spiele, die auch einzeln nutzbar sind. Der Multiplayer-Modus der *Total War* Reihe unterstützt nur die in Echtzeit ablaufenden Schlachten. Der rundenbasierte Kampagnenmodus ist hingegen auch ohne den Echtzeit-Part spielbar, indem die Ergebnisse der Schlachten simuliert werden. Auch dann, wenn sich nicht zwei unterschiedliche Zeitstrukturen ablösen, sind Kombinationen möglich. Biowares Rollenspiele *Mass Effect* und *Dragon Age* bieten während der in Echtzeit ablaufenden Kämpfe jederzeit die Möglichkeit, das Spiel zu pausieren, um den Spielfiguren neue Anweisungen zu geben. Die Geschicklichkeit und schnelle Reaktionsfähigkeit ist dadurch nicht so entscheidend wie die sorgfältige strategische Planung.

Nicht-zeitkritische Situationen können auch in Computerspielen auftreten oder bewußt erzeugt werden, die prinzipiell eine zeitkritische Struktur aufweisen. In Rollenspielen verlangen in der Regel nur die Kämpfe nach gut getimten Aktionen. Befinden sich die Spieler in Ruhezeiten, um Handel zu betreiben oder Gespräche zu führen, sind die Zeitpunkte ihrer Aktionen unerheblich. Selbst in Action-Spielen können Momente der Stasis erzeugt werden, indem z. B. der Avatar an einem Ort positioniert wird, der gegnerischen Figuren unzugänglich ist.

Die Unterscheidung zwischen einem zeitkritischen Rahmen und einer von zeitlichen Anforderungen losgelösten Spielkonzeption ist für die akustische Gestaltung von Computerspielen elementar. Aus den andersartigen Voraussetzungen folgen unterschiedliche funktionale An-

⁵http://site.wcbo.org/content/index_de.html (24.06.2013)

bindungsmöglichkeiten für akustische Elemente, die in 3.3 näher untersucht werden.

Trotz dieser Ungewißheiten über den zeitlichen Ablauf und die Länge des Spiels verlaufen mehrere Durchgänge des selben Spiels in der Regel ähnlich. So lässt sich der Spielablauf meistens in mehrere Phasen unterteilen. Im Schach unterscheidet man z. B. zwischen Eröffnung, Mittelspiel und Endspiel.⁶ Während die Eröffnung für versierte Spieler häufig reine Routine bleibt, treten im Mittelspiel entscheidende Situationen und Wendungen ein. Das Endspiel wiederum ist durch leichter vorhersehbare Konstellationen und erlernbare „finishing moves“ in der Regel weniger dramatisch als das Mittelspiel. Innerhalb von diesem übergreifenden Spannungsbogen, der das Spiel in mehrere Phasen aufteilt, können immer wieder spannungsgeladene Momente hervorstechen oder spannungsarme Momente das Spiel beruhigen. So treten in *Tetris* Kombinationen der Spielteile auf, die die Spieler problemlos verarbeiten können, während andere Konstellationen für eine Zuspitzung sorgen. Spiele besitzen Spannungsbögen, die aus den Regeln, Zielen und Zustandsübergängen des Spiels entstehen. Dieser aus dem Spiel entstehende Spannungsverlauf wird im Folgenden als *ludische Dramaturgie* bezeichnet, der Spannungsverlauf einer Erzählung hingegen als *narrative Dramaturgie*. Narrative Dramaturgie ist geplant. Ludische Dramaturgie ist zwar in den Regeln angelegt, wird aber stark von Spielvermögen und Zufall bestimmt.

Zeit der fiktiven Spielwelt

Neben Laufzeit, Echtzeit und Spielzeit enthalten Computerspiele Zeitvorstellungen, die sich aus der präsentierten fiktiven Welt ergeben. Dazu gehört sowohl das Zeitalter des Settings wie z. B. Mittelalter und Science Fiction, als auch die Zeit, die während des Spielablaufs im Rahmen der Fiktion vergeht. Letztere wurde in den Game Studies für verschiedenen Spieltypen bereits analysiert.⁷ Der häufigste Fall ist die Stauchung der fiktiven Zeit im Vergleich zur realen Zeit. Das simulierte Fußballspiel dauert 10 Minuten an Stelle von 90 Minuten, der Tag 40 Minuten an Stelle von 24 Stunden und für eine Runde vergeht ein Jahr.

⁶siehe z. B. (Pfleger, 2004)

⁷siehe (Wolf, 2002, 77-91) und (Juul, 2005, 141-156)

Die Zeit der fiktiven Spielwelt stellt einen wichtigen gestalterischen Einflußfaktor dar, an dem sich auch die akustische Ausgestaltung orientiert. Das zeitliche Setting entscheidet, ob die Geschichte im Mittelalter, in der Gegenwart, oder in der Zukunft spielt und legt damit eine Reihe von Konventionen für deren Darstellung nahe. Für das Zukunftsszenario ist z. B. ein elektronischer Soundtrack üblich, während mittelalterliche Fantasy-Szenarien auf konventionelle Instrumente und orchestrale Arrangements zurückgreifen. In feineren Abstufungen bestimmen die Tages-, Monats- und Jahreszyklen die jeweilige Ausgestaltung des Settings. Für die Nacht erwarten wir andere visuelle Gegebenheiten und andere Klangkulissen als für den Tag. In dem Rollenspiel *Risen*, das einen kontinuierlichen Wechsel zwischen Tag und Nacht simuliert, herrschen am selben Ort zu unterschiedlichen Tageszeiten dementsprechend ganz unterschiedliche Lichtverhältnisse. Am Tag bricht die virtuelle Sonne durch das Dickicht, während nachts nur vereinzelte Lichtquellen wie Fackeln die Szenerie beleuchten. In Ergänzung zu diesem Spiel aus Licht und Schatten sind in *Risen* verschiedene Tageszeiten mit unterschiedlichen Musikstücken verknüpft.

Vor allem Open-World-Spiele nutzen die Zeit der fiktiven Spielwelt als Auswahlkriterium der dargebotenen Inhalte. In *Grand Theft Auto IV* sind z. B. die Radiobeiträge nach Tageszeiten sortiert und werden innerhalb des Spiels entsprechend abgerufen.

Die Zeit der fiktiven Spielwelt ist die Domäne der narrativen Dramaturgie. Erzählerische Spannungsbögen entfalten sich darin als Abfolge von Geschehnissen. Im Gegensatz zur ludischen Dramaturgie sind die zeitlichen Abläufe der narrativen Dramaturgie einfacher planbar.⁸ Bei der Verwendung konventioneller Erzählweisen aus linearen Medien sind deren Höhepunkte häufig in Cut-Scenes, Skript-Sequenzen und Dialogen montiert. Ambitioniertere Erzählweisen, die an Stelle einer linearen Geschichte mehr Wert auf ein interaktives Erlebnis legen, erben die Planungswidrigkeiten der Dramaturgie des Spiels. Grundsätzlich überlagern sich im Computerspiel ludische und narrative Spannungsbögen. Ihr Verlauf und ihre Höhepunkte sind jedoch nicht unbedingt kongruent.

Selbst linear angelegte Zeit- und Erzählstrukturen werden häufig von der repetitiven Grundstruktur des Spiels eingeholt. So besteht ein Nebeneffekt bei zyklischen Zeitstrukturen darin, dass sich beim wiederhol-

⁸siehe auch 5.2

ten Durchspielen möglicherweise Bedeutungen von narrativen Komponenten verschieben. Wenn z. B. der Skelettkönig in *Diablo 3* den Bosskampf mit den Worten „You will never defeat me!“ einleitet, fühlen sich die Spieler beim ersten Mal vielleicht herausgefordert oder gewarnt. Im zweiten Durchgang erscheint ihnen die Aussage vielleicht unpassend, da sie ihn bereits beim ersten Durchspielen besiegten. Wenn der Skelettkönig im dritten Durchgang wieder vor ihnen steht und die Spieler mit den Worten „You will never defeat me!“ begrüßt, beschleicht sie das Gefühl, dass er Recht hat. Narrative Elemente durchlaufen im Geflecht der repetitiven Struktur des Spiels einen Bedeutungswandel.

3.2 Ereignisbasiertes Sounddesign

Aufgrund der Konstruktion zur Laufzeit in Abhängigkeit der Eingaben erfolgt in spielbaren Abschnitten und Menüs die Integration akustischer Elemente nicht in Form ihres Arrangements auf einer fixen Zeitachse. Stattdessen werden sie an Ereignisse und Zustände gebunden, die in unterschiedlichen Konstellationen und variablen Zeitabfolgen eintreten können. Dieser Umstand wirft sowohl für die Entwicklung, als auch für die Beschreibung der akustischen Dimension des Computerspiels eine Reihe von Schwierigkeiten auf. Entwickler sehen sich mit der Frage konfrontiert, wie weit bestimmte Konstellationen und Zeitabfolgen überhaupt planbar sind. Lässt sich die zeitliche Ungewißheit eingrenzen, so dass der entstehende Möglichkeitsraum innerhalb gesetzter Schranken vorhersehbar und kontrollierbar bleibt? Oder liegt gerade in der Variabilität die Chance einer freieren und dynamischen akustischen Gestaltung?

Für die kritische Auseinandersetzung mit dem Computerspiel und dessen akustischer Gestaltung ist die Kenntnis dieses Konstruktionsprozesses unabdingbar. Er stellt gewissermaßen das Pendant zur Montage des Films dar. Es gibt im Computerspiel jedoch keine zeitlich fixierten Szenen, die man untersuchen könnte, sondern vielmehr potentiell auftretende Konstellationen.⁹ Allein um die verhandelte Bedeutung akustischer Elemente zu erfassen, ist es notwendig, ihre Kopplung an Ereignisse und Zustände zu berücksichtigen. Die folgenden Abschnitte

⁹Cutscenes sind die Ausnahme. Siehe 5.2.

geben einen Überblick darüber, wie im Laufe des Entwicklungsprozesses akustische Bausteine und Parameter an Ereignisse und Zustände des Spiels gebunden werden.

Kopplung von Klangparametern und Ereignissen

Nicht nur im Computerspiel richtet sich die Vertonung nach Ereignissen. Wenn in einem Film eine Tür zugeschlagen wird, wird diesem Ereignis bei der nachträglichen Synchronisation ein entsprechendes Geräusch zugeordnet. Der Unterschied besteht jedoch darin, dass dort im Gegensatz zum Computerspiel der Zeitpunkt und der Kontext des Ereignisses bekannt sind. Wenn im Film ein Charakter nach einem niederschlagenden Gespräch eine Tür hinter sich schließt, kann das Sounddesign seine Gefühlslage antizipieren. Wenn die Spieler in der Rolle ihrer Avatare eine Tür in der fiktiven Welt schließen, können die jeweiligen Umstände hingegen ganz unterschiedlich sein - ebenso wie die Zeitpunkte. Während bei der Nachsynchronisation des Bilds der Film-Sounddesigner das Sample des Türgeräuschs an dem richtigen Zeitpunkt platziert, verknüpft der Computerspiel-Sounddesigner es lediglich mit dem Ereignis der schließenden Tür, ohne zu wissen, wann und wie oft es eintritt.

Grundsätzlich kann im Computerspiel jedes Ereignis und jeder Zustand vertont werden oder als Steuerelement für einen klanglichen Parameter wie Frequenz und Lautstärke fungieren. In der Regel ist es jedoch nicht praktikabel, alle Ereignisse und Zustände zu vertonen. Das ereignisbasierte Sounddesign birgt deshalb eine Selektion in sich. Einige Ereignisse werden vertont, andere bleiben stumm. In der Praxis haben sich einige Zuordnungen verfestigt und zu Konventionen verdichtet. Obwohl sich im Laufe der Geschichte des Computerspiels von der Klangerzeugung bis zu den produktionstechnischen Integrationsweisen viele Aspekte der akustischen Architektur verändert haben, blieben einige dieser Konventionen über Jahrzehnte intakt. Von wenigen Ausnahmen abgesehen, werden sämtliche Kollisionsergebnisse und alle Aktionen des Benutzers vertont. Die Vertonung der Aktionen dient als Feedback, das den Benutzern eine akustische Bestätigung ihrer Eingabe vermittelt. Sie ist vor allem dann unverzichtbar, wenn eine grafische Antwort nicht unmittelbar erfolgt. Vertonte Kollisionen bestätigen die Berührung zweier Dinge und bilden zentrale Steuerungspunkte in gra-

fischen Benutzeroberflächen. Die weitere Auswahl der zu vertonenden Ereignisse und Zustände richtet sich individuell nach den ludischen und narrativen Vorgaben des Computerspiels.

Die Kopplung von Klängen und Klangparametern mit Ereignissen stellt jene Schnittstelle dar, an der sich Sound und Code unweigerlich berühren. Letzten Endes taucht jede Soundoperation dieser Art als Befehl innerhalb des Code-Abschnitts auf, der beim Eintreten eines Ereignisses oder während der Dauer eines Zustands durchlaufen wird. Mit Hilfe von Audio-Middleware wie FMOD und Wwise ist es seit einigen Jahren möglich, die Zuordnungen von Klängen zu Ereignissen und Zuständen auch ohne Programmierkenntnisse durchzuführen. FMOD gliedert sich dazu in zwei Komponenten, die sich jeweils an unterschiedliche Expertenkreise richten. FMOD Designer ist ein Werkzeug für Sounddesigner, das speziell für die ereignisbasierte akustische Gestaltung ausgelegt ist. Die FMOD API ist eine Programmierschnittstelle, die den Audio-Programmierern die Nutzung des Sound-Systems ermöglicht. Auch wenn das ereignisbasierte Sounddesign in FMOD Designer außerhalb des Programmcodes definiert werden kann, erfordert die Einbindung letztendlich Eingriffe in den Code mit Hilfe der API. Geändert haben sich lediglich die Anforderungen an die Sounddesigner. Sie kommen mit weniger programmierspezifischem Wissen aus, um ein Spiel zu vertonen.

Wie gehen Sounddesigner nun mit den Ungewissheiten des ereignisbasierten Designs um? In den meisten spielbaren Abschnitten ist die zeitliche Abfolge der auftretenden Ereignisse schließlich nicht vorhersehbar. Dennoch sind zentrale Orientierungspunkte durch die Struktur des Spiels gegeben. Aufgrund der Repetition entstehen Muster, die immer wieder auftreten und deren Darstellung entsprechend gut planbar ist. In zeitkritischen Spielsituationen, die eine Reaktion der Benutzer erfordern, sind beispielsweise zwei wesentliche Ausgänge vorhersehbar. Entweder reagieren die Spieler rechtzeitig und meistern die Situation, oder sie reagieren zu spät und scheitern. Erfolg und Niederlage bilden wichtige Pole des Spiels, die nicht nur das letztendliche Ergebnis beschreiben, sondern den gesamten Spielverlauf durchdringen. Sie bilden die Grundpfeiler, auf denen verschiedene Gestaltungsvarianten aufgebaut sind.

Das Sounddesign antizipiert damit Teile der Programmierung des Spiels. Indem es die Eventualitäten der Spieldausgänge widerspiegelt, vertont

es nicht nur Sieg und Niederlage des Spiels, sondern auch *if* und *else* des Codes. Je detaillierter die programmierten Spielmodelle ausfallen, desto feiner kann auch das Ereignisgranulat und das daran ausgerichtete Sounddesign gestaltet sein. So treten Erfolge und Niederlagen in Hinblick auf ausgeführte Aktionen in der Regel in mehreren Abstufungen auf. Ein „kritischer“ Treffer¹⁰ wirkt sich deutlich positiver aus als ein „normaler“ Treffer und kann durch die Anpassung akustischer Parameter entsprechend hervorgehoben werden.

Gleichzeitig auftretende Ausgaben

Nicht nur das horizontale, sondern auch das vertikale Arrangement akustischer Ausgaben bringt Herausforderungen und Unwägbarkeiten mit sich. Die Vertonung gleichzeitig auftretender Ereignisse und Zustände unterliegt zwei Beschränkungen. Zum Einen gibt es eine technische Grenze, die bestimmt, wie viele Klänge übereinander geschichtet werden können. Zur Zeit der Soundchips war es die Anzahl der verfügbaren Kanäle. Heute sind es die für den Sound reservierten Rechen- und vor allem Speicherkapazitäten. Zum Anderen gibt es eine wahrnehmungsbedingte Grenze hinsichtlich des differenzierten Hörens übereinander geschichteter Sounds. Wird diese Schwelle überschritten, sind die funktionalen Aufgaben der akustischen Gestaltung stark beeinträchtigt. Die Benutzer bekommen einige Informationen vielleicht gar nicht mehr mit. Nichtsdestotrotz kann die gezielte Überladung und der damit einhergehende Informationsverlust als stilistisches Mittel fungieren.¹¹

Eine typische Gegenmaßnahme für die technisch auferlegte Grenze ist die Kombination mehrerer Klänge in einem Kanal oder Sample. Viele Chiptunes fassen z. B. den Bass und das Schlagzeug in einer Stimme zusammen, und auch die Samples aktueller Soundscapes enthalten Geräuschgemische. Um die Wahrnehmungsschwelle nicht zu überschreiten, sind Priorisierungen der auftretenden akustischen Elemente ein probates Mittel. Akustische Elemente mit niedrigen Prioritäten können bei

¹⁰Kritische Treffer sind in mehreren Genres wie z. B. dem Action-Adventure wichtige Komponenten des Spielmodells. Die typische Implementierung sieht so aus, dass der Avatar bei jedem Schwerthieb einen Schadenswert verursacht. Ein normaler Treffer entspricht 100 Prozent des Schadenswertes, während ein kritischer Treffer einen höheren Schadensanteil - beispielsweise 150 Prozent des Schadenswertes - ausmacht.

¹¹siehe auch 4.4 auf Seite 170

einer hohen Anzahl von Klangschichten ausgeblendet bleiben oder innerhalb des Mixes zurücktreten, indem ihre Lautstärke reduziert wird.

Auch wenn die Abstimmung verschiedener gleichzeitig auftretender Klänge bereits seit Jahrzehnten eine Herausforderung des Audio-Designs von Computerspielen ist, scheint die zunehmende Komplexität der Spieldarstellungen diese Herausforderung noch vergrößert zu haben. Je detaillierter die akustische Gestaltung ausfällt, und je feiner die vertonten Ereignisse segmentiert sind, desto herausfordernder ist es, einen ansprechenden Mix in Echtzeit zu realisieren.

Akustische Elemente sind selten die einzigen Darstellungsartefakte, die an ein Ereignis oder einen Zustand gebunden sind. In den meisten Fällen treten sie in Kombination mit Animationen auf. Aus der Kopplung akustischer und grafischer Ausgaben an Ereignisse entstehen audiovisuelle Blöcke, die in vielen Fällen exakt reproduzierbar sind. Wenn die Spieler in einem Jump'n'Run die Sprungtaste drücken und der Spielfigur keine Hindernisse im Weg sind, resultiert die Aktion in der Ausgabe des gleichen audiovisuellen Blocks. Ein audiovisueller Block besteht entweder aus einer eindeutigen Verbindung von Bild und Ton oder mehreren möglichen Zusammensetzungen. Vor allem bei Ereignissen, die sehr häufig auftreten, kommen Klanggruppen zum Einsatz, aus denen jeweils ein Element zufällig ausgewählt wird.

3.3 Wahrnehmungshilfen zeitlicher Abläufe

Musik ist Gestalt in der Zeit, und unser Hörsinn ist unser zeitlich genauester Sinn. Mit ihm können Zeitstrukturen erfasst werden, die anderen Sinnen verborgen bleiben.¹²

Da der Sinn des Hörens eine um „zwei bis drei Zehnerpotenzen größere Schnelligkeit“¹³ besitzt als das Sehen, trägt die akustische Gestaltung in zeitkritischen Spielsituationen maßgeblich zur besseren Einschätzung des zeitlichen Ablaufs bei. Unmittelbar deutlich wird diese zeitliche Diskrepanz unserer Sinneswahrnehmung, wenn man die Auflösung von Animationen und akustischen Elementen nebeneinanderstellt. Um dem

¹²(Spitzer, 2002, 18)

¹³(Spitzer, 2002, 61)

Auge einen flüssigen Bildablauf vorzugaukeln, gelten 24 Bilder pro Sekunde als Standard. Unser Gehör verlangt von digitalen Aufnahmen hingegen eine deutlich höhere Samplerate. Der für Audio-CDs genutzte Standard verwendet 44.100 Samples pro Sekunde.

Akustische Signale und Intervalle werden durch die bessere zeitliche Wahrnehmung unseres Hörvermögens schneller registriert als dargestellte Animationen und verschaffen den Spielern dadurch Vorteile. Die vermittelten Informationen lassen sich in Hinblick auf ihre Zeitlichkeit in drei Kategorien unterteilen. Sie unterstützen die Spieler bei der Einschätzung von *Zeitpunkten*, *Zeitspannen* und *Geschwindigkeiten*.

Timing: Den richtigen Zeitpunkt hören

Bei nahezu allen Computerspielen, die in Echtzeit ablaufen, nimmt das Timing der Eingaben eine herausragende Stellung ein. Gefragt ist die richtige Aktion zur richtigen Zeit. Oft sind es nur kurze Momente - vor allem in kompetitiv ausgerichteten Spielen - die über Siege und Niederlagen entscheiden. Die Vertonung dieser Momente, bzw. der in ihnen auftretenden Ereignisse und Aktionen, trägt maßgeblich zu ihrem Verständnis bei. Dabei ist es primär das akustische Zusammenspiel zwischen den von den Spielern ausgelösten Aktionen und den von der Spielwelt oder gegnerischen Akteuren ausgehenden Ereignissen, die den zeitlichen Ablauf interpunktieren und womöglich auch beeinflussen. Wenn die Spieler das Bröckeln einstürzender Plattformen, fallender Kisten, Schüsse oder Sprunggeräusche hören, helfen ihnen diese vertonten Momente bei der Einschätzung des Spielablaufs. Je besser die Spieler mit den jeweiligen Spielen vertraut sind, desto präziser können sie die akustischen Elemente ihren Ursprüngen zuordnen und die verbleibende Zeit für (Re-)Aktionen einschätzen.

Während in realen Sportarten das richtige Timing mit komplexen Bewegungsabläufen als Ergebnis jahrelangen Trainings einhergeht, sind in Computerspielen die motorischen Abläufe zur Bedienung des Hardware-Interfaces relativ simpel. Bei Tastatur, Maus oder Gamepad beschränkt sich die nötige zeitliche Abstimmung auf das Drücken eines Knopfes zur richtigen Zeit. Wenn jemand einen grünen Knopf drückt und dabei einen Analogstick in eine Richtung bewegt, anstatt mit seinem Körper einen Pass zum Mitspieler zu realisieren, nimmt das Timing prinzipiell

eine unberührtere Rolle ein. Es ist weniger eine motorische Herausforderung als eine geistige. Simple Eingaben lenken das Augenmerk auf die Ausgabe und den virtuellen Raum des Spielgeschehens. Neue Interface-Entwicklungen wie die *Wii Mote*, *PlayStation Move* oder *Kinect* brechen diese Betonung des Interaktionsvorgangs auf, indem sie den Körper der Spieler stärker einbeziehen. Nicht nur die Ausgabe des Programms steht im Mittelpunkt, sondern auch die zur Performance werdende Eingabe. Vor allem bei Party-Spielen wie *Raving Rabbids* oder Musik-Spielen wie *Rockband* und *Singstar* ist das, was sich vor den Bildschirmen und Boxen abspielt, ebenso faszinierend wie das Geschehen innerhalb des virtuellen Raums. Erfahrene Spieler neigen allerdings auch hier dazu, ihre Bewegungen zu minimieren. Um mit der *Wii-Mote* einen Tennisaufschlag durchzuführen, müssen die Spieler den Controller nicht über den Kopf schwingen. Eine Aufwärtsbewegung des Handgelenks reicht dazu aus. Je stärker sie darauf aus sind, das Spiel zu gewinnen, desto ökonomischer und effizienter gestalten sich ihre Bewegungen. Die Spieler treffen ihre Eingabe-Entscheidungen als Antwort auf die Darstellung des Spielgeschehens.

Sowohl bei konventionellen Eingabegeräten wie Gamepads und Tastaturen, als auch bei den neuen körperbetonen Interfaces ist nicht nur die Visuomotorik, insbesondere die Hand-Augen-Koordination gefragt. Die Einbeziehung akustischer Elemente und ihr interaktionsleitendes Spiel aus Feedback und Suggestion legt vielmehr eine Hand-Augen-Ohren-Koordination nahe. Die Überlegenheit des Gehörs in der Wahrnehmung zeitlicher Abläufe suggeriert überdies, dass dessen Anteil an diesen Prozessen keineswegs gering ist.

Powerup: Sonifikation von Zeitspannen

Ebenso wie akustische Inhalte mit der Bindung an Ereignisse bestimmte Zeitpunkte markieren, transportieren sie durch die Bindung an Zustände wichtige Informationen über zeitliche Intervalle. Besonders Musik ist durch ihre beliebig variierbare Länge dazu prädestiniert, die Länge eines Zustands abzubilden. Das Gleiche gilt für wiederholt abspielbare Geräusche wie Alarmsirenen und Herzklopfen. Ihr Abspielmoment informiert über den Zeitpunkt, zu dem der ausgewählte Zustand eintritt. Die Abspiellänge entspricht der Zeitspanne, in welcher der Zustand aktiv bleibt. Sobald die Spieler das an den Zustand gekoppelte



Abbildung 3.2: Motiv A und Motiv B aus *Donkey Kong*

akustische Element hören, erfahren sie von dem Inkrafttreten dieses Zustands. Wenn die Dauer eines wiederholt eintretenden Zustands konstant bleibt, dann ermöglicht dessen Sonifikation den Spielern darüber hinaus eine Einschätzung, wie lange er noch aktiv bleibt.

Ein typisches Beispiel für die explizite Bindung von Musik an einen Zustand bietet das Powerup. Für einen beschränkten Zeitraum verleiht es der Spielfigur bedeutende Kräfte wie Unverwundbarkeit oder die Vervielfachung des ausgeteilten Schadens. Aktivierte Powerups greifen direkt in das Regelsystem ein. Für die Dauer ihrer Wirkung gelten andere Gesetze, die nicht selten das Spielprinzip umkehren. Der Gejagte wird zum Jäger und der Jäger zum Gejagten. In *Donkey Kong* (1981) beginnt die chancenreiche Powerup-Phase, sobald die Spielfigur Jumpman einen Hammer aufnimmt. Anstatt den herabrollenden Fässern auszuweichen müssen die Spieler sie nun zertrümmern, Springen und Leitern steigen ist nicht möglich. Während dieser Zeitspanne ersetzt das in Abbildung 3.2 notierte Motiv B das Standardmotiv A. Die höhere Geschwindigkeit des eingeschobenen Motivs unterstreicht die Dringlichkeit der Spielsituation. Bevor sich dieses kairotische Zeitfenster wieder schließt, sollten die Spieler die Gelegenheit beim Schopf packen und den größtmöglichen Nutzen daraus ziehen. In den bis heute gepflegten Spielereihen von Nintendo, in denen der hammerschwingende Jumpman zum Klempner Mario wurde, blieb die musikalische Untermalung von Powerups durchgehend erhalten. Das musikalische Thema des Stern-Powerups ist seit Jahrzehnten ein fester Bestandteil des Mario-Universums.¹⁴

Frühe Jump'n'Runs wie *Super Mario Brothers* und *Giana Sisters* beschränkten die Zeit, in der ein Level beendet werden mußte. Auch hier hilft die Musik den Spielern dabei, die verbleibende Zeit einzuschätzen. In *Super Mario Brothers* startet bei Beginn eines Levels ein Countdown, der von 400 auf Null herunterzählt. Sobald drei Viertel der Zeit verstrichen sind, betont die Erhöhung des Musiktempos den Zeitdruck.

¹⁴siehe <http://www.mariowiki.com/Star> (14.03.2012)

Je komplexer das Modell des Spiels ausfällt, desto mehr Zustände stehen zur Verfügung, die mit Musik koppelbar sind. In dem Stealth-Shooter *Splinter Cell* ist die Musik an den Zustand endlicher Automaten gebunden, die das Verhalten von NPCs modellieren. Schöpft eine gegnerische Figur Verdacht und beginnt damit, die Umgebung nach dem Avatar der Spieler zu durchsuchen, startet ein neuer Loop. Bleibt die Suche erfolglos, springt der endliche Automat wieder in seinen Ausgangszustand zurück, und die Figur kehrt ahnungslos zu ihrer anfänglichen Routine zurück, während die Musik verstummt. Rollenspiele wie *Gothic* und *Risen* nutzen ein ähnliches Prinzip. Sobald ein NPC auf den Spieleravatar aufmerksam wird und sich für den Angriff bereit macht, antizipiert die Musik den Übergang in die Kampfphase.

In beiden Fällen vermittelt die Musik Informationen, die die Spieler ansonsten erst zu einem späteren Zeitpunkt wahrnehmen würden und sie nun zu ihrem Vorteil ausspielen können. Sie wissen in *Risen*, dass sie angegriffen werden, auch wenn sie den auf sie zustürmenden Angreifer nicht sehen. Genau so gut wissen sie, dass noch nicht alle Antagonisten aus dem Weg geräumt sind, so lange die Musik weiter spielt. Unmittelbar verbunden mit diesem Wissen sind die dadurch suggerierten Handlungsweisen. Bei der Überleitung in eine Kampfphase werden sich die Spieler im Raum bewegen und um die eigene Achse rotieren, um alle gegnerischen Figuren auszumachen. In *Splinter Cell* hingegen werden sie sich vor dem suchenden NPC verstecken oder ihn in einen Hinterhalt locken.

Durch die Schichtung von Klangebenen ist die Betonung mehrerer Zustände möglich, die sich zeitlich überschneiden. In dem Rollenspiel *The Witcher* signalisiert die Musik ebenso wie in *Risen* den Start, die Dauer, und das Ende eines Kampfes. Sinkt innerhalb dieser Phase der Gesundheitswert¹⁵ des Avatars unter eine bestimmte Grenze, erklingt zusätzlich zu der Musik ein wiederholt abgespieltes Sample mit dem Geräusch eines klopfenden Herzens. Ebenso, wie hier die Musik und ein Soundloop an unterschiedliche Zustände gekoppelt sind, könnte Musik alleine mehrere Zustände betonen, sofern sie als layerbasierte Komposition vorliegt.

¹⁵Der Gesundheitswert ist nicht nur für das Herzklopfen ein beliebter Sound-Auslöser. In *Street Fighter IV* triggert ein niedriger Gesundheitswert eine Musiktransition.

Agogische Synchronisation: Sonifikation von Geschwindigkeit

Veränderungen der Spielgeschwindigkeit sind in Computerspielen weit verbreitet. Sie können, wie in *Tetris*, durch das Spielsystem gesteuert sein, um die Schwierigkeit zu erhöhen, oder der Kontrolle der Spieler unterliegen. Möglicherweise können die Spieler die Zeit der fiktiven Spielwelt beschleunigen und verlangsamen, sie anhalten, vorspulen und zurückspulen. Musik ist ein Mittel, um diese Manipulationen der Zeit und Geschwindigkeit sinnlich erfahrbar zu machen.

Zur Sonifikation von Geschwindigkeitsänderungen wird das Tempo der Musik an das Tempo eines Spielaspekts, z. B. die Bewegungsgeschwindigkeit eines Akteurs, gekoppelt. In der Musik bezeichnet man die Lehre von der Gestaltung des Tempos mit dem Begriff Agogik.¹⁶ Typische agogische Gestaltungsmuster sind das *Accelerando* und das *Ritardando*, die graduelle Beschleunigung respektive Verlangsamung eines Abschnitts. Abgeleitet von dieser Terminologie wird die Kopplung des musikalischen Tempos an das Tempo eines Spielvorgangs im Folgenden als *agogische Synchronisation* bezeichnet.

Eine prominente Anwendung solch einer agogischen Synchronisation findet sich bereits in dem 1978 erschienenen Arcade-Bestseller *Space Invaders*. Je schneller sich die Invasoren bewegen, desto kürzer werden die Basstöne Ab, B, Bb, A des sich kontinuierlich wiederholenden Motivs. Es legt gleichzeitig mit den beschleunigten Animationsbewegungen deutlich an Tempo zu. Die Musik hilft den Spielern dadurch, die Geschwindigkeit der Invasoren besser einzuschätzen. Sie können die höhere Auflösung ihres Hörsinns zur Erkennung zeitlicher Muster einsetzen, anstatt alleine aus der visuellen Beobachtung der Animationen die Bewegungsgeschwindigkeit abzuleiten.

Die Geschwindigkeitsveränderung selbst war in *Space Invaders* ursprünglich nicht vorgesehen. Wie der Game Designer Tomohiro Nishikado ausführte, ergab sie sich aus technischen Einschränkungen:

Originally I wanted to move 55 Invaders at the same time,
but the hardware only let me move one Invader every 1/60

¹⁶(Riemann, 1884)



Abbildung 3.3: Die Geschwindigkeit des Basslaufs steigt in *Space Invaders* von 70 bpm bei 55 Figuren auf 172 bpm bei 15 Figuren.

second. As a result, Invaders began to move faster as they decreased in number.¹⁷

Wie man in Entwicklerkreisen zu sagen pflegt: It's not a bug, it's a feature. Obwohl es scheint, als ob sich die Geschwindigkeit der Invasoren und der Musik steigert, je weiter die Reihen nach unten rücken, sind also die Spieler die Auslöser. Wenn sie nicht schießen, behalten die Invasoren ihren langsamen Rhythmus bei. Mit jeder eliminierten Figur entlasten die Spieler die Maschine und beschleunigen die Bedrohung.

Viele Arcade-Spiele, die nach *Space Invaders* auf den Markt kamen, nutzten graduelle Tempoveränderungen der Musik, ohne dass die Geschwindigkeit eines Spielaspekts daran gekoppelt war. In diesen Fällen suggeriert die Geschwindigkeitssteigerung eine gewisse Dringlichkeit, die die Spieler unter zusätzlichen Druck setzt. Es ist durchaus denkbar, aber nicht empirisch bestätigt, dass dieser psychologische Trick Fehler von Seiten der Spieler verursachte und damit die Rentabilität der münzbetriebenen Automaten erhöhte.

¹⁷(Nishikado, 2004, 35)

Eine Variante agogischer Synchronisation nutzt das 2009 von Hemisphere Games veröffentlichte Spiel *Osmos*, in dem die Spieler mit einer Tastenkombination die Zeit der Spielwelt stauchen und dehnen können. In Abhängigkeit dieser Zeitmanipulationen verändern sich die Parameter eines Pitch-Shifting Effekts. Je langsamer die Zeit eingestellt ist, desto tiefer und langsamer erklingt der Ambient-Loop, während die Einstellung eines zügigeren Tempos seine Tonlage und Geschwindigkeit entsprechend anhebt. An den Endpunkten des Spektrums bewirkt der Effekt eine maximale Abweichung von einem Halbton nach unten bzw. einem Ganzton nach oben von der Originaltonlage. Mit den über 30 einrastenden Zwischenstufen bieten sich den Spielern sehr subtile Eingriffsmöglichkeiten in das Tempo und die Tonhöhe. Die Modifikation der Tonlage und die damit einhergehende Streckung oder Stauchung der Abspielzeit tragen nicht nur dazu bei, eine tiefere Verbindung zwischen der Bewegungsanimation und der Musik zu etablieren, sondern legen den Spielern ein effektiveres Mittel nahe, um mit ihren Sinnen den Grad der Zeitmanipulation wahrzunehmen.

I think pitch-shifting the music and sound effects helps players track time-warping in an intuitive way. Without these aural cues, the mental connection would be incomplete.¹⁸

Die Stärke des klassischen Pitch-Shifting Effekts, welcher ursprünglich durch das schnellere oder langsamere Abspielen von Tonbändern erzeugt wurde, liegt in dem Fall gerade in seiner Beeinflussung mehrerer klanglicher Parameter. Neben der angepassten Geschwindigkeit dient auch die entsprechend veränderte Tonlage als Wahrnehmungshilfe des Animationstempos.

3.4 Musik im Spiel

Ein definierendes Merkmal von Musik ist die zeitliche Organisation von Klangstrukturen. Anders als ein Bild, das mit einem Blick wahrgenommen werden kann, erfahren wir Musik über eine Zeitspanne. Während die Zeitspanne eines Stücks bei Live-Darbietungen variieren kann, ist sie bei einer Aufnahme immer identisch. Bei der Vertonung von Computerspielen herrscht nun eine hybride Situation vor. Einerseits gibt es mit

¹⁸Eddy Boxerman in (Kirn, 2009)

den Samples bzw. vorprogrammierten statischen Sounds gespeichertes Klangmaterial, das immer wieder abgespielt werden kann. Andererseits ähnelt die Zusammensetzung des Materials einer Live-Darbietung. Wenn die Band merkt, dass das Publikum begeistert mitsingt, wiederholt sie vielleicht den Refrain ein weiteres Mal oder schiebt die anschließende Ballade in der Setlist weiter nach hinten, um den Tanztrieb des Publikums nicht zu stoppen. Die Band beobachtet die Stimmung im Saal und trifft auf Grundlage dieser Beobachtungen Entscheidungen über den weiteren Verlauf. Analog dazu analysiert das Programm die Spielsituation und richtet je nach programmierter Vorgabe die Musik daran aus. Neben dem abspielbaren Musikmaterial enthält das Programm also Routinen, die die Wiedergabe und Zusammensetzung des Materials kontrollieren und steuern.

Die größte Herausforderung für die Integration und Verknüpfung des Musikmaterials mit dem Spielgeschehen ist die zeitliche Ungewißheit des Spielablaufs. Abgesehen von sehr stark eingeschränkten Interaktionsspielräumen ist unklar, zu welchem Zeitpunkt die Spieler welche Aktionen ausführen. Während im Film bis auf die letzte Millisekunde Klarheit darüber herrscht, was wann dargestellt wird, verlangt die Synchronisation von Musik und dem Spielgeschehen einen höheren Planungsaufwand. Ausgehend von dieser Problematik sind verschiedene Lösungen zur Integration von Musik entstanden, die sich sowohl in ihrer musikalischen Konzeption als auch in ihren Anbindungsmöglichkeiten an Narration und Spiel unterscheiden. Eine grundlegende Differenz dieser Lösungen liegt in der Länge der Musikbausteine, ihren Kombinationsregeln und ihren Kopplungen an Spielereignisse. Die folgenden Abschnitte demonstrieren ausgehend von diesen Parametern unterschiedliche Arten der Einbindung von Musik im Computerspiel. Sie zeigen die grundlegenden Strategien auf, um Musik in spielbare Abschnitte mit ungewissem zeitlichen Verlauf zu integrieren. Darüber hinaus wird diskutiert, an welchen narrativen und ludischen Aspekten sich die Kompositionen orientieren.

Strukturbasierte Musikbegleitung: Ein Lied für ein Level

Eine der gebräuchlichsten Realisierungen von Musik im Computerspiel sieht vor, dass ein Musikbaustein einem Stück entspricht. Auf diese

Weise sind keine Kombinationsregeln notwendig, die aus der Verknüpfung mehrerer Bausteine eine zusammenhängende musikalische Struktur erzeugen. Um das auskomponierte Stück in das Spiel einzubinden, wird es in den meisten Fällen an eine Strukturkomponente wie den Titelbildschirm, das Menü oder einen spielbaren Abschnitt gekoppelt - ein Lied für ein Level. Vor allem in den 1980er Jahren war diese Musikkonzeption im Rahmen einer kontinuierlichen, alle Bereiche des Spiels begleitenden Musikuntermalung stark vertreten. Für die kurze Spieldauer von Arcadeautomaten waren kurze Musikbausteine ausreichend. Mit der Entwicklung von komplexeren Spielen für Konsolen und Heimcomputer stieg die Anzahl der Strukturkomponenten und die Länge der Spielzeit. An die Stelle eines Stücks trat die Playlist. Auch die kompositorische Ausgestaltung der Stücke durchlief einen Wandel, dessen Rahmenbedingungen die in Kapitel 2 aufgezeigten technischen Entwicklungen bildeten.¹⁹

Die Kopplung eines auskomponierten Musikbausteins an ein Strukturelement läuft nach folgendem Schema ab: Beim Betreten der jeweiligen Strukturkomponente startet das Programm den Abspielvorgang des zugeordneten Bausteins und stoppt ihn beim Verlassen derselbigen. Abgesehen von zeitlich determinierten Spielformen wie dem Rail-Shooter²⁰ ist jedoch unklar, wie lange die Spieler zur Bewältigung des Levels brauchen. Sie könnten das Level beendet haben, bevor die Musik das Ende erreicht oder nachdem der Musikbaustein bereits abgespielt wurde. Mögliche Strategien zur Kaschierung dieser Variabilität sind Wiederholungen der Musikbausteine, Beschränkungen der Aufenthaltszeiten in einem Level²¹ sowie Fade Outs, Pausen und kurze Transitionsbausteine zur Glättung der Übergänge.

Neben der Herausforderung, die Übergänge ansprechend zu gestalten, gibt es lange Abschnitte, in denen Musik und Spielgeschehen unbeein-

¹⁹ Abgesehen von der gestiegenen Quantität des musikalischen Materials und technisch bedingten Veränderungen erlebte die Musik des Computerspiels innerhalb der vergangenen Dekaden einen Wandel, der auf die Kompositionspraxis der jeweiligen Zeit zurückzuführen ist. Die Musik der 1980er favorisierte andere Formen als die Musik der 1990er Jahre, was sich natürlich auch in der Spielmusik niederschlug.

²⁰ In diesem Genre bewegt das Programm den Avatar - typischerweise ein Raumschiff - wie auf Schienen, durch das Level. Wenn die Spieler weder das Tempo, noch die Position verändern können, herrscht zu jedem Moment Klarheit darüber, wo sich der Avatar befindet und wie lange er für das Durchqueren des Levels braucht.

²¹ In *Super Mario Brothers* und *Great Giana Sisters* ist dieses Prinzip z. B. durch einen eingeblendeten Countdown realisiert.

trächtig voneinander ablaufen. Die Musik besitzt keine Möglichkeit, auf Ereignisse zu reagieren und variabel eintretende Aspekte des Spielgeschehens unmittelbar zu kommentieren. Spielereignisse und -zustände nehmen keinen Einfluss darauf, wie sie sich weiterentwickelt. Synchronisationspunkte oder -intervalle zwischen der musikalischen Struktur und den Spielereignissen sind nicht geplant. Während die Musik läuft, vollzieht sich unabhängig davon die zeitliche Abfolge der Ereignisse.

Nicht immer ist die Musik so integriert, dass sie unterschiedliche strukturelle Abschnitte wie Levels, Maps oder Menüs klar voneinander abgrenzt. Manchmal läuft ein Stück über mehrere Abschnitte und verbindet sie damit. Wenn der Übergang zwischen zwei Strukturabschnitten nicht als Transitionspunkt in einen neuen Musikabschnitt herhält, schafft die Musik Kontinuität. Diese bindende Funktion von Musik ist vergleichbar mit Filmmusik, die mehrere Schnitte überlagert und sie dadurch zusammenhängend erscheinen lässt. Im Computerspiel wird diese Technik häufig dazu eingesetzt, um Ladevorgänge zu kaschieren.

Ein Vorteil der strukturbasierten Integration von kompletten Liedern liegt in der einfachen Vergabe von Auftragskompositionen, da bzgl. der Abspiel-Ereignisse keine weitere Abstimmung zwischen Komponisten und Programmierern erforderlich ist. Ebenso einfach wie die Integration in das Spiel gestaltet sich die Auskopplung der Musik aus dem Spiel heraus. In der Regel sind nur geringe Anpassungen nötig, um sie als Game Soundtrack separat zu vertreiben. Als grundsätzlich linear abspielbarer Datensatz entspricht das Format der Musik den Konventionen der Unterhaltungsindustrie. Sobald vermehrt Prozesse hinzutreten, wie in den nachfolgend vorgestellten Musikkonzeptionen, beginnt diese Kompatibilität zu bröckeln, auf deren Basis Musik einer medienübergreifenden Verwertung ausgesetzt ist.

Welche Funktionen kann die Musik in den spielbaren Abschnitten erfüllen, wenn sie an keine Ereignisse darin gebunden ist? Ohne die Kopplung an Spielzustände vermittelt sie offenbar keine Informationen, die für das Spiel an sich relevant sind. Die im vorangegangenen Abschnitt behandelten Funktionen akustischer Elemente als Wahrnehmungshilfen zeitlicher Abläufe stehen bei Diskussionen und Abhandlungen über die Musik des Computerspiels meistens im Schatten eines anderen Bedeutungskomplexes, der hier zum Tragen kommt. Musik wird eine große emotionale Wirkung nachgesagt. Sie bewirkt demnach deutlich mehr, als nur Informationen zu übermitteln. Mit Hilfe von kulturellen Codes

und Konventionen evoziert sie Stimmungen, Assoziationen und Emotionen und hilft dabei, die fiktionale Spielwelt und ihre Akteure zu charakterisieren.²² Gerade in medialen Kontexten, in denen Musik einen supplementären Charakter aufweist, überwiegt ihre Positionierung als Atmosphärenerezeuger und emotionaler Amplifikator. Lange Musikbausteine erlauben es den Komponisten, an die Narrative des Spiels anzudocken. So greifen sie z. B. in ihren Kompositionen Aspekte der virtuellen Welt auf wie das Setting eines Levels oder versuchen, bestimmte Stimmungen zu erschaffen.

Eine Spielart davon ist der Transport bzw. die Weiterführung narrativer Inhalte von Cut-Scenes in Spielabschnitte, die einer Erzählung wenig Raum eröffnen. Ein repräsentatives Beispiel dafür enthält das Echtzeitstrategiespiel *Starcraft II*, in dem drei verschiedene Fraktionen – Terraner, Zerg und Protoss – gegeneinander antreten. Das Spiel verwendet für die Integration der Musik eine Variante des strukturbasierten Ansatzes. Jede Fraktion erhält eine Reihe von auskomponierten Stücken, die während einer Partie längere Spielpassagen begleiten. Wählt man die Terraner, besteht die Musik aus mehreren Rock-Tracks mit leichten Country-Einschlägen. Die Musik der beiden außerirdischen Völker grenzt sich durch ihre Struktur und Instrumentierung deutlich davon ab. Die Tracks der aggressiven, insektoiden Zerg enthalten dissonante Flächen, wabernde synthetische Lead-Sounds und schnelle, häufig ungerade Drum & Bass Muster. Flirrende Geräusche durchbrechen in ruhigen Passagen immer wieder die liegenden Flächen. Durch wärmere Flächenounds, weniger Dissonanz und den weitgehenden Verzicht auf perkussive Elemente erhalten die Stücke der Protoss einen behäbigeren, verträumten Charakter.

Die Kompositionen der Fraktionen erfüllen narrative Zwecke, indem sie die in der Rahmenhandlung etablierten Grundzüge und Charakteristika der Fraktionen aufgreifen und in die spielbaren Abschnitte übertragen. Die Country-Licks der Terraner-Tracks verweisen auf den Rebellenfüh-

²²Nichtsdestotrotz scheint diese Dimension der Musik im Computerspiel weitaus weniger stark entwickelt zu sein, als beim Film. Ein wichtiger Unterschied zwischen Filmen und Computerspielen liegt in dieser Hinsicht darin, dass emotionale Höhepunkte mitunter andere Ursachen haben. In den seltensten Fällen sind es die dargestellten Szenen des Geschehens, sondern die Ausgänge von Spielsituationen. Jubel und Ärger bei Sieg und Niederlage sind die am deutlichsten beobachtbaren Ausbrüche der Spieler. Im Grunde genommen überlagern sich auch hier wieder die Strukturen des Spiels und der Erzählung, die in unterschiedlichen Facetten von der Komposition aufgenommen, reflektiert und erweitert werden.

rer Jim Raynor, der im Starcraft-Universum die Rolle des einsamen Cowboys spielt. Analog dazu spiegelt die Musik der Zerg und Protoss narrative Zuschreibungen ihres Charakters wider: ungezähmt, aggressiv und unberechenbar respektive hochmütig, stolz und mächtig.

Eine weitere Funktion ist die Antizipation des Gameplays. Musik kann eine durch die Spielmechanik vorgegebene Spielweise aufgreifen und betonen. Neben den angedeuteten narrativen Referenzen durch bestimmte Musikstile lassen sich in *Starcraft II* die Kompositionen darauf beziehen, wie sich die Fraktionen spielen lassen. Ein zentrales spielerisches Feature von *Starcraft II* ist gerade das ausgewogene Balancing der drei Fraktionen. Obwohl sie unterschiedliche Einheiten und Gebäude mit sich bringen und andere Taktiken erfordern, ist keine Partei den anderen beiden überlegen. Die Terraner sind dabei die Fraktion, die am ehesten dem Genre-Standard entspricht. Von daher ist es sehr passend, dass die ihnen zugewiesene Musik am stärksten auf Konventionen populärer Genres wie Rock und Country beruht. Die Einheiten der Zerg sind günstig und schneller zu produzieren, im Gegenzug aber auch schwächer. Wer mit den Zerg spielt, muß auf Masse setzen und viel Micromanagement betreiben. Die Zerg-Kompositionen unterstreichen diesen hektischer ausfallenden Spielstil. Die in der Starcraft-Erzählung als stolzes und hochtechnisiertes Volk beschriebenen Protoss besitzen die stärksten, aber auch die teuersten und schwerfälligsten Einheiten. Behäbig anschwellende Klangflächen und Choreinsätze spiegeln diesen Charakter wider.

Auch ohne die dezidierte Kopplung musikalischer Strukturen mit Zeitpunkten und -intervallen des Spielgeschehens legt die Musik bestimmte Spielweisen nahe. Ein treibender Heavy Metal-Track wird anders aufgenommen als ein Ambient-Track, Spieler adaptieren möglicherweise das Tempo und den Rhythmus des Stücks und richten ihre Spielhandlungen daran aus. Bei Spielen, die in Echtzeit ablaufen, hat die Antizipation der Musik durch die Spieler schwerwiegendere Folgen. Ein Beispiel hierfür bietet der First Person Shooter *Quake 2* von id Software. Der durchgängig temporeiche, von Sonic Mayhem komponierte Soundtrack lädt die Spieler dazu ein, durch die Level zu hasten. So heißt es in einem Kommentar zu dem Song „Descent into Cerberon“:

This is the song you play when you want to sprint through a level. You do not WALK through a level while playing this.

No, you move through the level like a mad righteous bastard on a mission from god to deliver lead to the unbelievers.²³

Die Musik trägt in *Quake 2* einen großen Teil dazu bei, das Spiel tempo-reich zu gestalten, obwohl ein schnelles Spiel in Hinblick auf die Spielmechanik nicht ausdrücklich erforderlich ist. Die Spieler könnten theoretisch auch langsam und vorsichtig vorgehen. Das hohe Tempo und der durchgängige Rhythmus suggerieren jedoch eine andere Handlungsweise.

Die durch Musik suggerierten Spielweisen können freie Angebote bleiben, die die Spieler freiwillig annehmen, oder durch entsprechend koordinierte Welt-Ereignisse durch die Entwickler forciert sein. In *Mirror's Edge* besitzt jedes Level beispielsweise Phasen der Erkundung und Phasen der Flucht, die von ruhiger, respektive hektischer elektronischer Musik begleitet werden. In den Erkundungsphasen sind keine gegnerischen Figuren vorhanden, während sie in den Fluchtphasen die Spieler zum schnellen Handeln zwingen.

Adaptive Musik

Viele Komponisten, die heute Musik für Computerspiele schreiben, haben in der Vergangenheit Musik für Filme komponiert oder sind für beide Medien tätig. Aufgrund der zeitlichen Fixiertheit der Tonspur ist es im Film möglich, bestimmte Szenen präzise mit speziell dafür ausgearbeiteten Musikpartien zu synchronisieren. Im Computerspiel ist dies offensichtlich nicht möglich. Um dem Ideal der Filmmusik näher zu rücken und auch in Computerspielen eine gezielte musikalische Untermauerung ausgewählter Momente zu ermöglichen, entwickelten die Komponisten unterschiedliche Konzeptionen sogenannter *Adaptiver Musik*. Wie der Begriff schon andeutet, liegt der Kernaspekt dieser Verfahren in der Anpassbarkeit der Kompositionen. Sie soll sich den dargestellten Szenen anpassen. Im Grunde genommen ersetzte der Begriff adaptive Musik nur den Begriff interaktive Musik und legte damit für Spielkomponisten eine Agenda mit einem bestimmten Schwerpunkt fest.²⁴ Das Ziel adaptiver Musik besteht demnach darin, eine anpassungs- und

²³(User Barricade, http://www.youtube.com/all_comments?v=WojopwjrsnY, 24.01.2013)

²⁴(Clark, 2007)

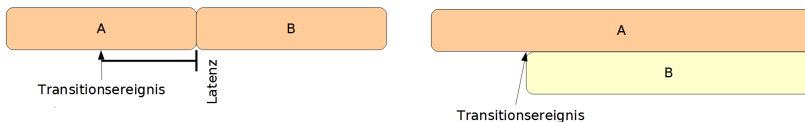


Abbildung 3.4: Transitionen zwischen horizontalen und vertikalen Bausteinen

reaktionsfähige Komposition zu realisieren, die trotz der zeitlichen Ungewißheit wie klassische Filmmusik funktioniert. Adaptive Musik etabliert eine klare Hierarchie. Die Musik hat sich nach dem dargestellten Geschehen zu richten, wie das folgende Statement von Richard Jacques verdeutlicht:

Basically, we are still composing music to picture here, even though the picture is not a locked linear picture (with the exception of some cut scenes), so it is still what is happening on screen that rules how I compose creatively.²⁵

Bei adaptiven Musikkonzeptionen enthalten die Bausteine keine kompletten Stücke mehr, sondern nur Teile davon. Die Bausteine sind nicht nur Materialien für den Aufbau der akustischen Ausgabe des Programms, sondern auch Versatzstücke einer aus ihnen zusammensetzbaren musikalischen Form. Adaptive Musik besteht aus horizontal und vertikal kombinierbaren Bausteinen und Regeln für deren Zusammensetzung. Die Übergangspunkte, an denen die Bausteine miteinander verknüpft werden, heißen *Transitionen*. Sie markieren jene Zeitpunkte, an denen eine Anpassung der Musik durch die Einsetzung eines entsprechenden Musikbausteins erfolgt. Das auslösende Ereignis, in dessen Folge sich das laufende Musikstück verändert, wird als *Transitionereignis* bezeichnet. Vertikal übereinander liegende Bausteine heißen auch *Layer*. Sie können ohne Verzögerung dazugenommen oder ausgeblendet werden zu können. In den meisten Fällen entsprechen die Layer verschiedenen Instrumentengruppen wie Perkussion und Bass. Transitionen zwischen vertikalen Bausteinen benötigen hingegen eine Latenz in der Länge der verbleibenden Dauer des Musikbausteins, in dem das Ereignis eintrat.

²⁵Richard Jacques in (Marks, 2008, 229)

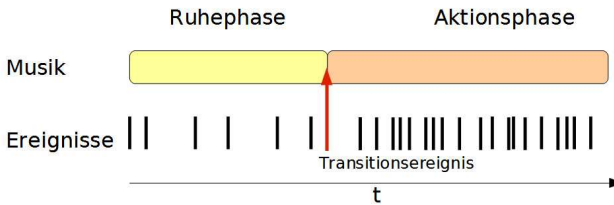


Abbildung 3.5: Musiktransition zwischen Spielphasen

Auskomponierte Transitionsbausteine, mit deren Hilfe abrupte Übergänge zweier Stücke verhindert werden sollen, benötigen also Zeit. Eine Möglichkeit um diese Zeit zu gewinnen besteht darin, den Entscheidungsspielraum für die Länge des Transitionsbausteins einzuschränken. In dem 2008 erschienenen *Prince of Persia Reboot* wurde dieses Prinzip geschickt eingesetzt, indem die Transitionen während unidirektionaler Kletterpartien ablaufen. Sobald der Prinz den Weg in einen neuen Landabschnitt einleitet, gibt es kein Zurück mehr. Für einige Sekunden muß er eine Reihe von Hindernissen bewältigen, welche der Musiktransition die nötige Zeit verschaffen. Das Leveldesign ist in diesem Fall auf die musikalische Ausgestaltung ausgerichtet.

Adaptive Musik kann ganz unterschiedliche Formen annehmen, je nachdem wie die Bausteine und die mit dem Spiel verwobenen Regeln für ihre Zusammensetzung gestaltet sind. Weit verbreitet ist beispielsweise der Einsatz adaptiver Musik in offenen Spielwelten. Die im letzten Abschnitt vorgestellte strukturbasierte Musikintegration stößt hier an ihre Grenzen. In offenen Spielwelten ist die Kopplung von Musik an Levels und Maps nicht mehr möglich. An Stelle des Übergangs in eine neue Strukturkomponente müssen andere Transitionspunkte definiert werden. Sie orientieren sich häufig an strukturellen Merkmalen der Spielwelt. So unterteilten die Entwickler von *Risen* die virtuelle Umgebung in mehrere Landabschnitte, denen jeweils unterschiedliche Stücke zugeordnet sind. Die Musik richtet sich dementsprechend nach dem Ort, an dem sich der Avatar befindet. Gleichzeitig charakterisiert sie die Fraktionen, in deren Einflußbereich dieser Ort liegt. Weitere Musiktransitionen treten auf, wenn der Tag in den Abend übergeht und die Nacht dem Morgen weicht. Das Transitionsraster der Musik ist damit sowohl an dem Raum als auch der Zeit der fiktiven Spielwelt ausgerichtet.

Eine andere Form adaptiver Musik kommt in Story-basierten First Person Shootern wie *Half Life 2* und *Mirror's Edge* zur Anwendung. Hier untergliedert sich das Spielgeschehen in verschiedene Phasen. Wenn keine Gegner anwesend sind, erhalten die Spieler den Freiraum für die Erkundung der Spielwelt. Diese Ruhephasen, in denen kaum zeitkritische Aktionen gefordert werden, sind häufig durch das Leveldesign implementiert, indem Barrieren einen offensichtlichen Weg versperren und zur Untersuchung der Spielwelt auffordern. Häufig bleibt in diesen Spielabschnitten die Musik komplett ausgeschaltet oder auf dezente Flächen reduziert. Der Übergang von diesen Ruhephasen in hektische, zeitkritische Aktionsphasen dient als Transitionspunkt der musikalischen Begleitung. Treibende Tracks suggerieren ein schnelles Handeln. Die musikalische Struktur antizipiert damit die sich verdichtende Ereignis- und Aktionsstruktur. Üblicherweise beginnt die Musik einige Sekunden früher als das hereinbrechende Ereignisgewitter. Die Spieler erhalten Zeit, um sich darauf einzustellen. Im Gegensatz zu dem an der virtuellen Welt ausgerichteten Transitionsraster in *Risen* orientiert sich die Musik bei der Vertonung verschiedener Phasen stärker an der Spielsituation. So können die Dauer und die besuchten Orte während einer ereignisreichen Spielphase variieren.

Häufig werden Transitionsergebnisse zwischen verschiedenen Phasen ohne bewusste Aktionen der Spieler ausgelöst. Es reicht aus, dass der Avatar eine unsichtbare Schwelle überschreitet oder im Hintergrund ein Timer ausläuft, um die Aktionsphase zu triggern. Andere Transitionsergebnisse sind als narrative Elemente der Erzählung des Spiels untergeordnet und an Aktionen der Spieler gebunden. In *Left 4 Dead 2* enthalten einige Level sogenannte *Crescendo-Events*, die von den Spielern zur Überwindung von Hindernissen ausgelöst werden müssen. Zu den Hindernissen gehören z. B. eine ausfahrbare Brücke und ein Karussell, das ein- und ausgeschaltet werden muß. Der damit verbundene Lärm fungiert als narrative Rechtfertigung für das Spawnen größerer Gegnerhorden, deren Bewältigung ein konzentriertes und gemeinsames Vorgehen erfordert.

Algorithmische Kompositionen

Mit Hilfe des Computers ist möglich, algorithmisch erzeugte bzw. prozedurale Musik zu generieren. Sie entsteht durch einen schrittweisen Ablauf zuvor festgelegter Regelsätze, auf deren Grundlage die Zusammensetzung musikalischen Materials geschieht. Die im vorherigen Abschnitt vorgestellten adaptiven Musikkonzeptionen sind in diesem Sinne ein Spezialfall algorithmischer Musik. Sie bestehen ebenfalls aus komponierten Bausteinen und Regeln, nach denen die Bausteine zusammengesetzt werden. Adaptive Musik ist immer auch eine algorithmische Komposition. Eine algorithmische Komposition ist aber nicht zwangsläufig adaptiv. Der Unterschied besteht darin, dass adaptive Musik stets in einem Funktionszusammenhang gedacht wird. Sie richtet sich durch ihre Anpassbarkeit per definitionem nach etwas anderem. Algorithmische Musik verfolgt hingegen nicht das Ideal der Filmmusik, bestimmte Szenen passgenau untermalen zu können. Statt dessen gibt sie einen offeneren Rahmen vor, der eine ungebundenerere Variantenbildung erlaubt.

Algorithmische Kompositionsverfahren sind keineswegs neu und prinzipiell auch ohne den Computer umsetzbar. Zu den bekanntesten Vertretern gehört das Musikalische Würfelspiel, das Ende des 18. Jahrhunderts aufkam. Die meisten Kompositionsspiele dieser Art enthielten eine Grundkomposition mit Lücken, die durch mehrere zufällig ermittelte Variationen gefüllt werden konnten. Nach dem Wurf des Würfels ließ sich die zu wählende Variation in einer Tabelle nachschlagen. Um das Ergebnis der erwürfelten Komposition begutachten zu können, war es erforderlich, die einzelnen Takte auf neuem Papier zusammen zu führen und sie anschließend am Klavier vom Blatt zu spielen. Computergestützte algorithmische Kompositionen entlasten die Benutzer, da die Maschinen sowohl das Würfeln und Zusammensetzen, als auch die Wiedergabe der fertigen Stücke erledigen. Die Ausführung in Echtzeit kaschiert den Aufwand, der zur Berechnung und Darlegung der Stücke erforderlich ist. Computerprogramme bieten damit einen neuen, deutlich unbeschwerteren Zugang zu algorithmischen Kompositionen. Während die auf Papier festgehaltenen Fassungen alle nötigen Prozesse beschreiben, sie aber nicht ausführen können, und Datenträger wie Audio-CDs nur die Ergebnisse algorithmischer Kompositionen wiedergeben können, sind Programme in der Lage, generative Kompositionssysteme in ihrer Gesamtheit abzubilden.

Abbildung 3.6: *C.P.U. Bach*

Neben avantgardistischer Musik, die sich in der zweiten Hälfte des 20. Jahrhunderts vermehrt mit algorithmischen Kompositionsverfahren auseinandersetzte, flossen prozedurale Techniken in die Musik des Computerspiels ein. Bereits 1984 veröffentlichte Lucasfilm Games mit *Ballblazer* einen Titel, für den Peter Langston eine algorithmisch generierte Melodielinie kreierte.²⁶ Langston führt die Idee seines Ansatzes, den er „riffology“ nennt, auf seine Erfahrung als improvisierender Lead-Gitarrist zurück. Ausgehend von mehreren vordefinierten Melodiefragmenten, die dem Riff-Repertoire des Gitarristen entsprechen, erzeugt der Algorithmus eine Verkettung der Riffs und variiert ihre Lautstärken sowie Geschwindigkeiten. Die Entscheidungen für den Verlauf der Melodie sind dabei unabhängig von Spielereignissen. Das Computerspiel bildet lediglich den operativen Rahmen des potentiell unendlich lang weiterlaufenden Stücks. Algorithmisch generierte Variationen sind ein Mittel, um Musik über längere Zeiträume interessant zu halten. Sie begegnen der zeitlichen Ungewisheit mit Abwechslung.

Eine stärkere Anlehnung an das Musikalische Würfelspiel zeigte ein Projekt, das Sid Meier nach der Fertigstellung von *Civilization* verfolgte. Gemeinsam mit Jeff Briggs programmierte er das Kompositionsprogramm *C.P.U. Bach* für die 1994 erschienene Konsole 3DO. Nach der Auswahl verschiedener Parameter wie z. B. der gewünschten Gattung generiert das Programm barocke Musik im Stil von Johann Sebastian Bach. Obwohl *C.P.U. Bach* kein Computerspiel ist sondern ein Kompositionsprogramm, zeigt es, wie innerhalb der Computerspielindustrie experimentelle künstlerische Konzepte erkundet wurden. Das Programm erschien schließlich auf einer Spielkonsole und richtete sich an deren Benutzer.

Neben Game Designern wie Sid Meier, die ausgehend von der Domäne des Computerspiels Exkurse in die algorithmische Komposition unter-

²⁶(Langston, 1989, 6)

nehmen, beschreiten Komponisten algorithmischer Musik den umgekehrten Weg. So wirkte Brian Eno an der akustischen Gestaltung des von Will Wright entworfenen Spiels *Spore* mit. In einem Vortrag an der Long Now Foundation erklären Eno und Wright ihre Faszination für generative Systeme.²⁷ Das System hinter *Spore* besteht demnach aus zellulären Automaten im Stil des „Spiel des Lebens“, das der Mathematiker John Conway 1970 entwarf. Ausgehend von wenigen Regeln, die durch die Analyse von Nachbarschaftsbeziehungen das Entstehen und Vergehen von Zellen in einem quadratisch segmentierten Spielfeld beschreiben, emergieren komplexe Muster. Eno beschreibt die Gestaltung der Muster mit folgender Metapher:

So this is the power of generative systems, that you make seeds rather than forests.

Nachdem die Anfangszustände (die Samen) und die Regeln des Systems gesetzt sind, entsteht das Muster (der Wald) aus dem schrittweisen Durchexerzieren der Regeln. Die Anfangskonfiguration wird auch als erste Generation bezeichnet. Der Zustand nach der n-maligen Anwendung des Regelsatzes beschreibt die n-te Generation. Zelluläre Automaten verwendete Wright bereits bei *Sim City* zur Modellierung von Kriminalität, Verkehr und Umweltverschmutzung.²⁸ Im Gegensatz zu Conways *Spiel des Lebens*, das nach dem Setzen der Anfangskonfiguration keine weitere Interaktion vorsieht, ermöglicht *Sim City* auch den Eingriff in späteren Phasen der ablaufenden Simulation.

Zelluläre Automaten können ebenso gut für algorithmische Kompositionen verwendet werden wie für die Erzeugung von Fraktalen und die Simulation von Spielaspekten. Sie bilden ein Modell, an das sowohl die Logik des Spiels, als auch die Spieldarstellung koppelbar ist. Anders als bei *Ballblazer*, das eine vom Spielgeschehen losgelöste algorithmische Komposition enthält, lassen sich Kompositionen dieser Art aus der Struktur der Spielereignisse ableiten. Zelluläre Automaten sind dabei nur ein Beispiel unter vielen. Die Symmetrie der daraus emergierenden Strukturen bietet sich an, um ästhetisch ansprechende Kompositionen zu konzipieren. Grundsätzlich können jedoch alle Spielzustände und -ereignisse als klanggenerierende und -modifizierende Auslöser herangezogen werden. Regeln für die Zusammensetzung der Musikbausteine sind bei dieser Konzeption entweder vollständig oder teilweise auch

²⁷http://fora.tv/2006/06/26/Will_Wright_and_Brian_Eno (10.09.2012)

²⁸ebd.

Regeln des Spiels. Die Musikbausteine selber sind nochmals feiner segmentiert als bei adaptiver Musik. Im Extremfall enthalten sie keinerlei auskomponierte Passagen mehr, sondern sind die atomaren Bestandteile der Komposition - Einzeltöne und ihre Parameter.

Die aus einem generativen System heraus emergierende Musik besitzt andere Charakteristiken als auskomponierte Musik, die in den musikalischen Vertonungen des Computerspiels zum Einsatz kommt. Brian Eno beschreibt letztere als *Narrative Music* und die generierte Musik als *Ambient Music*. Ambient Music besitzt im Idealfall keinen Anfang, kein Ende und keine Höhepunkte. Eno vergleicht sie mit der Geräuschkulisse eines Flusses, in die man jederzeit eintauchen oder sich ihr entziehen kann. Wenn die Musik aus den Regeln des Spiels hervorgeht, besitzt sie nichtsdestotrotz einen ludischen – wenn auch nicht narrativen – Spannungsbogen. Je nach der Dichte der in die Musikgeneration eingebundenen Ereignisse treten Höhen und Tiefen auf. Sie sind wie das An- und Abschwellen des Flusses, vom leisen Plätschern bis hin zum reißenden Strom. Eine exemplarische Analyse der akustischen Gestaltung von *Breakout* soll diese aus dem Spiel entstehende musikalische Dramaturgie verdeutlichen.

Breakout erschien 1976 als eine Variante von *Pong*. Das Ziel des Spiels besteht darin, mehrere Reihen horizontal angeordneter Blöcke mit Hilfe des Balls zu eliminieren. Sobald der Ball einen Block berührt, verschwindet der Block. Obwohl der Arcade-Automat lediglich über einen Schwarz-Weiß Bildschirm verfügte, erschienen damals die Steine durch aufgeklebte Kunststofffolien in Farbe. Die verschiedenfarbigen Reihen der Blöcke erzeugen dabei jeweils unterschiedliche Töne. Bei der ebenfalls im Jahr 1976 erschienenen Portierung für den Atari 2600 sind es sechs Reihen in den Farben eines Regenbogens. Je höher die Blöcke positioniert sind, desto höher ist der ihnen zugeordnete Ton. Legt man anstelle der Vorstellungen einer Narrativen Musik nun ein offeneres Musikkonzept zu Grund, lässt sich bereits diese simple Zuordnung von Tönen zu Kollisionsereignissen als algorithmische Komposition verstehen, die eng mit dem Spiel verschaltet ist. In jedem Spieldurchlauf variieren die Reihenfolgen und Zeitpunkte der ausgelösten Töne und ergeben eine durch die Interaktion beeinflusste musikalische Struktur. Höhepunkte treten dann auf, wenn sich die Kollisionsereignisse häufen. Eine beliebte Spielstrategie besteht darin, an den Rändern der Blockreihen eine Schneise freizulegen, so dass der Ball zwischen den oben gelegenen Blöcken und der Bildschirmbegrenzung hin- und her-

pendelt. Mit der Anzahl der rasch aufeinanderfolgenden Töne schnellert auch die Punktzahl in die Höhe und folgt so einem bis heute praktizierten Spielhallenschema, das jene hochfrequenten Tonsequenzen in Analogie zu klimpernden Münzen als gewinnbringend konnotiert. Die Schneisen-Strategie sorgt für eine Verdichtung der vertonten Ereignisse, die gleichzeitig einen Höhepunkt des Spiels und einen Höhepunkt der Musik generiert.

Algorithmische Kompositionen, die auf den Regeln des Spiels basieren und deren Bausteine die atomaren Bestandteile der Komposition bilden, sind eng mit der Tätigkeit des Spielens verbunden. Sie fungieren nicht als Begleitung wie adaptive Kompositionen, sondern sie emergieren aus dem Spiel heraus. Die Entscheidungen und das Timing der Spieler während der Laufzeit beeinflussen ihre konkrete Gestalt. Das Spiel macht die Musik. Beim rhythmusbasierten Musikspiel – einem speziellen Spieltypus, der in den vergangenen Jahren deutlich an Popularität gewann – ist es genau umgekehrt. Die Musik macht das Spiel.

Musikspiele: Ein Level für ein Lied

Spiel und Musik waren schon lange vor dem Auftreten des Computerspiels zwei eng miteinander verbundene Phänomene. Bereits im alltäglichen Sprachgebrauch finden wir Spuren, die auf ihre Nähe verweisen. Wir *spielen* Musikinstrumente und wenn uns dabei ein Fehler unterläuft, so haben wir uns *verspielt*. Die Entstehung des Musikspiels als eigenständiges Genre des Computerspiels scheint so gesehen eine sehr nahe liegende Entwicklung zu sein.

Der Durchbruch des Musikspiels in den Mainstream der Spielkultur geschah erst, nachdem das Sampling die technische Grundlage für die Integration von Aufnahmen ermöglichte. Rhythmus- und Karaoke-Spiele wie *Guitar Hero*, *Rockband* und *Singstar* sind heute so weit verbreitet, dass sie für viele Menschen eine wichtige Form der Auseinandersetzung mit Musik darstellen. Musikspiele dieser Art haben gemeinsam, dass sich die Musik während des Spielvorgangs nicht primär an dem Spiel und der Erzählung ausrichtet. Sie ist keine Begleitung, die sich der jeweiligen Spielsituation anpasst, sondern selbst der Mittelpunkt des Spiels. Die Musik erfüllt eine leitende Funktion in Hinblick auf die zeitliche Strukturierung des Spielgeschehens. Während bei adaptiven

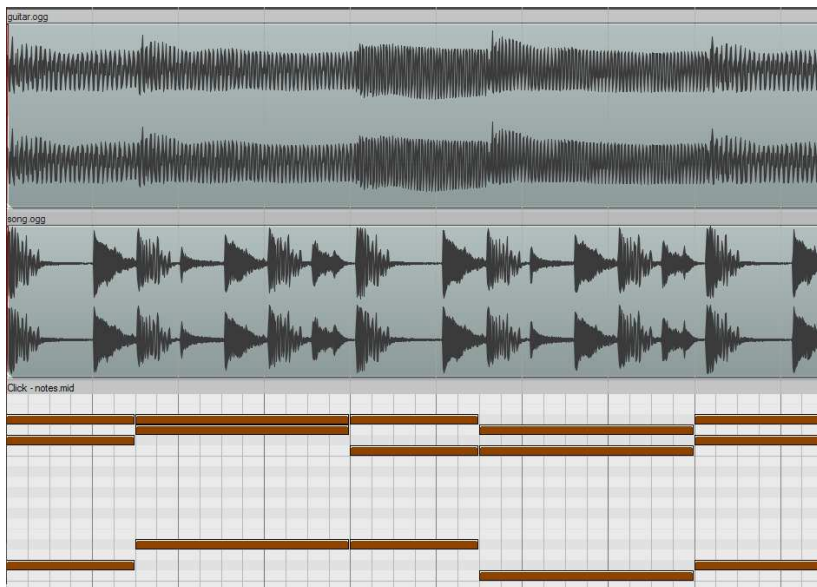


Abbildung 3.7: Dateien eines *Frets on Fire Levels*

Musikkonzeptionen die Zeitpunkte der Einsätze und Transitionen unklar sind, herrscht hier über den Ablauf große Gewißheit. Das Lied ist gegeben, was fehlt ist das Level.

Bei den oben genannten Titeln sind ausgehend von einem Stück bestimmte Eingabemuster abgeleitet, die mit einem System zur Punktevergabe gekoppelt sind. Eine richtige Eingabe gemäß dieses Musters wird mit Punkten und Beifall belohnt, während eine falsche Eingabe im Rahmen des narrativen Kontextes – eine Konzertsituation mit einem äußerst kritischen Publikum – als Verspieler kodiert ist. Die Hauptschwierigkeit besteht darin, das Timing des Musters zu verstehen. Wer den Song bereits kennt, dem fällt diese Aufgabe in der Regel ungemein leichter. Bei *Frets on Fire*, einer Open Source Variante von *Guitar Hero*, ist das Spielprinzip mit Hilfe von zwei Audio-Dateien und einer Midi-Datei als Behälter der Eingabemuster realisiert. Abbildung 3.7 zeigt jeweils den Beginn dieser drei Dateien für das Stück „Bang Bang Mystery Man“. Die mittlere Datei enthält den Background-Track des Stücks, während die obere Audio-Datei jene Spur darstellt, die die Spieler durch

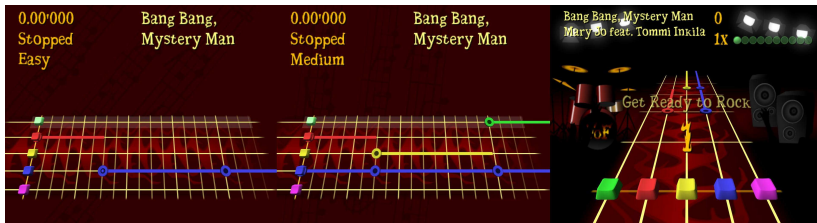


Abbildung 3.8: *Frets on Fire* Level-Editor und Spielansicht

korrekt getimte Eingaben nachspielen. Wie man an der Wellenform gut erkennen kann, handelt es sich hierbei um eine Bass-Stimme. Die unten abgebildete Midi-Datei enthält die Eingabemuster von zwei Schwierigkeitsgraden. Der obere Schwierigkeitsgrad verlangt jeweils zwei gleichzeitig gegriffene Tasten, während der untere mit jeweils einer Taste auskommt. Abbildung 3.8 zeigt beide Schwierigkeitsgrade im Leveleditor des Spiels sowie die Spielansicht des härteren Schwierigkeitsgrads. Der Quellcode verrät, dass die Lautstärke der Bass-Spur auf null gesetzt und ein Störgeräusch abgespielt wird, sobald die Spieler einen Punkt des Musters verfehlen. Beide Audio-Dateien werden über die gesamte Dauer eines Stücks parallel abgespielt, wobei die Bass-Spur in Abhängigkeit von den Eingaben lediglich leise gestellt wird oder laut bleibt.

Obwohl die Aufgabe des Spiels darin besteht, etwas hörbar zu machen, führt der Weg dazu primär über das visuell wahrnehmbare Muster. Die Spieler müssen es sehen, um es durch den Druck auf die richtigen Knöpfe zur richtigen Zeit bestätigen zu können. Ihr Ziel ist die Synchronisation der Eingaben mit dem vorgegebenen Muster. Theoretisch ist das Spiel auch mit abgeschaltetem Sound spielbar. Die Hintergrundspur hilft jedoch dabei, die rhythmische Struktur des Musters zu erfassen und die Zeitpunkte besser einzuschätzen. Nur diejenigen Spieler, die das Muster eines Songs auswendig gelernt haben, sind in der Lage, es ohne die grafische Darstellung zu reproduzieren.²⁹ Der Entwurf des nachzuspielenden Musters stammt von den Entwicklern und engagierten Mitgliedern der Community, die mit Hilfe eines beiliegen-

²⁹ Dieses Auswendiglernen des Musters ist eine Form der Spieler, Virtuosität zu demonstrieren. Weitere Spielformen von *Guitar Hero* und das daraus entstehende Spannungsfeld zwischen Performativität, Spiel und Musik erkundet die Musikethnologin Kiri Miller in (Miller, 2009).

den Leveleditors jedes beliebige Stück für das Spiel aufbereiten können. Das Muster an sich ist eine Art rhythmische Interpretation der nachzuspielenden Stimme. Musikalische Strukturelemente innerhalb des Lied-Bausteins werden über eine Metastruktur mit Spielereignissen verbunden. Kopplungen dieser Art sind neben dem Musikspiel auch vereinzelt in anderen Computerspielen anzutreffen. In *New Super Mario Brothers* für das Nintendo DS sind Akzente der Hintergrundmusik mit Aktionen computergesteuerter Akteure gekoppelt. Goombas springen bei Chöreinsätzen in die Luft, während die Schildkrötenartigen Koopa Troopas einen Tanzschritt ausführen. Während Spielereignisse üblicherweise den Ausschlag für Klangereignisse geben, liegt hier der umgekehrte Fall vor. Musikalische Elemente sind die Auslöser von Spielereignissen. Je besser die Spieler mit der Musik des Spiels vertraut sind, desto präziser können sie die Zeitpunkte dieser Abweichungen von den üblichen Laufwegen der antagonistischen Akteure einschätzen. Die Musik ermöglicht ihnen, präzise Vorhersagen zu treffen. Gleichzeitig wird das Spiel zu einem musikalischen Erlebnis, da die Spieler die Choreografie der Figuren nicht nur beobachten, sondern ihren Tanz in ihr Spiel miteinbeziehen und antizipieren. Die musikalische Struktur kann neue Ereignisse und Zustände in das Computerspiel einspeisen.

Während *Frets on Fire* dank des Leveleditors um beliebige Stücke erweiterbar ist, erleichtern einige andere Musikspiele die Integration der Musik des Benutzers noch stärker, indem der Aufbau eines Levels an den Computer delegiert wird. Ein Vertreter dieser Art ist das Spiel *Audiosurf*, in dem die Spieler einen Raumgleiter durch eine dreispurige Bahn führen, in der es farbige Blöcke einzusammeln gilt. Sowohl die Gestalt der Bahn als auch die Positionen der Blöcke und das Tempo des Raumgleiters werden algorithmisch auf Basis der akustischen Daten generiert. Die Spieler können beliebige Audiodateien in den gebräuchlichsten Formaten einspeisen, auf deren Basis das Programm einen Level erzeugt. Die Faszination des Spiels geht von dem deterministischen Algorithmus aus, der für beliebige Tonaufnahmen eine Visualisierung und ein an Regel- und Punktesystem gekoppeltes Eingabemuster erzeugt. Obwohl der Algorithmus nicht einsehbar ist, verrät sein Output, dass er Parameter wie das Volumen und die Dynamik der Audio-Datei berücksichtigt. In ruhigen Passagen eines Stücks sind wenige Blöcke untergebracht, während Spannungsspitzen um so mehr von ihnen enthalten. Bei einem Hip Hop-Track pumpt der Raumgleiter in Synchroni-

sation zu dem Beat sprunghaft nach vorne, während er bei Rhythmen mit weniger starken Betonungen gleichmäßig über die Bahn rollt.

Der Schöpfer des Spiels, Dylan Fitterer, gibt in einem Interview an, dass das Ziel der Entwicklung eine Art spielbare Musikvisualisierung war.

[..] it came from my fascination with music visualizers. I wanted to merge a visualizer with gameplay to get more invested in it and sharpen my emotional response.³⁰

Mit dem Interesse für Musikvisualisierungen betrat Fitterer ein Terrain, das in der Geschichte des Computerspiels immer wieder durchstreift wurde. Bereits 1978 veröffentlichte Atari ein Zusatzgerät unter der Kennzeichnung *Video Music C-240*, das die akustischen Signale einer angeschlossenen Stereo-Anlage in grafische Muster auf einem Fernseher wandelte.³¹ Die *Virtual Light Machine* der 1993 erschienenen Konsole Atari Jaguar 64 griff das Feature der Musikvisualisierung erneut auf³², bevor Ende der 1990er Jahre durch die Verbreitung von Audiospielern wie WinAmp die Musikvisualisierung auf Windows-PCs zum Standardfeature wurde. Mit *Rez* erschien 2001 ein Computerspiel, dessen Stil stark an typische Musikvisualisierungen anlehnt und das Fitterer ebenfalls als Inspiration angibt. *Audiosurf* ist ein Beispiel dafür, wie Computerspiele Algorithmen aus anderen Anwendungsbereichen adaptieren und als Grundlage für neue Spielsysteme annehmen. Das Prinzip, aus der Analyse akustischer Daten eine grafische Repräsentation zu gewinnen, wird insofern erweitert, als dass die Analyse gleich ein Level erzeugt. Die Möglichkeit zum Einsammeln der farbigen Blöcke macht die Visualisierungen spielbar. Der deterministische Algorithmus, der für jede Audio-Datei immer wieder die gleiche Strecke erzeugt, macht das Spiel kompetitiv. Sowohl die eigenen Spieldurchläufe lassen sich miteinander vergleichen, als auch die eigene Performance mit derjenigen von anderen Spielern auf der ganzen Welt abgleichen. Ein webbasiertes Highscore-System listet zu jeder Audio-Datei, die je gespielt wurde, die erzielten Bestzeiten auf. Neben einigen mitgelieferten Stücken und der Musiksammlung der User steht durch *Audio Surf Radio* und eine Anbindung an *Last.fm* eine ständig wechselnde Song-

³⁰(Gamasutra, 2007)

³¹(Wolf, 2003, 56)

³²http://www.audiovisualizers.com/toolshak/vidsynth/jag_vlm/jag_vlm.htm
(12.09.2012)

auswahl zur Verfügung. Durch diese webbasierten Features besitzt die Gemeinschaft der Spieler eine gemeinsame Datenbank aller bereits gespielten und gerade aktuellen Songs, über die in den angeschlossenen Foren ein reger Austausch stattfindet.

Die *Audiosurf*-Foren sind die Schauplätze, an denen ein Diskurs über Musik im Allgemeinen und das damit verbundene Gameplay im Speziellen stattfindet. Spieler geben sich untereinander Tipps, welche Musik für eine intensive Spielerfahrung geeignet ist.

The new Animal Collective album „Meriweather Post Pavilion“ is a GREAT Ride on Audiosurf. It has the perfect balance of slow/fast, uphill/downhill, calm/bouncy/frantic etc.³³

Limewax - Cracking Core. Put this song in Audiosurf. Makes it look like the Matrix, it's totally crazy!³⁴

try classical music – crazy tracks³⁵

Auch der Schöpfer des Spiels hat seine Favoriten. Als Frage auf den seltsamsten Song, den er je gespielt hat, antwortet Flitterer:

Ich komme immer wieder zu einem Song von Lisa Gerrard namens "Sacrifice" zurück. Der hat wirklich wunderschöne Bergab-Areale mit gleißend roten Tunneln und eine Menge Kurven. Die Strecke passt also sehr gut zur majestätischen, gefühlvollen Stimmung des Songs. "Meltdown" von Orbital ist eine fantastische, 10-minütige Erfahrung. Bob Dylans Harmonika macht wundervolle Tunnel. Fast alles von Daft Punk [sic] ist perfekt. Das wohl seltsamste ist es, einen Podcast herunterzsurfen.³⁶

Im Mittelpunkt der Kommentare steht weniger die Musik selbst als das, was durch die Musik generiert wird. Es fällt auf, dass persönliche musikalische Präferenzen kaum zum Ausdruck kommen. Die Möglichkeit des Programms, aus jeder beliebigen Audio-Datei einen Level zu

³³<http://www.lastfm.de/group/Audiosurf> (User kmanning2008, eingesehen am 02.07.2009)

³⁴ebd. (User fleadilla, eingesehen am 11.06.2012)

³⁵ebd. (User molingrad, eingesehen am 11.06.2012)

³⁶<http://www.pcgames.de/Audiosurf-PC-213173/none/Unser-Interview-mit-dem-Audiosurf-Schoepfer-Dylan-Fitterer-632002/> (12.09.2012)

erzeugen, weckt vielmehr die Neugierde über die Ergebnisse des Algorithmus. Zu diesem Zweck testen die Spieler mitunter auch Stücke aus Genres, mit denen sie weniger vertraut sind. Von Interesse sind entweder Audio-Dateien, auf deren Basis der Algorithmus möglichst stimmige durchfahrbare Strecken generiert, oder deren Ergebnisse außergewöhnlich und „seltsam“ ausfallen. Programme wie *Audiosurf* sind gleichzeitig Spiele, Musikvisualisierungen und Experimentierstätten. Allen voran stellen sie neue Formen der Auseinandersetzung mit Musik dar, die über das Hören hinausgehen. Sie eröffnen die Möglichkeit, bereits bekannte Musik auf eine neue Art und Weise zu erleben und neue Musik kennen zu lernen. Wenn Musik auf diese Weise spielbar wird, entstehen von Seiten der Benutzer ganz neue Anforderungen und Bewertungskriterien an die Kompositionen. Auch für *Guitar Hero* und *Frets on Fire*, deren nachspielbare Muster nicht algorithmisch erzeugt sind, gelten ähnliche Kriterien. Die ausgewählten Stücke müssen zumindest eine gewisse Plausibilität aufweisen, um mit einem interessanten Muster versehen zu werden. Hier gehören insbesondere Titel mit Gitarrensoli dazu, die den Kriterien eines abwechslungsreichen und fordernden Spiels genügen, sowie Raum für eine spektakuläre Performance lassen.

Wenn man die Metastruktur zur Verknüpfung innermusikalischer Strukturen und Spielereignisse ausblendet, entspricht die Summe der zusammengemischten Bausteine des Musikspiels einem Baustein der Level-Begleitung. Dieser für die Musik in Computerspielen keineswegs selbstverständliche Umstand ist entscheidend für die Wechselbeziehungen zwischen Musikspielen und der Musikindustrie. So sind die Stücke der *Guitar Hero* Serie auch allesamt auf Tonträgern erhältlich. Der Release von *Guitar Hero III - Legends of Rock* sorgte geradezu für eine Renaissance älterer Rock-Klassiker, die auch im Downloadhandel der Musikindustrie deutliche Umsatzsteigerungen hinterließ. Beispielsweise stieg der Absatz des ursprünglich 1974 erschienenen Titels „Same old song and dance“ von Aerosmith in zwei Monaten nach dessen Erscheinen in *Guitar Hero III* um 446%.³⁷ Als vermeintlich schwerster Song in *Guitar Hero III* verbuchte vor allem der Titel „Through the fire and flames“ der Powermetal-Band Dragon Force großen Erfolg. Chart-Platzierungen und die Integration in weitere Musikspiele wie *Guitar-Freaks*, *DrumMania V6* und *Brütal Legend* waren die Folge.

³⁷(Christian Höferle, 2008)

4 Räumliche Konstruktionsprinzipien

Die Simulation navigierbarer Räume stellt eine genuine Errungenschaft der Neuen Medien dar.¹ Sowohl in Anwendungsprogrammen und Hypertext, als auch speziell in Computerspielen gilt die Traversierung unterschiedlicher Bereiche – seien es Menüs, Textabschnitte oder repräsentierte Räume – als grundlegendes Prinzip ihrer Nutzung.

The new digital environments are characterized by their power to represent navigable space. Linear media such as books and films can portray space, either by verbal description or image, but only digital environments can present space that we can move through.²

Wenn wir virtuelle Räume betreten und durchqueren, bewegen wir uns im Falle des Computerspiels immer auch durch Klangräume. Stand bei der zeitlichen Dimension der akustischen Architektur noch die Frage im Vordergrund, wie die Zeitpunkte von Ereignissen und Aktionen in das akustische Arrangement hineinspielen, richtet sich das Augenmerk nun auf die Gestaltung von Orten. Welche Prinzipien liegen dem räumlichen Aufbau der akustischen Struktur zugrunde, und welche Spielformen werden durch diese räumlichen Arrangements möglich?

Die Untersuchung der akustischen Gestaltung virtueller Räume verspricht dabei nicht nur, Einblicke in die Mechaniken des Spiels zu geben. Die zum großen Teil durch Computerspiele herausgebildeten Verfahren zur Darstellung virtueller Räume begegnen uns heute in vielen verschiedenen Wirtschafts- und Lebensbereichen. So gehört z. B. die Architekturvisualisierung zu den gängigen Verfahren der Bauplanung. Einige dieser Visualisierer bieten ihren Benutzern die Möglichkeit, die entworfenen Gebäude wie in einem First Person Shooter zu betreten. Weitere

¹siehe (Manovich, 2001, 244-285)

²(Murray, 1997, 79)

Programme wie *Google Maps*, *Earth* und *StreetView* ermöglichen es uns, mit Hilfe ähnlicher Navigationsverfahren die Aufnahmen realer Orte zu durchschreiten. Wir zoomen in Satellitenbildern auf Länder, Städte und Hinterhöfe und klicken uns durch bekannte oder unbekanntere Straßenzüge. Die Erfassung unserer Welt in solchen Anwendungen scheint immer weiter voranzuschreiten.

Unsere akustische Umwelt blieb von den systematischen Vermessungen unserer Informationsgesellschaft bislang größtenteils verschont.³ Die oben genannten Anwendungsprogramme beschränken sich allesamt auf visuelle Informationen, während akustische Daten außen vor bleiben. Man kann sich eine Gegend zwar aus mehreren Perspektiven anschauen, aber nicht anhören. Es gibt virtuelle Entwürfe ganzer Neubausiedlungen, in denen Fassaden glitzern und großzügige freie Flächen das Auge zieren, ohne dass die akustischen Auswirkungen solcher Bauprojekte auch nur thematisiert werden.⁴

Computerspiele hingegen enthalten nicht nur grafisch aufbereitete Repräsentationen und Simulationen von Räumen, sondern auch deren akustische Gegenstücke. Sie stellen nicht nur grafische Umgebungen dar, sondern auch begehbare Klangräume. Dabei handelt es sich zwar in den meisten Fällen nicht um Nachbildungen realer Schauplätze⁵, sondern um fiktive Räume, die mit Fantasie und künstlerischem Anspruch in Hinblick auf ein ästhetisch ansprechendes Spielerlebnis entworfen werden. Wenn jedoch grafische Darstellungstechniken Eingang in ganz unterschiedliche Formen von Anwendungssoftware halten, liegt der Gedanke nahe, dass akustische Repräsentationstechniken des Spiels in Zukunft ebenfalls in anderen Kontexten zur Anwendung kommen. Der in diesem Kapitel ausgebreitete Überblick über die Verbindung akustischer Elemente und räumlicher Aspekte des Spiels soll die wichtigsten

³Nichtsdestotrotz scheint es nur eine Frage der Zeit zu sein, bis der öffentliche Raum von Mikrofonen ebenso erfasst wird wie von Überwachungskameras. Bereits heute sind in einigen US-amerikanischen Städten sogenannte „Gun Detection Systems“ im Einsatz. Sie bestehen aus einer in der Stadt verteilten Anordnung von Mikrofonen, die ununterbrochen nach Schüssen horchen und sie orten sowie bei Bedarf Polizei und Rettungskräfte automatisch alarmieren. Auch handelsübliche Überwachungskameras verfügen mittlerweile über Mikrofone, mit denen z. B. zerbrechendes Glas erkannt werden kann.

⁴Initiativen im Stile des von Raymond Murray Schafer Ende der 1960er Jahre ins Leben gerufene *World Soundscape Project*, das sich kritisch mit den uns umgebenden Soundscapes auseinandersetzt, blieben bislang aus.

⁵Eine Ausnahme ist beispielsweise *Grand Theft Auto IV*, dessen Schauplatz Liberty City dem realen New York nachempfunden ist.

Konstruktionsprinzipien virtueller Klangräume zeigen und, darauf aufbauend, eine Reihe von Spielformen in diesen Räumen erkunden.

4.1 Die Vertonung von Spielräumen

Spiel- und Klangräume des Computerspiels weisen eine ebenso große Variabilität auf wie die unterschiedlichen Zeitstrukturen des Spiels. Bevor das raumbasierte Sounddesign in den Fokus der Betrachtung rückt, sollen Überlegungen zur Beziehung zwischen realen und virtuellen Klang- und Spielräumen sowie eine Skizze der historischen Entwicklung virtueller Raumstrukturen einen ersten Überblick verschaffen.

Klang in realen und virtuellen Räumen

Ein Computerspiel findet stets an zwei Orten gleichzeitig statt. Es gibt einen realen Spielraum, in dem sich die Spieler mit ihren Ein- und Ausgabegeräten befinden, und es gibt einen über die Ausgabegeräte gezeigten virtuellen Raum, in den sich die Spieler hineinversetzen. Zu den realen Räumen gehören Spielhallen, Wohnzimmer oder U-Bahn-Stationen, während die virtuellen Räume die Orte des Spielgeschehens und der Fiktion repräsentieren.⁶ Je nach der Konzeption des Programms und den Nutzungsweisen der Spieler treten beide Räume unterschiedlich in Erscheinung. So tritt der reale Raum in Single-Player-Games in der Wahrnehmung der Spieler weit zurück, wenn sie sich ungestört in die virtuellen Räume hineinbegeben. Sie bewegen sich durch virtuelle Räume und sitzen dabei größtenteils unbeweglich vor den Ausgabegeräten. Bei lokalen Multiplayer Games wie *Rockband* steht der reale Raum hingegen im Mittelpunkt der Spielperformance. Die Bewegung findet parallel im virtuellen und realen Raum statt. Hier und dort steht jemand

⁶Realer und virtueller Spielraum lassen sich bei Computerspielen klar voneinander unterscheiden. Es gibt jedoch subtilere Formen von Software, die unsere Wahrnehmung von Räumen beeinflussen, ohne dass wir sie unmittelbar wahrnehmen. Kitchin und Dodge nennen dieses Phänomen Code/Space: „Code/space occurs when software and the spatiality of everyday life become mutually constituted, that is, produced through one another. [...] For example, a check-in area at an airport can be described as a code/space. The spatiality of the check-in area is dependent on software. If the software crashes, the area reverts from a space in which to check in to a fairly chaotic waiting room.“ (Kitchin u. Dodge, 2011, 16f)

mit seinem Instrument auf der Bühne. Weiterhin stellt die Verwendung von Eingabegeräten wie der *WiiMote*, *PS Move* oder *Kinect* Anforderungen an den realen Spielraum, in dem sich die Körperbewegungen der Spieler vollziehen. Der Wohnzimmertisch wird aus dem Weg geräumt, und die Zuschauer nehmen einen Sicherheitsabstand zu den Spielern ein.

Auch die akustischen Ausgaben stellen gegebenenfalls Anforderungen an den realen Raum. Mehrkanalige Lautsprecher-Setups wie 5.1 oder 7.1 verlangen bestimmte räumliche Anordnungen. Manche Eingabegeräte wie die *WiiMote* enthalten zusätzliche Lautsprecher, deren Positionen sich durch die Gesten der Spieler ändern. Die Einbeziehung des realen Spielraums wird durch akustische Elemente dementsprechend stärker forciert als durch grafische. Gerade der Kontrast zwischen der flachen Scheibe des Bildschirms und dem in Klang gehüllten Raum verdeutlicht, dass die Geräuschkulisse des Spiels weitaus weniger abgeschottet ist als die grafische Darstellung. Wenn die Spieler keine Kopfhörer einsetzen, vermischen sich die Klänge des Spiels mit den akustischen Gegebenheiten ihrer Umgebung. Ein Resultat dieser akustischen Kontamination ist die Wahrnehmung des Computerspielounds in benachbarten Räumen und die Vermischung von Geräuschen der realen Umgebung mit denen der virtuellen Welt. Wenn die Spieler in einem urbanen Zentrum bei offenem Fenster *Grand Theft Auto* spielen, dann akzeptieren sie zu einem gewissen Grad die Vermischung beider Soundscapes ohne sie als Störung zu empfinden. Verwechslungen zwischen den Ursprüngen der Klangquellen sind potentiell möglich. Ein polternes Geräusch aus dem Nebenzimmer kann die Spieler eines Horror-Adventures zu einer Eingabe bewegen, wenn sie erst kurze Zeit später den wahren Ursprung des Geräuschs registrieren. Es gibt Wechselwirkungen zwischen dem Klang des Spiels und der akustischen Umgebung der Spieler.⁷

Während in den meisten Fällen die akustischen Gegebenheiten der Umgebung der Spieler den Entwicklern unbekannt sind, gibt es einige Fälle, in denen sie berücksichtigt werden und in das Sounddesign einfließen. So lautet eine grundsätzliche Anforderung an Mobile Games, dass sie auch mit abgeschaltetem Sound gut spielbar sein müssen. Akustische Elemente sollten dementsprechend keine Zusatzinformatio-

⁷In geringerem Maße trifft das auch auf grafische Gestaltung zu. So beeinflussen z. B. die Lichtverhältnisse des realen Spielraums das Spielerlebnis. Vor allem Titel des Horror-Genres legen nahe, im Dunkeln zu spielen.

nen transportieren, auf deren Grundlage Entscheidungen getroffen werden müssen. Auch bei Arcade Games, die für den öffentlichen Raum bestimmt sind, berücksichtigt das Sounddesign die akustischen Grundzüge der Aufstellungsorte.

A fairly successful method is to use sounds that closely resemble typical noises of the location the arcade machine is likely to be installed to. If, for instance, the game is some kind of racing simulation, it is usually wise to make it sound like TV or radio broadcast of a racing event instead of trying to reach perfection in creating to create that special 8,000-RPM engine sound.⁸

Im Gegensatz zu den realen Spielräumen, die den Entwicklern – abgesehen von den typischen Aufstellplätzen der Plattformen – unbekannt bleiben, unterliegen die virtuellen Räume ihrer uneingeschränkten Kontrolle. Im Lauf der über 40-jährigen Geschichte des Computerspiels entstand eine Reihe unterschiedlicher simulierter Raumstrukturen. War zunächst nur die Darstellung von zweidimensionalen bildschirmgroßen Spielfeldern möglich, können sich die Spieler heute in den Weiten dreidimensionaler virtueller Welten verlieren. Zwischen diesen beiden Polen öffnet sich ein Spektrum räumlicher Darstellungsverfahren, das von horizontalen und vertikalen Sidescrollern über isometrische und Top-Down-Perspektiven bis hin zu First- und Third-Person-Perspektiven reicht. Bemerkenswert an dieser Entwicklung ist, dass neue Verfahren zur Darstellung von Räumen ihre Vorgänger nicht ausnahmslos ersetzen. Vielmehr gehen sie gemeinsam in einem Tableau realisierbarer Darstellungsmöglichkeiten auf.⁹ So erscheinen auch im Jahr 2012 neben den Spielen mit gigantischen 3D-Welten immer noch Titel, bei denen der Bildschirm das gesamte Spielfeld zeigt. Analog zu den Zeitstrukturen, die in unterschiedlichen Abschnitten von Computerspielen variieren, treten in vielen Titeln auch unterschiedliche Raumstrukturen auf. Ebenso wie z. B. die Spiele der *Total War*-Reihe eine rundenbasierte Planungs- und Verwaltungsphase mit echtzeitbasierten Schlachten kombinieren, oszilliert der Raum dieser Spiele zwischen einer zweidimensionalen Kartenansicht und der Darstellung dreidimensionaler Geländeabschnitte. Weiterhin unterstützen einige Computerspiele das Umschal-

⁸(Weske, 2000)

⁹Mark J. P. Wolf listet in (Wolf, 2002, 51-76) von textbasierten bis hin zu 3D-Umgebungen die gängigsten Repräsentationstechniken von Räumen in Computerspielen.

ten zwischen mehreren unterstützten Perspektiven. In Rollenspielen ist z. B. die Wahl zwischen einer First und einer Third Person-Perspektive üblich. Die Raumstrukturen unterschiedlicher Computerspiele weisen ebenso wie die Zeitstrukturen ein hohes Maß an Variabilität auf.

Die Entwicklungen virtueller Raumstrukturen sind primär auf Innovationen im Bereich der Computergrafik zurückzuführen. Die akustische Gestaltung passte sich den neuen Gegebenheiten jeweils im Nachhinein an. Den größten Einfluss auf die Vertonung von Räumen übten die im Zuge der Erschließung dreidimensionaler Repräsentationsverfahren entwickelten Audio-Technologien aus.

Die Erschließung der dritten Dimension

Die Einbeziehung der dritten Dimension gilt als einschneidenste Entwicklung des Computerspiels, in deren Folge sowohl die grafischen und akustischen Ausgaben, als auch die Eingabeformen massive Änderungen erfuhren. Heute ist es kaum noch vorstellbar, welche Sensation die Einführung der dritten Dimension in Spielwelten bedeutete und welche Anforderungen sie an die Spieler stellte. David Kushner ruft in seinem Buch *Masters of Doom* einige Symptome in Erinnerung, die der First Person Shooter *Wolfenstein* bei den Spielern hervorgerufen haben soll. Obwohl das Spiel noch keine Kopfbewegungen nach unten und oben integrierte, ist in vielen Berichten die Rede von Übelkeitserscheinungen.

In fact, the sense of immersion was so real that many began complaining of motion sickness. Calls were coming in even at the Apogee office saying that people were throwing up while playing the game. *Wolfenstein* vomit stories became items of fascination online. Theories abounded. Some players thought the game's animation was so smooth that it tricked the brain into thinking it was moving in a real space. Other gamers thought it had something to do with the "jerkiness" of the graphics, which induced the feeling of seasickness. Some felt it was simply disorienting because there was no acceleration involved; it was like going from zero to sixty at light speed. Gamers even exchanged tips for how to play without losing one's Doritos.¹⁰

¹⁰(Kushner, 2004, 94)

Die Steuerung des Avatars in einem First Person Shooter verlangte eine gehörige Portion Übung und Gewöhnung. Ließen sich auf dem PC die meisten Spiele entweder komplett mit Tastatur oder Maus bedienen, wurde nun die gleichzeitige Nutzung beider Eingabegeräte Pflicht. Mit der Maus bewegen die Spieler die am Kopf des Avatars befestigte Kamera, während sie mit der Tastatur die Figur innerhalb des Raumes bewegen.

Die grafische Erweiterung des Raums in die dritte Dimension legte auch für die akustische Gestaltung eine räumliche Erweiterung nahe. Nachdem zunächst Stereo-Schwenks zum Einsatz kamen, entwickelten sich bald elaboriertere Lösungen. In ihrem Mittelpunkt stehen die sog. Head Relative Transfer Functions (HRTF). Sie bezeichnen ein psychoakustisches Verfahren, das unserem Hörsinn die Position eines Klangs suggeriert. Die Ausgangsidee bestand darin, mit Hilfe spezieller Algorithmen die Wiedergabe von dreidimensionalem Sound auf zwei Lautsprechern zu ermöglichen:

After all, humans only have two ears, and if the algorithms used to signal the brain about sound location can be reverse-engineered, then two speakers should be able to reproduce those algorithms and essentially fool the brain into thinking that a sound is originating from somewhere other than the speakers. These algorithms are generally known as HRTF¹¹

Die Umsetzung gestaltete sich als komplexes mathematisch-physikalisches Problem, das ohne die Messung der individuellen Kopf- und Ohrformen der Hörer nur approximativ gelöst werden kann.¹²

HRTF-Algorithmen stellen das Fundament von Programmbibliotheken wie *DirectSound*¹³ und *Open AL*¹⁴ dar, mit deren Hilfe sich dreidimensionale Klangräume gestalten lassen. Die Bibliotheken stellen Funktionen zur Platzierung von Klangquellen innerhalb des Raums und Steue-

¹¹(Boer, 2002, 25)

¹²In Hinblick auf neue Eingabegeräte ist es zumindest theoretisch denkbar, dass zukünftige HRTF-Algorithmen auch die sensorisch erfasste, individuelle Konstitution der Ohren berücksichtigen.

¹³*DirectSound* ist ein seit 1995 kontinuierlich weiterentwickelter Bestandteil der von Microsoft entwickelten Programmierschnittstelle *DirectX*. Der 2008 eingeführte Nachfolger von *DirectSound* heißt *XAudio 2*.

¹⁴Die *Open Audio Library* entstand im Jahr 2000. Ihr Aufbau orientiert sich an der seit 1992, ursprünglich von Silicon Graphics entwickelten Bibliothek *Open Graphics Library (Open GL)*.

rung ihres Verhaltens zur Verfügung. So können sich auch während des Spielverlaufs die Eigenschaften der Klangquellen verändern und z. B. ihre Positionen verschieben. Darüber hinaus simulieren die Bibliotheken akustische Phänomene wie den Doppler-Effekt.

Am Beispiel von *Open AL* soll der grundlegende Aufbau dieser Bibliotheken kurz skizziert werden. *Open AL* unterstützt nicht nur PC-Spiele, sondern auch Titel gängiger Konsolen wie der *Xbox 360*, *Playstation 3* und *Wii*. Von *Jedi Knight 2* über *Doom 3* bis hin zu *Bioshock* und *America's Army* bildet die Bibliothek seit über einer Dekade das Grundgerüst für die akustische Gestaltung vieler Computerspiele. Die an Programmierer gerichtete Dokumentation fasst die Inhalte der Bibliothek folgendermaßen zusammen:

OpenAL is a cross-platform three-dimensional audio API. The API's primary purpose is to allow an application to position audio sources in a three-dimensional space around a listener, producing reasonable spatialization of the sources for the audio system (headphones, 2.1 speaker output, 5.1 speaker output, etc.) [...] Use of OpenAL revolves around the use of three fundamental objects – Buffers, Sources, and a Listener. A buffer can be filled with audio data, and can then be attached to a source. The source can then be positioned and played. How the source is heard is determined by its position and orientation relative to the Listener object (there is only one Listener). Creating a number of sources and buffers and a single listener and then updating the positions and orientations of the sources and listener dynamically can present a convincing 3D audio world.¹⁵

Während ein Buffer nur wenige editierbare Eigenschaften besitzt, wie z. B. die Frequenz, Bittiefe und Anzahl der Kanäle, besitzt eine Klangquelle über 20 änderbare Zustände. Abbildung 4.1 zeigt die in (Hiebert u. a., 2007, 13) gelistete Beschreibung der Eigenschaften. Neben Attributen zur Kontrolle des Abspielvorgangs (z. B. `AL_LOOPING` und `AL_SOURCE_STATE`) befinden sich darunter vor allem Attribute zur Positionierung und Verhaltenssteuerung der Klangquelle. So lassen sich u. a. ihre Koordinaten, ihr Emissionskegel sowie ihre Richtung und Geschwindigkeit beeinflussen.

¹⁵(Hiebert u. a., 2007, 7-8)

Property	Data Type	Description
AL_PTCH	f, fv	pitch multiplier
AL_GAIN	f, fv	source gain
AL_MAX_DISTANCE	f, fv, i, iv	used with Inverse Clamped Distance Model to set the distance where there will no longer be any attenuation of the source
AL_ROLLOFF_FACTOR	f, fv, i, iv	the rolloff rate for the source
AL_REFERENCE_DISTANCE	f, fv, i, iv	the distance under which the volume for the source would normally drop by half
AL_MIN_GAIN	f, fv	minimum gain for this source
AL_MAX_GAIN	f, fv	maximum gain for this source
AL_CONE_OUTER_GAIN	f, fv	gain when outside the oriented cone
AL_CONE_INNER_ANGLE	f, fv, i, iv	gain when inside the oriented cone
AL_CONE_OUTER_ANGLE	f, fv, i, iv	outer angle of the sound cone
AL_POSITION	fv, 3f	X, Y, Z position
AL_VELOCITY	fv, 3f	velocity vector
AL_DIRECTION	fv, 3f, iv, 3i	direction vector
AL_SOURCE_RELATIVE	i, iv	determines if the positions are relative to the listener
AL_SOURCE_TYPE	i, iv	source type - AL_UNDETERMINED, AL_STATIC, or AL_STREAMING
AL_LOOPING	i, iv	turns looping on (AL_TRUE) or off (AL_FALSE)
AL_BUFFER	i, iv	id of the attached buffer
AL_SOURCE_STATE	i, iv	state of the source (AL_STOPPED, AL_PLAYING, ...)
AL_BUFFERS_QUEUED	i, iv	number of buffers queued on this source
AL_BUFFERS_PROCESSED	i, iv	number of buffers in the queue that have been processed
AL_SEC_OFFSET	f, fv, i, iv	playback position, expressed in seconds
AL_SAMPLE_OFFSET	f, fv, i, iv	playback position, expressed in samples
AL_BYTE_OFFSET	f, fv, i, iv	playback position, expressed in bytes

Abbildung 4.1: Eigenschaften einer *Open AL* Klangquelle

Das Listener-Objekt repräsentiert schließlich den Hörer. Es wird direkt an die Position und Ausrichtung des Avatars gekoppelt. Bewegen die Spieler mit der Maus die am Kopf des Avatars befestigte Kamera oder ändern mit der Tastatur die Position, ergibt sich jeweils eine andere Hörperspektive. Ein Listener-Objekt besitzt lediglich vier Eigenschaften. Das Attribut `AL_POSITION` repräsentiert analog zu den Klangquellen die Position des Hörers innerhalb des dreidimensionalen kartesischen Koordinatensystems. `AL_ORIENTATION` enthält einen Vektor mit der aktuellen Blickrichtung und einen dazu orthogonalen Vektor. Darüber hinaus enthält das Listener-Objekt lediglich einen Gain-Regler und einen Geschwindigkeitsvektor.

Die Gestaltung begehrter Klangräume durch Systeme wie *OpenAL* resultiert in neuen räumlichen Wahrnehmungsmöglichkeiten der fiktiven Welt. In traditionellen audiovisuellen Medien basiert die Zuordnung von Geräuschen zu ihren Ursprüngen innerhalb des Bildes auf einem psychoakustischen Phänomen, das Michel Chion im Kontext des Films als „spatial magnetization of sound by image“¹⁶ bezeichnet. Sehen die Zuschauer auf dem Bildschirm eine Explosion und hören dabei aus den Lautsprechern ein Explosionsgeräusch, dann verknüpfen sie diese beiden Sinneseindrücke. Obwohl die Emission des Klangs von den Lautsprechern und nicht der Bildschirmposition ausgeht, an welcher die Explosion stattfindet, verorten sie den Ursprung von Beidem an der betroffenen Position ohne deren physikalische Trennung zu hinterfragen. HRTF-Algorithmen erweitern dieses psychoakustische Prinzip. Sie berechnen in Echtzeit, wie die virtuelle Welt aus unterschiedlichen eingenommenen Hörperspektiven klingt. Alle positionierten Klangquellen besitzen nun tatsächlich einen Ort innerhalb des modellierten Raums. In zweidimensionalen Spielewelten mit Mono-Ausgabe galt hingegen noch das gleiche Prinzip wie beim Film. Das gleichzeitige Abspielen einer Animation und eines Geräusches fiel in unserer Wahrnehmung am Ort des Dargestellten zusammen.

Der entscheidende Unterschied zwischen zwei- und dreidimensionalen Soundscapes besteht darin, dass dreidimensionale Geräusche direkter mit dem Modell der Spielwelt verschaltet sind. Sie fungieren nicht nur als Zeichen, sondern verweisen darüber hinaus auf Zahlen, allen voran auf ihre Koordinaten. Sie werden durch ihre Positionierung und Ausrichtung innerhalb des Raummodells erfahrbar. Eine Konsequenz

¹⁶(Chion u. a., 1994, 69-71)

daraus besteht in der Verschiebung zwischen zustandsvarianten und -invarianten Darstellungselementen.¹⁷ Im Vergleich zu zweidimensionalen Räumen ändert sich in der akustischen Dimension das Verhältnis von Vorder- und Hintergrund. Obwohl es nach wie vor noch Umgebungsgeräusche gibt, die keine ortbare Position besitzen, treten viele der zuvor räumlich unbestimmbaren Geräusche nun als Klangquellen mit ausmachbaren Koordinaten in Erscheinung. Geräusche, die in zweidimensionalen Spielwelten im Hintergrund lagen, werden Stützen für die räumliche Orientierung. Die ehemalige Dekoration erfüllt nun wichtige Funktionen innerhalb des Spielablaufs. Das vertonte Flackern von Neonleuchten und das Rattern von Generatoren unterstützt die Spieler bei der Orientierung innerhalb des virtuellen Raums. Während Soundscapes in zweidimensionalen Spielewelten als atmosphärische Hintergründe fungieren, übernehmen sie in dreidimensionalen Umgebungen zusätzliche Orientierungsfunktionen. Sie helfen den Spielern dabei, sich in den virtuellen Räumen zurechtzufinden.

Mark Grimshaw berücksichtigt diese Funktion des 3D-Sounds in seiner Arbeit über die akustische Ökologie¹⁸ des First Person Shooters, indem er die für den Film konzipierte Theorie unterschiedlicher Hörmodi um den Modus des *navigational listening* erweitert.¹⁹ Chion fasst in (Chion u. a., 1994, 25-34) drei verschiedene Hörmodi zusammen: Kausales Hören, Semantisches Hören und Reduziertes Hören. Kausales Hören fragt nach den Ursprüngen des Gehörten und schließt dadurch auf die Eigenschaften der Klangquellen. Semantisches Hören bezeichnet die Wahrnehmung von Sprache und allen weiteren akustischen Kommunikationssystemen, die Botschaften transportieren. Der von Pierre Schaeffer konzipierte Modus des Reduzierten Hörens beschreibt eine auf den Klang selber gerichtete Aufmerksamkeit, die sowohl dessen Ursprünge,

¹⁷Verschiebungen sind auch auf grafischer Ebene vorhanden. Während der visuelle Hintergrund z. B. in 2D-Plattformern einen sehr prominenten, exponierten Platz einnimmt, verschwindet das Konzept der Hintergrundgrafik in 3D-Titeln größtenteils. Der Himmel oder das entfernte Landschaftspanorama sind häufig nur in seltenen Momenten einsehbar und verlangen als unerreichbare Orte nicht mehr jene perzeptuelle Differenzierungsarbeit, die für die Unterscheidung von Vorder- und Hintergrund in zweidimensionalen Darstellungen notwendig war. An die Stelle der Hintergrundgrafik treten nun Oberflächentexturen, die den architektonischen Raum bepflastern und geometrische Details simulieren.

¹⁸Engl. „acoustic ecology“. Der Begriff bezeichnet die Erforschung von Soundscapes. Als Ausgangspunkt der Disziplin gilt das von R. Murray Schafer Ende der 1960er Jahre initiierte *World Soundscape Project*.

¹⁹(Grimshaw, 2007, 99)

als auch dessen Bedeutungen ignoriert. Das Konzept des *navigational listening* entwirft Grimshaw nun als eine über das Kausale Hören hinausgehende Form der Frage nach dem Ursprung von Klangquellen. Es dient dem Zweck, die Spieler bei der Wegfindung zu unterstützen. Im Gegensatz zu den drei anderen Hörmodi impliziert es Aktionen von Seiten der Hörenden.

Umgebungsgeräusche

In Folge der Entwicklung dreidimensionaler Räume begann die Integration von Umgebungsgeräuschen als einer bis dato stark vernachlässigten Klangkategorie. Bis Anfang der 1990er Jahre waren Umgebungsgeräusche der fiktiven Spielwelten in Computerspielen kaum vorhanden. Die durch Hintergrundgrafiken suggerierten Räume und Landschaften blieben größtenteils still. Kontinuierliche Hintergrundmusik und Soundeffekte, die an zentrale Aktionen und Ereignisse gebunden waren, bestimmten den Ton. Vereinzelte Ambient-Sounds, wie das mit dem Rauschgenerator des SID-Chips erzeugte Windgeräusch in *Giana Sisters*, bildeten die Ausnahme. In dieser Phase kompensierte die Musik durch ihre Prominenz im Mix in weiten Teilen die Abwesenheit von zustandsinvarianten Umgebungsgeräuschen. Viele Kompositionen orientierten sich an der Level-Gestaltung im Zuge des in 3.4 auf Seite 122 skizzierten Untermalungsansatzes. Betrat ein Spieler z. B. in *Super Mario Brothers* ein Level der Unterwelt, änderte sich nicht nur die Hintergrundgrafik, sondern auch das musikalische Thema. Auch heute richten sich viele Kompositionen für Computerspiele vorrangig nach den Orten, an denen sie abgespielt werden.

Ab Mitte der 1990er Jahre erfolgte die Erschließung und kontinuierliche Erweiterung von Umgebungsgeräuschen. Seitdem sind im Computerspiel vermehrt Geräusche hörbar, die weder an Aktionen noch an Ereignisse des Spiels gebunden sind. Sie charakterisieren Orte und zielen darauf, die virtuelle Welt glaubhafter erscheinen zu lassen. Im Kontext der Computerspielvertonung wird diese spezielle Klangkategorie häufig als *Soundscape*, und einzelne Elemente daraus als *Ambient Sounds* bezeichnet. Maßgeblichen Anteil an dieser Entwicklung hatte zum Einen die Durchsetzung des Samplings, mit dessen Hilfe aufgenommene Umgebungsgeräusche einfach in das Spiel integriert werden konnten. Zum Anderen legte die Entwicklung möglichst glaubwürdiger

dreidimensionaler Umgebungen die akustische Simulation des virtuellen Raums nahe.

In dreidimensionalen Spielwelten, insbesondere im Rahmen des First Person Shooters, verdrängte die akustische Simulation der Räume größtenteils das Konzept der kontinuierlichen Hintergrundmusik. Die Übergänge zwischen beiden Kategorien sind jedoch prinzipiell fließend. Vor allem dann, wenn die Komposition stark an der Spielwelt ausgerichtet ist und stärker mit Geräuschen operiert als mit traditionellen Musikstrukturen, verschwimmen die Grenzen. Eine Komposition dieser Art realisierte Martin Stig Anderson in dem Spiel *Limbo*. Die Musik entsteht hier aus einer Reihe verfremdeter Geräusche, deren Ursprünge auf bestimmte Aspekte der Spielwelt zurückzuführen sind.

[.] in audiovisual work you can make seamless transitions between realism and abstraction, and make sound travel smoothly between the diegetic and non-diegetic space of a represented world. For me it has a much bigger psychological impact when you turn a naturalistic soundscape into abstraction by making your sound effects play as “music” rather than adding some traditional background music.²⁰

Während in den Kompositionen der 1980er Jahre die Hintergrundmusik die dargestellten Räume charakterisierte, bildet diese Konzeption der Musik auf Basis der Umgebungsgeräusche zu einem gewissen Grad den entgegengesetzten Pol. Die Musik emergiert aus der abstrakt gewordenen Geräuschkulisse.

In Ergänzung zu Umgebungsgeräuschen, die innerhalb des virtuellen Raums platziert sind, adaptierte das Computerspiel ab Ende der 1990er Jahre eine weitere Technik zur Charakterisierung von Räumen. Mit Hilfe der *Aureal 3rd Dimension Technology* (A3D) und den *Environmental Audio Extensions* (EAX) von Creative Labs ließen sich die Hall-Eigenschaften der simulierten Räume definieren. Technisch gesehen funktioniert der Umgebungshall dabei ähnlich wie das zuschaltbare Hall-Plugin einer *Digital Audio Workstation* (DAW). In Abhängigkeit der eingestellten Parameter werden gesampelte Audio-Daten mit zusätzlichen Effekten versehen. Für unterschiedliche Räume lassen sich mit diesen Systemen unterschiedliche Hall-Eigenschaften festlegen. In einer Höhle erhalten Geräusche eine andere Qualität als in einem Bü-

²⁰(Kastbauer, 2011)

rokomplex. Der Umgebungshall betrifft allen voran die von Avataren und computergesteuerten Akteuren ausgelösten Sounds. Fußschritte, Schüsse, Sprung- und Atemgeräusche sind in Computerspielen an vielen Orten zu hören. Vor der Einbindung von Hall-Effekten blieben sie an allen Orten der Spielwelt identisch. Nun durchlaufen sie Transformationen in Echtzeit, die das akustische Verhalten der betretenen Räume widerspiegeln. Die Einbindung von Umgebungshall schafft nicht nur glaubhaftere Räume, sondern stellt eine hohe Bandbreite an Varianten dieser repetitiv auftauchenden Klänge bereit. Ebenso wie Klanggruppen, die für ein Ereignis mehrere Vertonungen beinhalten, kaschieren sie die Statik aufgenommenener Samples.

EAX wurde seit 1998 kontinuierlich weiterentwickelt. In der neuesten Version EAX Advanced HD 5.0 können bis zu 128 Kanäle gleichzeitig verarbeitet werden, während es bei Version 1.0 lediglich acht waren. Neben Verfeinerungen des Umgebungsklangs durch Techniken wie *Environment Occlusion*²¹ berücksichtigt die Version 5.0 auch die spezielle Kommunikationssituation von Multiplayer-Games. Die Stimme der Spieler ist mit den gleichen Effekten belegbar wie die Geräusche innerhalb der virtuellen Umgebung.

EAX® Voice allows you to literally become part of the game! Using a microphone connected to an EAX ADVANCED HD 5.0 compliant audio device in any game title, you can speak and hear your voice with the same effects as the environment your character is in.²²

Auch wenn dieses Feature bisher nicht weit verbreitet ist, soll es in Multiplayer-Games in Zukunft möglich sein, die Avatare der Spieler durch ihre Stimmen innerhalb der virtuellen Welt orten zu können.²³ Die akustische Simulation des virtuellen Raums geht damit noch einen Schritt weiter. Erfasste der Umgebungshall zunächst nur die Geräusche des Avatars, transformiert er nun auch die über das Mikrofon erfassten Laute der Spieler und platziert sie in der virtuellen Welt an den Koordinaten des Avatars.

²¹Environment Occlusion simuliert, wie Objekte aus festen Materialien wie Holz und Stein den durch sie passierenden Schall beeinflussen.

²²http://www.creative.com/soundblaster/technology/welcome_flash.asp?j1=eax (10.10.2012)

²³Ob sich die Spieler darauf einlassen, oder aufgrund der spielerischen Nachteile auf externe VOIP-Lösungen zurückgreifen werden, sei dahingestellt.

4.2 Raumbasiertes Sounddesign

Ereignisbasiertes Sounddesign koppelt akustische Elemente und Parameter an Aktionen. Raumbasiertes Sounddesign koppelt sie an Orte. Dabei kann es sich einerseits um Strukturabschnitte des Programms wie das Hauptmenü, einen Level und den Game Over-Bildschirm oder andererseits um Areale und Positionen der fiktiven Welt handeln. Erstere Verknüpfungen sind in der Regel direkt innerhalb des Quellcodes gesetzt, während letztere in aktuellen Entwicklungswerkzeugen über Leveleditoren bestimmbar sind. Sehr häufig sind den Gegenständen, welche die Leveldesigner an Orten der virtuellen Welt platzieren, bereits Geräusche zugewiesen. Darüber hinaus unterstützen Level-Editoren das Positionieren von Klangquellen, die nicht mit Gegenständen verknüpft sind. Eine ausführliche Darstellung der Praktiken und Strategien des raumbasierten Sounddesigns sprengt den Rahmen dieser Arbeit. Stattdessen sollen zwei Beispiele einen kursorischen Überblick über die Spannweite der räumlichen Vertonungskonzepte geben.

Das erste Beispiel untersucht die an Räume gekoppelte akustische Ausgestaltung des rundenbasierten Strategiespiels *Civilization 5*. Das zweite Beispiel zeigt, wie innerhalb des First Person Shooters *Left 4 Dead* eine Soundscape realisiert wurde.²⁴

Ambience in Civilization V

Das rundenbasierte Strategiespiel *Civilization V* besteht aus mehreren verschiedenen Ansichten, die jeweils unterschiedliche Aspekte des Spiels darstellen. Im Mittelpunkt des Spielgeschehens liegt eine isometrisch dargestellte Weltkarte. Darüber hinaus gibt es weitere Ansichten und Menüs wie den Diplomatie- und Städtebildschirm. Die Umgebungsgereusche richten sich jeweils stark danach, was die Spieler gerade sehen. Im Dialog mit den anderen Anführern, in der Städteansicht und in der Kartenansicht kommen dabei jeweils unterschiedliche Vertonungsstrategien zur Anwendung.

Der Diplomatiebildschirm zeigt den jeweiligen Anführer einer Nation vor einem im Laufe des Spiels unverändert bleibendem Hintergrund.

²⁴In beiden Fällen begünstigen offen zugängliche Skript-Dateien, Level-Editoren und Anleitungen zur Anfertigung von Modifikationen die Analyse.

```

- <Script2DSound>
  <ScriptID>AS2D_AMBIENCE_CITY_ERA_RENAISSANCE_LARGE</ScriptID>
  <SoundID>SND_AMBIENCE_CITY_MED_LARGE</SoundID>
  <SoundType>GAME_AMBIENCE</SoundType>
  <bLooping>True</bLooping>
  <bDontTriggerDuplicates>True</bDontTriggerDuplicates>
  <iDontPlayMoreThan>1</iDontPlayMoreThan>
  <iMaxVolume>60</iMaxVolume>
  <iMinVolume>60</iMinVolume>
</Script2DSound>
- <Script2DSound>
  <ScriptID>AS2D_AMBIENCE_CITY_ERA_RENAISSANCE_SMALL</ScriptID>
  <SoundID>SND_AMBIENCE_CITY_MED_SMALL</SoundID>
  <SoundType>GAME_AMBIENCE</SoundType>
  <bLooping>True</bLooping>
  <bDontTriggerDuplicates>True</bDontTriggerDuplicates>
  <iDontPlayMoreThan>1</iDontPlayMoreThan>
  <iMaxVolume>60</iMaxVolume>
  <iMinVolume>60</iMinVolume>
</Script2DSound>

```

Abbildung 4.2: Soundscape-Skript aus *Civilization V*

Neben dem Start eines Musikstücks, das der jeweiligen Nation zugeordnet ist, erklingt eine Geräuschkulisse, die sich an dem Hintergrundbild orientiert. Der aztekische Herrscher Montezuma posiert vor einer Feuerschale, deren Knistern mit den Geräuschen eines Schlachtgetümmels begleitet wird, und Mahatma Ghandi steht vor dem rauschenden Meer. Als klare Erkennungsmerkmale der Herrscher bleibt diese grafische und akustische Darstellung über alle Strecken des Spiels identisch.

Anders sieht es bei den Städtebildschirmen aus. Hier entscheiden sowohl die Größe der Stadt als auch die Epoche, in der sich die Spieler gerade befinden, welche Soundscape jeweils abgespielt wird. Abbildung 4.2 zeigt einen Ausschnitt des Skripts²⁵ für die Städte-Soundscapes im Zeitalter der Renaissance. Die Script2DSound-Tags definieren dabei die Soundscape-Eigenschaften für große, respektive kleine Städte des Zeitalters. Das Tag SoundID verweist jeweils auf bereits fertig abgemischte Audio-Dateien, die entsprechend der Stadtgröße ein unterschiedliches Maß an Geschäftigkeit suggerieren. Beide Soundscapes werden kontinuierlich wiederholt, sofern sich die Spieler in den Städtebildschirmen befinden.

Den höchsten Aufwand investierten die Entwickler in die Vertonung des in Hexagonen gerasterten Kartenbildschirms, der den größten Raum des Spiels einnimmt. Audio-Designer Michael Curran erklärt das angewandte Konzept mit folgenden Worten:

²⁵CIV5DIR/Assets/Sound/XML/AudioSoundscapeScripts.xml

With the ambience we wanted to try to represent the whole picture of what you see in the world. If the player can see a forest, a desert, an ocean, and a river on screen, we want to hear all of these elements as well. In *Civilization IV*, we just had an ambient soundscape of the tile that you were centered on. In *Civilization V*, each quadrant of the screen calls a soundscape in addition to the center tile, rivers, and resources. It's really a much more believable experience for the player.²⁶

Abhängig von den sichtbaren Geländetypen und Ressourcen ergibt sich daraus eine flexible Geräuschkulisse, die selten repetitiv wirkt. Durch die Regeln für ihre Zusammenstellung geht sie aus jeder beliebigen Karte in immer ähnlichen, aber leicht abgewandelten Varianten hervor.²⁷ Im Gegensatz zu den Städte-Soundscapes ist die Geräuschkulisse der Karte nicht an die Epochen gebunden, welche die Spieler durchlaufen. Die Welt klingt immer gleich, egal ob man sich im Mittelalter, in der Renaissance oder der Moderne befindet.

Jede dieser drei Ansichten - Diplomatiebildschirm, Städteansicht und Weltkarte - besitzt ein eigenes Vertonungskonzept. Der Diplomatiebildschirm enthält fixierte Loops, die Städteansicht hingegen variable Loops in Abhängigkeit der simulierten Einwohnerzahl und Epoche. Die Weltkarte bietet schließlich eine emergierende Klangkulisse, die aus den zugeordneten Sounds des hexagonalen Rasters und den von den Spielern gewählten Bildschirmausschnitten hervorgeht. *Civilization V* zeigt dabei, mit welchem Aufwand die Räume und Orte eines Computerspiels vertont werden, das in der Tradition des Brettspiels steht. Das Spielbrett besteht nicht mehr nur aus abstrakten grafischen Symbolen, sondern kombiniert aufwändige grafische Darstellungen im Stil einer Satellitenaufnahme mit ausgefeilten Klangkulissen.

²⁶(Greening)

²⁷Der einzige Kritikpunkt aus der Spielergemeinde betrifft die Vertonung des Viehbestands, der aus einigen häufig auftretenden Ressourcen-Feldern hervorgeht. Das Blöken und Muhen von Schafen und Kühen nimmt einen derart prominenten Platz in der Geräuschkulisse ein, dass in den Foren von www.civfanatics.com ein Thread beschreibt, wie es durch die Modifikation einer XML-Datei abgeschaltet werden kann.

Soundscapes in Left 4 Dead

Ein First Person Shooter wie *Left 4 Dead* weist im Gegensatz zu einem Strategiespiel wie *Civilization V* wiederum eine andere Raumstruktur auf, die andere Vertonungskonzepte mit sich bringt. Abbildung 4.3 zeigt die Platzierung einer Soundscape aus einem *Left 4 Dead*-Level innerhalb des *Valve Hammer*-Leveleditors. Das im Bild ausgewählte Lautsprecher-Icon repräsentiert ein Objekt namens „env_soundscape“, das den Wechsel einer Soundscape veranlasst. Das Entwicklerwiki von Valves Source Engine beschreibt das Objekt folgendermaßen:

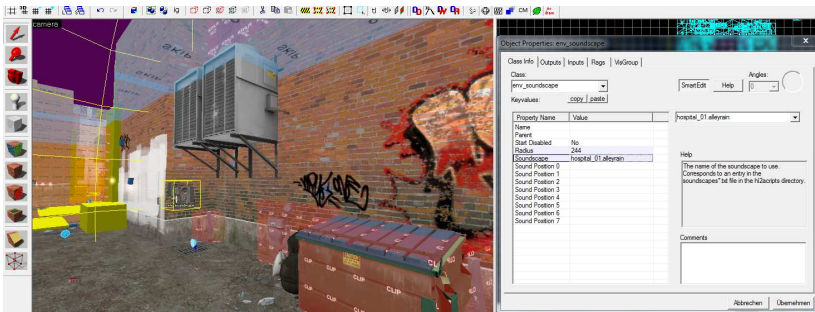


Abbildung 4.3: Soundscape „alleyrain“ in *Left 4 Dead*

env_soundscape is a point entity available in all Source games. It changes the active soundscape of players, and defines where positioned sounds within it play from.

A player must be within the entity’s radius and be able to see it for the change to take place. The entity’s soundscape will then play until another one takes over. If more than one soundscape entity is fighting for control, the nearest wins.²⁸

Das in der linken Hälfte des Bildes erkennbare Gitter lässt die umspannende Fläche des Objekts erahnen, die sich aus dem im rechten Eigenschaften-Fenster entnehmbaren Radius „244“ ergibt. Sobald die Spieler in dem Kapitel „The Apartments“ der „No Mercy“-Kampagne aus einem Haus auf die gezeigte Gasse treten, startet das Programm einen Crossfade zu der Soundscape „hospital_01.alleyrain“.

²⁸https://developer.valvesoftware.com/wiki/Env_soundscape (15.10.2012)

Das angegebene Skript²⁹ ist allerdings lediglich eine Weiterleitung auf das wesentlich umfangreichere Skript „urban.alleyrain“.

```
"hospital_01.alleyrain"
{
    "dsp" "1"
    "playsoundscape"
    {
        "name" "urban.alleyrain"
        "volume" "1"
    }
}
```

„urban.alleyrain“ definiert mehrere Bestandteile der Soundscape.³⁰ Die von den Entwicklern enthaltenen, jeweils mit den Zeichen // eingeleiteten, Kommentare und die Namen der Wave-Dateien geben einen guten Überblick über den Aufbau der Soundscape. Sie ist aufgeteilt in mehrere Blöcke, welche die Abspielmethoden „playlooping“, „playrandom“ und „playsoundscape“ bezeichnen. Innerhalb der Blöcke werden einer Reihe von Parametern konkrete Werte zugeordnet.

```
//
// Generic alley ambience
// "urban.alleyrain"
{
    "dsp" "1"

    // underlying area "tone"
    "playlooping"
    {
        "volume" "1"
        "pitch" "100"
        "wave" "ambient/weather/
                crucial_rumble_rain_nowind.wav"
    }

    // wind gusts
    "playrandom"
    {
        "time" "11,24"
        "volume" "0.8"
        "pitch" "90,110"
        "rndwave"
        {
            "wave" "ambient/wind/wind_snippet1.wav"
            "wave" "ambient/wind/wind_snippet2.wav"
            "wave" "ambient/wind/wind_snippet3.wav"
        }
    }

    // creatures
    "playrandom"
```

²⁹/left4dead/scripts/soundscapes_hospital.txt: 179-188

³⁰/left4dead/scripts/soundscapes_urban.txt: 1779-1906

```

{
    "time"      "5,20"
    "volume"    ".6,.8"
    "pitch"     "90,100"
    "rndwave"   {
        "wave"  "ambient/creatures/flies1.wav"
        "wave"  "ambient/creatures/flies2.wav"
        "wave"  "ambient/creatures/flies3.wav"
        "wave"  "ambient/creatures/rats1.wav"
        "wave"  "ambient/creatures/rats2.wav"
        "wave"  "ambient/creatures/rats3.wav"
    }
}

// animals
"playrandom"
{
    "position"  "0"
    "time"      "20,40"
    "pitch"     "95,102"
    "volume"    "0.6"
    "soundlevel" "SNDLVL_NORM"
    "rndwave"   {
        "wave"  "ambient/animal/crow_1.wav"
        "wave"  "ambient/animal/crow_2.wav"
    }
}

// metal sfx
"playrandom"
{
    "position"  "random"
    "time"      "14,23"
    "pitch"     "75,100"
    "volume"    "0.6,.9"
    "rndwave"   {
        "wave"  "ambient/Random_Amb_SFX/
                Cable_Rattle01.wav"
        "wave"  "ambient/Random_Amb_SFX/
                Cable_Rattle02.wav"
        "wave"  "ambient/Random_Amb_SFX/
                Cable_Rattle03.wav"
        "wave"  "ambient/Random_Amb_SFX/
                Cable_Rattle04.wav"
        "wave"  "ambient/Random_Amb_SFX/
                Cable_Rattle05.wav"
    }
}

"playsoundscape"
{
    "name"  "urban.util_distantcombat"
    "volume" "1"
}

"playsoundscape"
{

```

```

        "name" "urban.util_distantexplosions"
        "volume" "1"
    }

    // Position 0: Flies
    "playsoundscape"
    {
        "positionoverride" "0"
        "name" "urban.flies"
    }

    "playlooping"
    {
        "volume" ".95"
        "pitch" "110"
        "wave" "ambient/Spacial_Loops/
                MediumFire_Loop.wav"
        "origin" "2657.182861,1841.603394,61.281250;"
        "soundlevel" "SNDLVL_70dB"
    }

    // Fire by gas station hospital03 alleyrain
    "playlooping"
    {
        "volume" "1"
        "pitch" "110"
        "wave" "ambient/Spacial_Loops/
                MediumFire_Loop.wav"
        "origin" "11818.375977,6007.173828,65.281250;"
        "soundlevel" "SNDLVL_70dB"
    }

    // Fire by gas station lift
    "playlooping"
    {
        "volume" "1"
        "pitch" "110"
        "wave" "ambient/Spacial_Loops/MediumFire_Loop.wav"
        "origin" "12675.755859,6965.870117,61.281250;"
        "soundlevel" "SNDLVL_70dB"
    }
}

```

Bereits dieser kurze Ausschnitt einer vertonten Gasse lässt vermuten, wie aufwändig die akustische Umgebungssimulation eines gesamten Levels ausfällt. Neben den Abspielsegmenten „playlooping“ und „playrandom“, die direkt auf Audio-Dateien zugreifen, enthält die Soundscape Verweise auf weitere Soundscapes wie z. B. „urban.util_distantcombat“. Das referenzierte Klangmaterial an sich unterscheidet sich kaum von der Vertonung einer filmischen Szene. Es gibt Regen- und Windgeräusche, metallische Klangeffekte zur Etablierung des urbanen Settings sowie das Summen von Fliegen und das Fiepsen von Ratten. Die Organisation des Klangmaterials unterscheidet sich jedoch signifikant von der Vertonung einer Filmszene. Schließlich ist die Dauer, innerhalb derer

die Spieler im Dunstkreis der Soundscape bleiben, unbekannt. Die Soundscape enthält deswegen viele Wiederholungen und Zufallsprozesse. Alle metallischen Soundeffekte erhalten z. B. zufällige Positionen im direkten Umkreis des Avatars zugewiesen. Darüber hinaus kennzeichnen alle doppelt angegebenen Parameter wie „pitch“ „75, 100“, dass der erste Wert die untere Grenze und der zweite Wert die Obergrenze einer Zufallszahl markiert. Auf diese Weise variieren nicht nur die abgespielten Wave-Dateien und die ihnen zugewiesenen Positionen innerhalb des virtuellen Raums, sondern auch ihre Tonhöhen und Lautstärken.

4.3 Navigation und Positionierung

Bestimmte Konstellationen im virtuellen Raum können als Auslöser von Ereignissen fungieren. Eine besondere Rolle nimmt in dieser Hinsicht die Position des Avatars ein. Überschreitet die Spielfigur unsichtbare Schwellen, so werden Skript-Sequenzen abgespielt und Übergänge zwischen Soundscapes realisiert. Neben dem Aufbau von Klangräumen müssen deshalb auch die Navigationsweisen der Spieler durch diese Klangräume berücksichtigt werden. Hier lassen sich, ausgehend von der Spiel- und Spielraum-Struktur, grundlegende Navigationsprinzipien ausmachen.

Sichtbares hören und Hörbares sehen

In den meisten Computerspielen ist die fiktive Spielwelt größer als der auf dem Bildschirm dargestellte Bereich. Die Spieler sehen und hören jeweils nur Ausschnitte daraus, die sie über die gegebenen Navigationsweisen und Möglichkeiten der Kameraführung bestimmen. Abgesehen von wenigen Ausnahmen wie Rail-Shootern, die die Spieler wie auf Schienen durch ein Level lotsen, entscheiden sie selbst, in welche Richtung sie gehen, welchen Raum sie als nächstes betreten, welchen Ausschnitt des Spielfelds sie durch einen Klick auf die Minimap anspringen und in welches Teilsegment einer Karte sie hineinzoomen. Das Erlernen eines Computerspiels umfasst immer auch das Erlernen der Kameraführung und Navigation durch den virtuellen Raum. In Strategiespielen besteht das Standardrepertoire zur Bestimmung des Bildschirmausschnitts z. B. aus drei Operationen. Über das Scrolling verschieben die

Spieler den Bildschirmausschnitt, indem sie mit dem Mauszeiger einen der Bildschirmränder berühren. Das Zoomen ermöglicht die Vergrößerung oder Verkleinerung des sichtbaren Ausschnitts.³¹ Die dritte Operation bildet der Sprung an andere Orte der Karte mit Hilfe der Minimap. Einige Echtzeitstrategiespiele wie *Paraworld* erlauben darüber hinaus die Rotation der Karte und die Fixierung der Kamera auf eine Einheit, die damit verfolgt werden kann. Andere Genres besitzen wiederum andere Konventionen bezüglich der Navigation und Kameraführung. Steuern die Spieler einen Avatar aus der dritten Person, wird die Spielfigur häufig in der Mitte des Bildschirms fixiert, während sich die Welt um sie herum bei Bewegungen verschiebt. Beim First Person Shooter legen die mit der Maus gesteuerte Kamera und die Position des Avatars den sichtbaren Bildschirmausschnitt fest.

Bei all diesen Verfahren zur Bestimmung des sichtbaren Bildschirmausschnitts verändert sich potentiell auch der hörbare Ausschnitt. Computerspiele besitzen Hörperspektiven, die von den Spielern durch die Navigation und Kameraführung beeinflusst werden. Welche Faktoren wirken nun auf die Entscheidungen, bestimmte Ausschnitte auszuwählen?

Einen zentralen Faktor bildet das Spiel als regelbasiertes System. Die Spieler bewegen sich nicht nur *durch* Räume, sondern operieren auch mit Konfigurationen und Konstellationen *in* Räumen. Gerade in kompetitiven Spielsituationen ist dieser Punkt nicht zu unterschätzen. In Strategiespielen ist die Herrschaft über Raumabschnitte absolut entscheidend. So gelten beim Schachspiel die mittleren Punkte des Felds als wichtige Stützpunkte, da die Einheiten von diesen Zentren aus den größten Einfluss ausüben. Der geübte Schachspieler wird zudem darauf bedacht sein, Linien unter seine Kontrolle zu bringen. Bestimmte Positionen des Raums besitzen im Rahmen des Spielsystems einen höheren Wert als andere. Das ist nicht nur bei Strategiespielen der Fall, sondern z. B. auch beim First Person Shooter. Erfahrene Spieler kennen die Punkte einer Map, die ihnen Vorteile verschaffen. Sie wissen genau so gut, welche Areale möglichst zu meiden sind.

Das Computerspiel stellt also nicht nur navigierbare Räume dar, sondern es legt durch seine Regelsätze bestimmte Arten der Navigation durch sie und der Positionierung in ihnen nahe. Spielregeln struktu-

³¹ Titel wie *Sins of a Solar Empire* erlauben den stufenlosen Zoom von der Ansicht einer Raumschiff flotte bis hin zur Ansicht ganzer Sonnensysteme.

rieren zu einem bedeutenden Anteil die Bewegungen der Spieler durch die virtuellen Räume. Ein typisches Beispiel hierfür ist der "Nebel des Krieges", der in Strategiespielen jenen Teil des Spielfelds bedeckt, der außerhalb des Sichtradius der eigenen Einheiten liegt. Bereits das Spiel *Adventure* für den Atari 2600 simulierte die beschränkte Sichtweite der mit einer Fackel ausgerüsteten Spielfigur im Dunkeln. Eine wichtige Aufgabe der Spieler besteht in diesem Fall darin, möglichst viele Bereiche der Karte sichtbar und damit in der Regel auch hörbar zu machen. Das Spielkonzept, den Spielern über die Ausblendung bestimmter Areale nur unvollständige Informationen zukommen zu lassen, hält sich bis heute. Vor allem in Echtzeitstrategiespielen gehört es zum Standard und macht die Aufklärung zu einem wichtigen Bestandteil des Spiels. Das Aufdecken von Unbekanntem ist ein weit verbreitetes Spielkonzept, das u. a. auch in Kartenspielen eine zentrale Rolle einnimmt.

Ein weiteres Prinzip, das die Navigation durch den virtuellen Raum strukturiert, besteht aus Wechselbeziehungen zwischen dem Sicht- und dem Hörbaren. Sichtperspektive und Hörperspektive sind in den seltensten Fällen deckungsgleich. Objekte der virtuellen Welt, die in weiter Ferne liegen, können die Spieler womöglich sehen, aber nicht hören. Dinge, die sich außerhalb des Bildschirmausschnitts befinden können sie entsprechend hören, aber nicht unmittelbar sehen. Als Konsequenz daraus treten in der Navigation und Kameraführung engmaschige Wechselbeziehungen zwischen akustischen und grafischen Elementen auf. Sichtbares führt zu Hörbarem und Hörbares führt zu Sichtbarem. Nähern sich die Spieler einem entfernten Objekt, so hören sie ab einer bestimmten Schwelle dessen emittiertes Geräusch. Hören sie etwas, das außerhalb des sichtbaren Bereichs liegt, können sie durch das Verändern des Bildschirmausschnitts die Klangquelle sichtbar machen. Dieses Zusammenspiel zwischen den akustischen und grafischen Repräsentationen von Objekten und Spielvorgängen konstituiert einen weiteren Wegweiser, der die Positionierung und Navigation in virtuellen Räumen beeinflusst.

In vielen Spielen reicht der hörbare Ausschnitt weiter als der sichtbare Ausschnitt. Je kleiner die Fenster der grafischen Darstellung ausfallen, desto stärker entfaltet sich die suggestive Wirkung des Hörbaren, das sichtbar gemacht werden kann. Bereits in der Kartenansicht vieler Strategiespiele wie *Starcraft II* hören die Spieler Geräusche, die nicht sichtbar sind, sich dem gewählten Bildschirmausschnitt aber unmittelbar anschließen. Der Hörradius ist größer als das rechteckige Fenster des

Bildschirms. Noch gravierender tritt diese auditive Weitsicht in dreidimensionalen Umgebungen zu Tage. Der sichtbare Ausschnitt erfasst zu jedem Zeitpunkt nur einen Bruchteil der virtuellen Umgebung. Das, was sich neben, hinter, über, und unter dem Avatar abspielt, bleibt dem Auge verborgen.

The viewer only has a constricted window into this game world – auditory clues are critical in helping to establish a real-world proximity to objects and the surrounding environment. Because of this, sound is arguably a much more important component in modern games than it was in those games of just a few years ago.³²

Neben den Informationen über Bereiche, die sich unmittelbar dem Bildschirmausschnitt anschließen, verweisen akustische Elemente potentiell auch auf weiter entfernte Orte. In Strategiespielen berichten Statusmeldungen z. B. über Geschehnisse, die sich am anderen Ende der Karte ereignen und der Aufmerksamkeit der Spieler bedürfen. Wenn sich der Avatar eines Action-RPGs über sein volles Inventar beschwert, ist dies immer auch eine Aufforderung dazu, das Inventarmenü aufzurufen und Abhilfe zu schaffen.

Zusätzlich zu diesen Verweisen auf Entferntes, das sich dem ausgewählten Bildschirmausschnitt entzieht, zeigt die akustische Gestaltung auch Dinge aus der Nähe. Akzentuierte Sounds, die aus dem Mix hervortreten, können die Aufmerksamkeit der Spieler auf bestimmte Aspekte der virtuellen Umgebung lenken. Sie kompensieren damit zu einem gewissen Grad die Einschränkungen einer gleichbleibenden Perspektive in spielbaren Sequenzen. Während in einem Film das Tippen auf einer Schreibmaschine mit einem Schnitt und einer Nahaufnahme der Tastatur inszeniert werden kann, bietet in einem First Person Shooter die Hervorhebung des Tippgeräuschs ein ähnliches erzählerisches Mittel.

Akusmatischer Sound

Die Möglichkeit der Wahl von Sicht- und Hörperspektiven hat zur Folge, dass viele intentional einsetzbare Gestaltungsprinzipien des Filmsounds nicht ohne Weiteres auf das Computerspiel übertragbar sind.

³²(Boer, 2002, 157)

Die Differenz zwischen Onscreen und Offscreen Sound löst sich im Spiel zwischen Feedback und Suggestion auf. Insbesondere bedarf das Akusmatische, das in der Theorie audiovisueller Medien Klänge bezeichnet, deren Ursprünge im Bild nicht sichtbar, bzw. nicht eindeutig identifizierbar sind³³, im Kontext des Computerspiels einer Rekonzeptualisierung. Sobald die Spieler über Sicht- und Hörperspektiven frei verfügen können, ist eine (De-)Akusmatisierung als intentionales stilistisches Mittel schwerer einsetzbar. Die Spieler entscheiden selber, welche Klangquellen sie in den Blick nehmen. Geleitet werden sie dabei maßgeblich durch das Arrangement aus akustischem Feedback und suggestivem Sound. Eine intentionale Akusmatisierung ist im Computerspiel nur durch die Einschränkung des Navigationsspielraums oder die Unzugänglichkeit der Klangquelle realisierbar.

In dem Spiel *Portal*, das die künstliche Intelligenz Glados mit Hilfe einer akusmatischen Stimme inszeniert, bleibt Glados z. B. in allen Levels außer dem letzten hörbar, aber unzugänglich und unsichtbar. Alle Levels definieren einen klar abgesteckten Raumabschnitt, in dem sich die Spieler bewegen können. All das, was außerhalb dieses Abschnitts liegt, bleibt unerreichbar. Die Stimme von Glados ist zudem nicht an räumliche Parameter gekoppelt. Unabhängig davon, welche Hörperspektive die Spieler eingenommen haben, erschallt die Stimme aus dem Center-Speaker. Sie ist zwar da, aber gleichzeitig der Spielwelt entrückt. Zu beachten ist weiterhin, dass akusmatische Stimmen dieser Art in Computerspielen häufig dazu eingesetzt werden, um das Spiel zu erklären. Erst im Laufe des Spielgeschehens erfahren die Spieler, dass Glados tatsächlich innerhalb der fiktiven Welt des Spiels existiert.

Die vollständige Abwesenheit der Klangquelle ist mit Ausnahme von Ambient-Sounds mitunter narrativ schwer zu rechtfertigen. Aus diesem Grund treten intentionale Akusmatisierungen vorrangig im Horror-Genre auf, in dem sie als Illusionen der Spielfigur interpretierbar sind. Die Eigenheiten von akusmatischem Sound im Interaktionsgeflecht des Computerspiels zeigen, dass sich Sound-Theorien aus dem Film nicht ohne Weiteres auf das Spiel übertragen lassen. So ist z. B. auch die Übernahme von Musikkategorien problematisch.

Bei der Untersuchung von Filmmusik gehört die Unterscheidung, ob die Musik in der Handlung verortet ist oder darüber schwebt, zu den grundlegendsten Einordnungskriterien. Für das Computerspiel ist diese

³³(Chion u. a., 1994, 71)

Aufteilung in intra- und extradiegetische Musik hingegen hochproblematisch. Einerseits greift die Unterscheidung nur bei Computerspielen, die eine stark ausgeprägte Erzählung aufweisen. Ohne eine halbwegs komplexe Diegese ist sie trivial. In Musikspielen wie *Audiosurf*, in denen ein Level auf Basis der Musik erzeugt wird, ergibt sie keinen Sinn. Die Musik ist weder außerhalb noch innerhalb einer Erzählung verortet, sondern sie konstituiert das Spiel. Andererseits verfehlt diese Beschreibung den wichtigsten Aspekt der Musikimplementierung. Wenn in *Grand Theft Auto* die Musik über das in der Spielwelt verortete Radio abgespielt wird, ist deren Bezeichnung im Sinne der Filmterminologie als intradiegetisch wenig aussagekräftig. Viel wichtiger ist die Tatsache, dass die Spieler sie durch die Wahl der Radiosender beeinflussen können und diese Wahlmöglichkeit ggf. für eigene Inszenierungen ihres fiktiven Ganovendaseins einsetzen. Im Gegensatz zu der Frage nach dem Ort der Musik rückt die Frage nach deren Manipulierbarkeit durch die Aktionen der Spieler in den Vordergrund.

4.4 Spiel- und Erzählformen in Klangräumen

Innerhalb der vergangenen Jahre sind vermehrt Computerspiele erschienen, welche die Navigation in Klangräumen in ihre Spielkonzepte integrieren. Die Möglichkeiten des Horchens und darauf aufbauenden Interaktionsstrukturen werden durch die Einbeziehung von Klangräumen in Spielmechaniken und das damit verbundene Level-Design ganz unterschiedlich ausgelotet. Einige dieser Ausprägungen, von der kontrastiven Verwendung navigationsdienlicher und -hindernder Geräusche bis hin zum Schleichspiel, werden im Folgenden vorgestellt.

Gruseln: Überladene Klangkulissen und falsche Fährten

In dreidimensionalen Spielwelten wie sie der First Person Shooter bereitstellt, ist der sichtbare Ausschnitt auf das Fenster des Bildschirms beschränkt. Die Spieler sehen nur das, was sich in dem Sichtkegel ihrer gewählten Perspektive befindet. Der hörbare Ausschnitt hingegen ist kugelförmig und damit bedeutend größer als der Sichtkegel. Spieler

orientieren sich demnach nicht nur an dem, was sie sehen, sondern auch an den hörbaren Eigenschaften der dargestellten Szenerie. Genau diese Orientierungsfunktion ist der Ansatzpunkt einer Gestaltungsstrategie, die häufig zur Erzeugung von Horror-Erfahrungen eingesetzt wird. Die grundlegende Idee besteht darin, den Spielern diese auditive Orientierungsmöglichkeit zu verwehren oder sie damit in die Irre zu führen.

Überladung

Eine Variante dieser Strategie sieht eine dichte und laute Gestaltung der Ambient-Kulisse vor, um das Orten von Gegnern zu erschweren. Ein Musterbeispiel dafür findet sich in *Doom 3* in dem Level Mars City. Nachdem das Tor zur Hölle geöffnet wurde und die Marsstation im Chaos versinkt, sind die Spieler in der Rolle eines Soldaten auf sich allein gestellt. Die schwache Beleuchtung des Levels legt den Spielern nahe, ihren Hörsinn zur Orientierung zu nutzen, zumal die Taschenlampe nicht gleichzeitig mit einer Waffe getragen werden kann. Das Stampfen und Scheppern von Maschinen, das periodische Zischen von austretendem Dampf, zerstückelte Funksprüche und das Brummen defekter Lampen und Geräte macht dies jedoch kaum möglich.

Abbildung 4.4 zeigt die durch einen Konsolenbefehl sichtbar gemachten Klangquellen eines Maschinenraums in *Doom 3*. Die Pfade verweisen direkt auf die abspielenden .ogg-Dateien, während die transparenten Würfel die Positionen kennzeichnen. Die Klangquelle links im Bild (loop_airflow_02.ogg) enthält ein 18 Sekunden langes Rauschen im mittleren Frequenzbereich, das zwischendurch um unregelmäßig fließende Töne ergänzt wird. Der mittig platzierte Sound „chirp_02.ogg“ besteht aus einem tieffrequenten Rumoren mit starkem Hall. Ebenso wie diese beiden Klänge wird auch der rechts im Bild zu sehende Sound „loop_machineroom_06.ogg“, ein 5 Sekunden langes Ventilatorrauschen, kontinuierlich wiederholt. Hört man die Sounds zunächst außerhalb des Spiels und bewegt sich danach mit sichtbar gemachten Klangquellen durch den Raum, lässt sich gut nachvollziehen, wie ihre Lautstärken und Mischverhältnisse in Abhängigkeit von der Spielerposition variieren. Ohne die Kenntnisse dieses Aufbaus sind einzelne Klangquellen der Spielwelt hingegen nur äußerst schwer einzeln auszumachen, da sich die verwendeten Geräusche ähneln und ihre simultane Wiedergabe einen dichten Klangteppich erzeugt.



Abbildung 4.4: Im Raum positionierte Klangquellen aus *Doom 3*

Vor dem Hintergrund dieser Geräuschkulisse ist es sehr schwer, Geräusche im Vordergrund zu erkennen. Sobald gegnerische Figuren auftauchen, verschmelzen ihre Geräusche unmittelbar mit dem akustischen Gefüge der Kulisse, eine auditive Ortung ihrer Position bleibt größtenteils verwehrt. Die Überladung des akustischen Outputs durch den Lärm der Ambient-Kulisse torpediert sämtliche Bemühungen der Spieler hinsichtlich eines differenzierten Hörens. In Kombination mit immer wieder eingeflochtenen Skriptsequenzen, die auch im Rücken der Spieler Feinde auftauchen lassen, ergibt sich für sie dadurch eine permanente Unsicherheit.

Falsche Fährten

Eine weitere Variante dieser Gestaltungsstrategie tritt in dem 3rd Person Action-Adventure *Dead Space* auf. Ebenso wie *Doom 3* erzeugt *Dead Space* immer wieder Spielpassagen, in denen laute Maschinen-

geräusche die präzise Ortung gegnerischer Figuren einschränken. Dazu gesellen sich orchestrale, vor allem von Streichern getragene musikalische Akzente. Im Kontrast zu diesen sehr voluminösen und verdichteten Klangkollagen stehen akustisch sehr reduzierte, fast stille Momente des Spielgeschehens. In diesen Erkundungsabschnitten erhalten die Spieler wieder die Chance, die im simulierten Raum positionierten Klangquellen als Orientierungshilfe zu nutzen. Nicht selten entlädt sich die während des Horchens aufgeladene Spannung in kurzen Skriptsequenzen, die sehr laute, unerwartet eintretende Klangereignisse mit sich bringen. Plötzlich platzt im Rücken der Spieler ein Rohr oder vor ihnen springen Bildschirme an, die in voller Lautstärke zerstückelte Videos abspielen.

Darüber hinaus kommt in diesen ruhigen Passagen ein subtilerer, rein akustischer Effekt zur Anwendung. In die Soundscape des Raumschiffs *Ishimura* sind immer wieder dezent heraustretende Schepper- und Schlurf-Geräusche integriert, welche auf die Präsenz eines Anderen verweisen. Die Suche nach der Herkunft dieser beunruhigenden Klänge verläuft meistens, aber nicht immer, im Nichts. Das mit ähnlichen Klängen begleitete Auftauchen der deformierten *Necromorph* lässt sich dadurch auditiv äußerst schwer vorhersagen. Die Spieler bleiben durch diese Irritationen kontinuierlich in einem Feld der Unsicherheit, in dem Geräusche des Hinter- und Vordergrunds miteinander verschwimmen. Zustandsvariante und zustandsinvariante akustische Elemente sind kaum voneinander zu unterscheiden, so sehr sich die Spieler auch bemühen.

Auch wenn Computerspiele wie *Doom 3* und *Dead Space* offensichtlich Gestaltungsstrategien des Films adaptieren, in dem z. B. der kontrastreiche Einsatz von Stille und plötzlich einsetzendem Lärm ein probates Mittel des Horror-Genres bildet, entfalten diese Strategien im Computerspiel eine andere Intensität. Schließlich nutzen die Spieler die Soundscape der virtuellen Umgebung, um sich in ihr zurecht zu finden. Sie lauschen auf Hinweise über Ereignisse und Zustände des Spielgeschehens, auf deren Basis sie ihre Entscheidungen treffen. Anstatt sich in unheimlichen Spielsituationen die Ohren zuhalten zu können, müssen sie gerade genauer hinhören, damit ihre Spielfiguren keinen Schaden erleiden.

Horchen und Finden

Auch dann, wenn die Geräuschkulisse nicht die Mehrdeutigkeiten der oben beschriebenen Vertonungsstrategien aufweist, treten Situationen auf, in denen die Ursprünge von Klangquellen gesucht werden. Eine davon ist die Vertonung von Bonusgegenständen. Bonusgegenstände sind Objekte innerhalb der Spielwelt, die die Spieler üblicherweise mit zusätzlichen Punkten oder Ressourcen belohnen. Sie unterscheiden sich von anderen Spielgegenständen dadurch, dass sie häufig versteckt sind oder sich an schwer erreichbaren Orten befinden. Die Spieler müssen sich bemühen, um in ihren Besitz zu gelangen. Gerade in weiträumigen und verwinkelten dreidimensionalen Spielwelten gleicht die Suche nach einem versteckten Bonusgegenstand derjenigen nach der Nadel im Heuhaufen. Um etwaigen Frustrationen der Spieler vorzubeugen, nutzen Game Designer eine auf den ersten Blick paradoxe Gestaltungsstrategie: Bonusgegenstände werden versteckt und gleichzeitig hervorgehoben, sie glitzern und emittieren Klang. In *Stranglehold* sind es Kraniche, in der *Assassin's Creed*-Reihe klingende Kisten und Federn.

Vor allem die akustischen Eigenschaften der Bonusgegenstände unterstützen die Spieler dabei, sie aufzuspüren. Sie sind bei der Suche nicht auf den Sichtkegel des grafischen Ausschnitts angewiesen, sondern profitieren von dem raumgreifenderen Hörradius. Hören die Spieler bei ihren Streifzügen durch die virtuelle Umgebung das Geräusch eines Bonusgegenstands, den sie nicht unmittelbar sehen, können sie ihn hörend ertasten. Mit den Bewegungen ihres Avatars nähern sie sich dem Bonusgegenstand oder entfernen sich von ihm. Das Geräusch wird dementsprechend lauter oder leiser, bis es bei einer zu großen Entfernung vollständig verstummt. Verschiebungen der Blickrichtung und der damit verbundenen Hörperspektive ermöglichen eine Einschätzung der Richtung, aus der das Geräusch erschallt. Mit Hilfe dieser zwei Operationen lässt sich die Klangquelle auch dann orten, wenn sie hinter Wänden verborgen ist. Die experimentelle Verschiebung der Hörperspektive unterstützt die Spieler dabei, die Koordinaten der Klangquelle auszumachen und den damit verbundenen Gegenstand zu finden.

Ein ähnliches Prinzip, bei dem die Spieler kontinuierlich Hörperspektiven verschieben, um verdeckte Gegenstände zu Tage zu fördern, nutzt das Spiel *Psychonauts*. Unter Zuhilfenahme einer Wünschelrute, die einen anhaltenden, vibrierenden Ton absondert, machen sich die Spieler

auf die Suche nach vergrabenen Pfeilspitzen. Nähern sie sich ihnen, steigt die Frequenz des Tons, während sie sinkt, wenn sie sich von ihnen entfernen. Auch hier geht die Bewegung des Avatars im virtuellen Raum in ein Spiel mit Klang über. In vereinfachter Form tritt diese Suche nach Bodenschätzen durch die Wahrnehmung veränderter akustischer Parameter auch außerhalb dreidimensionaler Umgebungen auf. In *Mass Effect II* tasten die Spieler Planeten nach Ressourcen ab, indem sie mit der Maus über deren Oberflächen fahren und dem sonifizierten Ausschlag des Scanners lauschen.

Die Sonifikation der Nähe eines Gegenstands zu den Koordinaten des Avatars wird zuweilen auch eingesetzt, um die Spieler zu warnen. In *Fallout 3* und *Batman Arkham City* sind an manchen Stellen der Spielwelt Minen platziert, die sich visuell nur schwer ausmachen lassen. Sobald sich der Avatar ihnen nähert, beginnen sie zu ticken. Die Spieler erkennen dadurch die drohende Gefahr und können entsprechend reagieren, indem sie die Mine entschärfen oder den Abstand zu ihr wieder vergrößern. Im Gegensatz zu den Bonusgegenständen ist der Radius, in dem das Geräusch hörbar wird, deutlich verkürzt.

Sonntagsfahrten

Niko Bellic macht Urlaub. Anstatt Drogendealer zu überfahren, mit einer irischen Verbrecherbande Banken zu überfallen oder Juwelen aus Mülltransportern zu fischen, gönnt er sich eine Auszeit. Gekleidet in seinen Sonntagsanzug steigt er in den Sportwagen vor dem Apartement und beginnt eine Fahrt ins Blaue. Ohne Hektik und Ziel lässt er sich durch den Strom des Verkehrs treiben. Als er merkt, dass er sich direkt am Meer befindet, fährt er mit seinem Wagen auf die Fußgängerpromenade, steigt aus und schlendert zur nächsten Parkbank.

Was für Niko Bellic, den Protagonisten von *Grand Theft Auto IV* ein freier Tag ist, ist für die ihn steuernden Spieler ein freies Spiel. Wenn die Spieler ihren Avatar nun auf die Parkbank manövriert haben und den Sonnenuntergang sowie vorbei schlendernde Passanten betrachten, tritt das Spiel in eine Phase, die Alexander Galloway den „ambience act“ nennt.

The world of the game exists as a purely aesthetic object in the ambience act. It can be looked at; it is detached from the world, a self-contained expression.³⁴

Die Simulation der fiktiven Welt läuft auch ohne die Eingaben der Spieler weiter in jenen Momenten, die weder Pausen noch Progressionen des Spiels sind. Soundkulisse und grafische Gestaltung tragen zu gleichen Teilen dazu bei, die Welt lebendig erscheinen zu lassen. Neben diesen Momenten der bewegten Stasis ist es gerade die Bewegung durch den Raum, die größere Freiräume eröffnet. Der „ambience act“ bildet damit einen Gegenentwurf zu Spielmodi, die auf dem Erkennen zustandsvarianter Darstellungselemente und den damit verbundenen Prozessen aufbauen. Er relativiert funktional orientierte Interpretationen des Spielablaufs wie die folgende Ausführung von McKenzie Wark über die Darstellung in dem *Grand Theft Auto IV* Vorgänger *Vice City*:

The artful surfaces of the game are just a way for the gamer to intuit their way through the steps of the algorithm.³⁵

In Hinblick auf die übergeordneten Spielziele - „a vast accumulation of cash, cars, and cronies, of weapons and real estate“³⁶ - ist Warks Diagnose nicht von der Hand zu weisen. Im „ambience act“ sind diese Ziele aber außer Kraft gesetzt. Hier rücken gerade jene Algorithmen des Programms in den Vordergrund, die für die multimediale Darstellung der Stadt verantwortlich sind - von den kontinuierlich wechselnden Lichtverhältnissen des simulierten Tag-Nacht-Zyklus bis hin zu den prozedural erzeugten Soundkulissen.

Die Simulation der in Anlehnung an New York gestalteten Stadt Liberty City ist in *Grand Theft Auto IV* so umfangreich, dass alleine die Navigation durch sie für viele Spieler ein ansprechendes Erlebnis darstellt. Ihre Motivation liegt in der Erkundung der virtuellen Welt und der freien Interaktion mit ihren Objekten und Figuren. Ihr Modus Operandi zielt nicht auf das Gewinnen des Spiels oder die Progression einer Geschichte, sondern auf das Sein in der Spielwelt. Open World Spiele dieser Art, die ausreichend Interaktionsmöglichkeiten und eine freie Navigation innerhalb großer Areale bereitstellen, legen neue Spielweisen nahe.

³⁴(Galloway, 2006, 11)

³⁵(Wark, 2007, Absatz 120)

³⁶ebd.

Eine davon ist das Sightseeing prominenter Orte, das sich vor allem bei Nachbildungen realer Schauplätze anbietet. In Liberty City besuchen die Spieler markante Plätze wie den Time Square, die sie vielleicht schon einmal während ihres New York-Urlaubs besichtigt haben. Spiele der *Assasins Creed* Reihe ermöglichen es sogar, historische Schauplätze zu betreten. Während der erste Teil die Städte Jerusalem, Akkon und Damaskus zur Zeit des dritten Kreuzzugs simuliert, spielt die nachfolgende Trilogie im Italien der Renaissance. Ubisoft Montreal konsultierte während der Entwicklung Historiker³⁷, welche die inhaltliche Gestaltung unterstützten. An vielen Orten innerhalb der Spielwelt befinden sich Einstiegspunkte für Hintergrundinformationen über Schauplätze, Personen und Geschehnisse, die über Menüs abrufbar sind. Durch diese eingebauten Reiseführer wirkt das Spiel stellenweise wie eine Lernumgebung, in der sich die Spieler über historische Begebenheiten informieren können.

Beluschen

In Computerspielen, die auf die Darstellung einer glaubhaften fiktiven Welt aus sind, umfasst die Spielwelt neben dem Avatar und Quest-bezogenen Spielfiguren eine Reihe weiterer Nichtspieler-Charaktere (NPCs). Sie sind die Statisten des Spiels, die einem vorprogrammierten Tagesgeschäft nachgehen und deren Verhalten meist nur geringfügig durch die Spieler beeinflusst werden kann. Bereits der 1997 erschienene First Person Shooter *Jedi Knight* bereicherte in manchen Level-Abschnitten die Spielwelt mit NPCs, denen kurze Sprachsamples zugeordnet waren. Wenn sich der Avatar in ihrer Nähe befand, konnten die Spieler sie sprechen hören. Bis heute wurden diese Statistenrollen und ihr Sprachumfang konsequent ausgebaut, so dass in aktuellen Open-World-Spielen ganze Konversationen zwischen ihnen zu hören sind. Solche Unterhaltungen bilden ein weiteres erzählerisches Mittel, das verschiedene Funktionen erfüllen kann. Einerseits vermitteln die Konversationen Zusatzinformationen über die fiktive Welt und ihre Bewohner, die einen Beitrag zur Versenkung im virtuellen Raum leisten. Sie können auf diese Weise den aktuellen Status der fiktiven Welt referieren, vergangene Geschehnisse kommentieren oder auf Mißstände hinweisen, die kommende Quests stärker motivieren. In diesem Sinne fungieren sie als Reflektionsinstanz, welche die Entscheidungen der Spieler potentiell auf sie

³⁷(Hsu, 2010)

zurückwirft. Darüber hinaus lassen sich über die Gespräche aber auch Informationen transportieren, die in Hinblick auf die Spielmechanik förderlich sind. So können sich NPCs z. B. auch über die Schwachstellen eines Gegnertyps unterhalten oder die Orte von Schätzen verraten, die den Spielern nützliche Ressourcen verschaffen. Obwohl die Gespräche überwiegend narrative Funktionen übernehmen, sind Verweise auf das Spielmodell dementsprechend durchaus vorhanden.

Konversationen zwischen NPCs sind optionale Inhalte, welche die Spieler entweder ignorieren oder auf die sie sich einlassen können. Während des Flanierens durch den virtuellen Raum schnappen sie vielleicht nur Gesprächsfetzen auf, oder ihr Interesse wird geweckt, so dass sie stehenbleiben und zuhören. Im Grunde genommen bilden die Konversationen kleine theatralische Versatzstücke, die die Spieler für kurze Zeit in die Rolle des Beobachters einer Darbietung versetzen. Da die Spieler ihre Sicht- und Hörperspektive selber bestimmen, ähnelt die Situation stärker derjenigen eines Theaterstücks, in dem die freie Platzwahl möglich ist, als einer kinematographischen Inszenierung, in der die Einstellungen fest montiert sind. Sobald die NPCs ihr Skript nur dann aufführen, wenn sich der Spieleravatar außerhalb ihrer simulierten Sinneswahrnehmungen befindet, tritt zusätzlich eine voyeuristische Komponente hinzu. Aus dem Passanten, der ungezwungenen Gesprächen zuhört, wird ein Spion, der eine nicht für ihn bestimmte Unterredung gezielt belauscht. Sowohl die Inhalte der Konversationen, als auch die an ihre Ausführung geknüpften Bedingungen beeinflussen den narrativen Gehalt der Gespräche und deren Anknüpfungspunkte an das Spielmodell.

Stealth

Eine weitere Spielart, in der die Navigation durch Klangräume und die Simulation der akustischen Gegebenheiten fiktiver Welten eine zentrale Rolle spielt, ist das Schleichspiel. In reinsten Form tritt es in Spielereihen wie *Thief* und *Splinter Cell* auf, die maßgeblich das Stealth-Genre etablierten. Ein langsames und bedachtes Vorgehen wird in diesen Titeln stark durch die Spielmechanik forciert. Wer nicht schleicht, der scheitert in der Regel an den übermächtigen Gegnern im offenen Schlagabtausch. In abgeschwächter Form tritt das Schleichspiel auch als optionale Spielweise in Rollenspielen wie *Deus Ex* und *Fallout 3* auf. Hier können die Spieler durch den Ausbau des Schleich-Attributs

ihren Charakter gezielt darauf ausrichten, sich unbemerkt fortbewegen zu können.

Da die Spieler in diesen Titeln nicht gegen menschliche, sondern computergesteuerte Gegenspieler antreten, nimmt die Simulation der Sinne dieser künstlichen Intelligenzen einen wichtigen Platz ein. Der Begriff „Sinn“ steht in diesem Kontext als metaphorische Beschreibung für die Weiterleitung von Zustandsinformationen der Spielsituation an die Routinen der künstlichen Intelligenz. Werfen die Spieler beispielsweise in *Crysis* einen Stein in die Nähe einer Wache, geht sie dem Geräusch nach. Sie „hört“ den Aufprall. Im Grunde genommen handelt es sich um das einfache Weiterreichen von Zahlenwerten, welche die in der Regel als endliche Automaten modellierten KIs in potentiell neue Zustände übergehen lassen.³⁸

This engineering metaphor is not best applied too literally. In spite of the seemingly physical nature of the AIs in the game world, the analogy of game AI senses is not a physiological or neurological one. The line between "sense" and "knowledge" in a game is a blurry one. Sense incorporates the idea of awareness of another entity in the game, includes elements of value and knowledge, and can have game-relevant logic wired directly in.³⁹

Viele Computerspiele versehen die Übergänge zwischen Zuständen der KI mit Sprachsamples und machen sie damit für die Spieler transparent. Auf diese Weise wird ihnen z. B. mitgeteilt, dass sie von gegnerischen Figuren entdeckt wurden. Wenn in dem Star Wars-Spiel *Dark Forces* ein Sturmtruppler den Spieler sieht, triggert er ein Sprachsample wie „Da ist er!“, bevor er das Feuer eröffnet. Analog dazu spiegeln in *Portal* die Sprachsamples der Wachtürme wie „Is anyone there?“ und „I see you“ ihr Wissen über die Position des Spieleravatars wider. Die Mitteilung dieser Zusatzinformationen hat einen einfachen Zweck. Sie soll den Spielern Gelegenheit zur Reaktion geben und eventuellen Frustrationen vorbeugen, die eine nicht sichtbare Attacke mit sich bringt. Darüber hinaus kompensieren Sprachsamples möglicherweise ein schlechtes Verhalten der KI oder suggerieren mehr Verständnis als in den KI-Routinen implementiert ist.

³⁸ An Stelle von endlichen Automaten kommen auch weitere Planungssysteme, wie z. B. der Stanford Research Institute Problem Solver (STRIPS), zur Anwendung.

³⁹ (Leonard, 2003)

Die simulierten Sinne der KI lassen sich konzeptuell in die Bereiche Hören und Sehen einteilen, wobei der größte Unterschied zwischen ihnen darin besteht, welche Informationen der virtuellen Welt an die KI weitergeleitet werden. Der Sehsinn ist auf eine Richtung und einen Blickwinkel beschränkt, während der Hörsinn einen Kreis um die computergesteuerte Spielfigur in der Größe ihres Hörradius spannt. Es ist unschwer erkennbar, dass die Implementierung eines Sehsinns aufwändiger ist, da die jeweilige Blickrichtung der Figur nachgehalten werden muß. Beim Hörsinn müssen die Koordinaten der zu entdeckenden Figur des Spielers lediglich innerhalb des Hörradius der computergesteuerten Figur liegen. Bei einfacheren Implementierungen der KI-Sinne ist eine Differenzierung zwischen Sehen und Hören deshalb nicht vorgesehen. Sie statten die Spielfiguren der Einfachheit halber mit einem Wahrnehmungsradius aus, der alle Sinne in sich trägt. Genau genommen ist es also erst die Implementierung eines gerichteten Sehsinns, die aus dem Wahrnehmungsradius einen Hörsinn macht und unterschiedliche Handlungen für unterschiedliche Sinneswahrnehmungen ermöglicht.

Die Nutzung solcher Sinnessimulationen ist nicht auf Spiele mit einer dreidimensionalen Raumstruktur beschränkt. Bereits das in einer isometrischen Perspektive gehaltene Strategiespiel *Jagged Alliance* simuliert für jede Spielfigur einen Seh- und einen Hörsinn, die durch ausrüstbare Gegenstände wie Nachtsicht- und Abhörgeräte zusätzlich geschärft werden können. Wenn die Söldner des Spielers einen Gegenspieler sichten oder auffällige Geräusche hören, geben sie die Information über ein Sprachsample an den Spieler weiter. Der entscheidende Unterschied dieser Sinnessimulation zu derjenigen im First Person Shooter besteht darin, dass sich die Wahrnehmungsperspektiven des Avatars und der Spieler überlagern. Der Avatar ist keine Spielfigur, die den Spielern mitteilen muß, was sie gehört und gesehen hat. Die Spieler sind selber diejenigen, die in der Rolle des Avatars spähen und horchen.

Audiologs

In mehreren virtuellen Spielwelten der vergangenen Jahre wie u.a. *Doom 3*, *Batman Arkham Asylum*, *Fallout 3*, *Dead Space* und *Bioshock* finden die Spieler besprochene Tonbänder und Aufnahmen. Diese abspielbaren Notizen erfüllen primär narrative Funktionen. Sie etablieren die Hintergrundgeschichte, erzählen von vergangenen Ereignissen und bieten

intime Einblicke in die Psyche und das gegenseitige Verhältnis von noch auftauchenden oder längst verschollenen Charakteren. Meistens liegen die Inhalte in persönlicher Form wie einem Tagebucheintrag oder einem kurzen Memorandum vor. In *Bioshock* erfahren die Spieler z. B. über diverse verstreute Fetzen von Audio-Tagebüchern, wie es zu der verheerenden Lage der Unterwasserstadt Rapture und ihrer dem Irrsinn verfallenen Bevölkerung kommen konnte. Die 2009 erschienene Comic-Adaption *Batman Arkham Asylum* legt den Spielern mit ebenfalls in der Spielwelt verteilten Tonbändern mehrere Patientengespräche von Batmans Widersachern an die Hand, die sie über deren kriminellen Werdegang aufklären. In *Doom 3* enthalten die Audiologs Codes, mit denen die Spieler Container öffnen können, die hilfreiche Accessoires wie Munition und Medikits enthalten. Analog zu belauschbaren NPC-Konversationen vermitteln auch Audiologs mitunter spielrelevante Informationen.

Die narrative Technik, Story-Fragmente in Form von auflesbaren Gegenständen innerhalb der Spielwelt zu platzieren, besitzt im Genre des First Person RPG Shooters eine langjährige Tradition. Nachdem bereits im 1994 erschienenen Titel *System Shock* von Spieldesigner Warren Spector eine Reihe von Datendisks in der Spielwelt auftauchen, die die Spieler lesen können, wurde vier Jahre später in *System Shock 2* das Konzept erneut aufgegriffen und in Form von hörbaren Sprachaufnahmen umgesetzt. Diese Audiologs lassen Sprache dort sein, wo keine Personen sind. Sie erlauben keine Rückfragen, keine Interaktion und eignen sich dadurch ausgezeichnet für ein Genre wie den First Person Shooter, in dem die Interaktionsfähigkeiten des Avatars größtenteils auf die Positionierung innerhalb des dreidimensionalen Raums sowie den geschickten Einsatz von Gegenständen und Waffen beschränkt sind. Am konsequentesten verwendet bislang *Dead Space* die Erzähltechnik, indem dort mit Text-, Audio- und Videologs gleich drei verschiedene mediale Typen auftreten, die sowohl als Gegenstände in der Spielwelt verteilt sind, als auch über Funkübertragungen die Spieler erreichen.

Die einzelnen Einträge sind dabei in der Regel nicht über einen linearen Handlungsstrang miteinander verknüpft, sondern bilden bruchstückhafte Erinnerungen vergangener Ereignisse. Dadurch ist die Reihenfolge, in welcher die Spieler die Einträge aufsammeln, nicht entscheidend. Die Erzählweise funktioniert wie ein Puzzle, das die Spieler nach und nach zusammensetzen, und das ihnen ein fragiles Bild der Spielwelt eröffnet. Die durch den fragmentarischen Charakter bedingte Unvollständigkeit,

die auch nach dem Einsammeln aller Puzzle-Stücke bestehen bleibt, suggeriert den Spielern, die Lücken durch eigene Vorstellungen zu füllen. Neben den gedanklichen Freiräumen, die sich für die Spieler öffnen, trägt die Existenz archivierter Daten zur Etablierung des Spieluniversums bei. Es besitzt eine erinnerbare Geschichte und war schon da, bevor es betreten wurde.

Aus Gameplay-Sicht stellen die Bänder sammelbare Gegenstände dar. In *Batman Arkham Asylum* fügt jedes gefundene Band einen Eintrag auf der Personenseite des jeweiligen Bösewichts in der freispielbaren Galerie hinzu. Die Patientengespräche werden dadurch zu sammelbaren Artefakten, welche die Galerieseiten der Antagonisten wie Trophäen schmücken und Spieler zur intensiven Suche nach ihnen in der Spielwelt motivieren. In *Bioshock* erhalten die Spieler bei der Xbox 360 Version z. B. das „Historian Achievement“⁴⁰, wenn sie alle der 122 Audiologs einsammeln.⁴¹

Die Einbindung der Sprach- und Gesprächsfetzen geschieht in allen oben genannten Fällen ähnlich. Das Tonband oder die Aufnahme befindet sich an einer bestimmten Position innerhalb der Spielwelt und wartet auf die Entdeckung durch die Spieler. Die Darstellung reicht von CD-ähnlichen Datenträgern bis hin zu kompletten Wiedergabegeräten, in denen die Bänder jeweils integriert sind. In manchen Fällen sind die Aufnahmen bereits in symbolischer Form als Dateien in anderen Geräten wie den PDAs aus *Doom 3* enthalten und entbehren einer materiellen Form. Ähnlich wie bei der Jukebox⁴² lässt sich eine Präferenz für betagtere physikalische Datenträger feststellen, die aufgrund ihrer Größe leichter in der Spielwelt zu orten sind.

Nehmen die Spieler das Speichermedium auf, so verschwindet dessen visuelle Form in der Spielwelt und ein Audio-Eintrag in einem Menü wird freigeschaltet. Dieser Wechsel von der visuellen Ebene des Ortes innerhalb der Spielwelt auf die funktionale Abstraktionsebene der Benutzeroberfläche löst den sichtbaren Gegenstand mit fixen Koordinaten auf und verwandelt ihn in eine gegenstandslose Datei, die mit der Funktion Abspielen gekoppelt ist. Häufig taucht der Gegenstand nicht im Inventar, sondern in einem speziell dafür vorgesehenen Menü

⁴⁰Achievements sind vorgegebene optionale Spielziele, die häufig unkonventionelle Spielweisen im Sinne von emergent gameplay fördern.

⁴¹siehe http://bioshock.wikia.com/wiki/Audio_Diaries

⁴²siehe auch 6.3 auf Seite 240

auf und büßt damit seine simulierte Materialität komplett ein. Ähnlich wie bei textuellen Zusatzinformationen, die vor allem in Rollenspielen in Menüstrukturen ausgelagert werden, erfolgt die Transformation des Audio-Eintrags in einen Menüpunkt. Derartige Menüs dienen nicht der Konfiguration. Vielmehr bilden sie eine Form von Gedächtnis, das sämtliche innerhalb der virtuellen Welt aufflesbaren Informationen – seien es textuelle, akustische oder audiovisuelle Artefakte – speichert und abfragbar macht. Einmal eingesammelt, lässt sich der jeweilige Eintrag nicht mehr entfernen, nicht mehr vergessen. Der eigentliche konfigurative Aspekt der Audiologs liegt in ihrer Verteilung, ihrer Konfiguration innerhalb der Spielwelt während des Entwicklungsprozesses.

Beim Vorgang des Abspielens besteht nun eine unmittelbare Kongruenz zwischen dem, was die Spieler im Spiel bezwecken und dem, was tatsächlich in den Interna des ausführenden Codes vonstatten geht: die Wiedergabe einer Audio-Datei. Es gibt keinen internen Konstruktionsakt, der den hörbaren Output aus mehreren Stücken zusammensetzt. Abgesehen von möglichen Umgebungseffekten wie Hall handelt es sich um eine eins-zu-eins Relation. Der einzige Unterschied besteht in der Art des Mediums und der Beschaffenheit des Abspielgeräts. Anstelle eines Tonbands, das von einer analogen Bandmaschine abgespult wird, hören die Spieler den Output ihrer Soundkarte, welche die digital kodierten Wellenformen in analoge Signale rückübersetzt.

Theoretisch könnte man die narrative Strategie der Audiologs auch problemlos in Spielen anwenden, die vor dem 20. Jahrhundert und der Erfindung des Phonographen spielen. Der Träger wäre die Schrift. Bei der Wiedergabe würde der Spieleravatar lesen, und der Urheber der Nachricht schreiben und dabei laut nachdenken. Im Film ist dieses Verfahren bereits etabliert. Auf gleiche Weise könnte in einem historischen Setting ein Videolog untergebracht werden, indem es den Schreibenden zeigt, der vor dem inneren Auge des Lesenden zum Leben erwacht.

5 Adaption kinematographischer Gestaltungsmuster

Die akustische Gestaltung des Computerspiels besitzt Vorbilder, die ihre Entwicklung prägen. Orientierte sich das Sounddesign von Computerspielen Anfang der 1970er Jahre an elektromechanischen Spielautomaten, gilt vor allem ab den 1990er Jahren der Film als Vorbild. Wechselbeziehungen zwischen Computerspielen und Filmen prägen die Entwicklung beider Medien. Sie zeigen sich einerseits in Computerspiel Fassungen von Filmen, Filmadaptionen von Computerspielen sowie den Bemühungen um interaktive Filme, und andererseits in der gegenseitigen Übernahme von Gestaltungsmustern. Filme experimentieren mit der Kontinuität der First Person Perspektive und Computerspiele adaptieren filmische Gestaltungsmuster wie den „over the shoulder shot“. Trotzdem scheint der Film für die Gestaltung des Computerspiels eine stärkere Vorbildsfunktion einzunehmen, als umgekehrt. Viele Game Designer beziehen sich auf Filme oder eine bestimmte Ära des Films, wenn sie die Inhalte, Gestaltungsrichtlinien und den Stil ihrer Spiele beschreiben. Auch die Rahmung von Spielen im Stil des Films mit scrollenden Credits im Abspann, oder die Segmentierung eines Spiels in mehrere Episoden wie bei einer Fernsehserie¹ sind keine Seltenheit. Viele Gründe lassen sich anführen, warum der Film zu den größten Inspirationsquellen der Gestaltung des Computerspiels gehört. Er verwendet die gleichen Ausgabegeräte² und bietet eine elaborierte, historisch gewach-

¹Ein Beispiel dafür ist die parallel stattfindende Film- und Computerspiel-Adaption der Comicreihe *The Walking Dead*. Der adaptierte Stoff erscheint als Fernsehserie auf dem US-amerikanischen Sender AMC und als episodentartige Spielereihe auf den einschlägigen Spielevertriebsplattformen.

²Während Spielekonsolen seit ihrer Entstehung für den Anschluß an Fernseher und Stereoanlagen konzipiert waren, unterstützt die Vernetzung dieser Geräte neuerdings auch die Darstellung von Browsern und PC-Games. Siehe z. B. *Steam Big Picture*: <http://store.steampowered.com/bigpicture/> (23.10.2012)

sene Formsprache, die aus der Verschmelzung von Bild und Ton im Audiovisuellen aufgeht. Das große Repertoire ästhetischer Strategien und die qualitativen Maßstäbe von Hollywood-Produktionen scheinen für Computerspiele ebenso erstrebenswert zu sein wie das Prestige des Films als gesellschaftlich anerkannte Kunstform. Die Form der akustischen Architektur des Computerspiels ist dementsprechend stark durch filmische Vertonungsmuster beeinflusst.

Vor allem nach dem Jahrtausendwechsel adaptierte das Computerspiel in hohem Maße Produktionsprinzipien und Darstellungsformen des Films, die insbesondere auch die akustische Gestaltung betreffen. So bemerkt der Sounddesigner Rob Bridgett:

It became clear to me around 2004 [...] that video games were indeed consciously seeking the production values of Hollywood cinema. [...] The move was also being mirrored significantly in sound technology, with many traditional film sound technologies and standardizations quickly becoming adopted by video games.³

Trotz aller Ähnlichkeiten und dem dadurch entfachten Diskurs der Medienkonvergenz dürfen die Unterschiede nicht außer Acht gelassen werden. Filme bieten schließlich weder Eingriffsmöglichkeiten, die den weiteren Verlauf des Geschehens beeinflussen, noch enthalten sie Spielsysteme, deren Regeln, Ziele und Freiräume die Eingriffe motivieren. Ästhetische Gestaltungsmuster des Films durchlaufen Anpassungen, wenn sie innerhalb des Computerspiels und dessen akustischer Architektur aufgegriffen werden. Sie treten in neue Funktions- und Bedeutungszusammenhänge, die sowohl durch die medialen Charakteristiken, als auch die inhaltlichen Unterschiede des Computerspiels zustande kommen.

Drei Adaptionfelder sollen im Folgenden in Hinblick auf die Transformationen adaptierter kinematographischer Gestaltungsmuster durchschritten werden. Das erste Feld untersucht die Adaption nonverbaler Vertonungsprinzipien. In Rückgriff auf Kapitel 2 steht die Beziehung zwischen dem in Filmen und Computerspielen verwendeten Klangmaterial zur Disposition. Im Anschluß daran folgen Überlegungen zu den Anpassungen übernommener Vertonungsprinzipien, die anhand der Leitmotiv-Technik exemplifiziert werden. Das zweite Feld behandelt die Frage, inwiefern die Gestaltung inszenierter Szenen vor dem Hintergrund der

³(Bridgett, 2010, 9)

Interaktion möglich ist. Nach dem Abschreiten von Strukturabschnitten des Spiels mit unterschiedlichen Interaktionsspielräumen werden Techniken zur szenischen Inszenierung von spielbaren Abschnitten vorgestellt. Der Inhalt des letzten Feldes ist die Integration gesprochener Sprache. Verschiedene Adaptionen der Voice-Over Technik rücken nach einer Skizze der Funktion von Sprache im Kontext des Spiels in den Fokus.

5.1 Anpassung filmischer Vertonungsprinzipien

Konvergenz des Klangmaterials

Eine wichtige Brücke zwischen Filmen und Computerspielen bildet die Übertragung von Musik und Soundeffekten. Bereits zu den Zeiten der Klangsynthese imitierten Computerspielsounds markante Geräusche aus der Domäne des Films. Sie kamen überwiegend in Computerspieladaptionen erfolgreicher Hollywood-Filme zur Anwendung. So enthielt die 1992 für das Sega Mega Drive erschienene Adaption des ersten Terminator-Films ein Pixel-Gemälde von Arnold Schwarzenegger und eine Chiptune-Fassung der Titelmelodie des Films. Auch Computerspiele ohne direkten Filmbezug integrierten synthetisch erzeugte Abbildungen stereotyper Filmsounds. Wenn sich in *Pitfall* (Atari 2600, 1982) die Spielfigur mit der Liane über Sümpfe schwingt, erklingt eine synthetische Version des berühmten Tarzan Rufes von Johnny Weissmüller. Wie nah die adaptierten Filmsounds und Soundtracks an ihre Vorbilder heranreichten, ergab sich einerseits aus den technischen Möglichkeiten der Soundchips und daraus, wie versiert die Sound Designer im Umgang mit ihnen waren. Andererseits bestimmte auch die Komplexität der nachzubildenden Wellenformen die erreichbare Nähe zum Original. Viele Geräusche von Spezialeffekten, insbesondere im Genre des Science Fiction-Films, die ohnehin synthetischen Ursprungs waren, ließen sich dementsprechend überzeugender umsetzen als z. B. die Klanglandschaft einer natürlichen Umgebung. Die Adaption der Filmmusik reichte in Abhängigkeit der verfügbaren Kanäle von der einstimmigen Abbildung markanter Melodien bis hin zu mehrstimmigen Arrangements, die ihren orchestralen Vorbildern nachempfunden waren.

Die Einführung des Samplings knüpfte ein noch festeres Band, da Filme und Computerspiele von da an auf die gleichen Sample-Bibliotheken zugreifen konnten. Computerspiele, die im Star Wars Universum spielten, verwendeten nun die aus Filmen bekannten Geräusche von Lichtschwertern, Lasern und Tie-Fightern. Darüber hinaus ließ sich nun die orchestrale Musik von John Williams einbinden, die vorher nur in Form von Chiptunes oder Midi-Tracks zur Verfügung stand. Das Material, auf dessen Basis Filme und Ausgaben des Computerspiels entstehen, ist heute stellenweise identisch. Sample-Bibliotheken richten sich gleichzeitig an das Sounddesign von Filmen und an das Sounddesign von Computerspielen. Auch dann, wenn die akustischen Bausteine nicht identisch sind, weisen sie starke Ähnlichkeiten auf. Computerspiele adaptieren die hyperreale Geräuschästhetik und die musikalische Untermalung im Stil des Hollywood-Films. Selbst Hans Zimmer war zwischen seinen Kompositionen für *Inception* und *The Dark Knight Rises* an dem Game-Soundtrack von *Crysis 2* beteiligt.

Bei der Gegenüberstellung von Spielmitschnitten und Filmaufnahmen kam lange Zeit kein Zweifel auf, was Spiel und was Film ist. Abgesehen von wenigen Ausnahmen wie dem 1983 von Cinematronics veröffentlichten Arcade-Automaten *Dragon's Lair*, der Zeichentrick-Sequenzen in Disney-Manier enthielt, wiesen Computerspiele kaum eine Ähnlichkeit zu Filmen auf. Mit fotorealistischer Grafik und hochauflösenden Samples, die Hollywood in nichts nachstehen, beginnt diese Differenz zu bröseln. Im Jahr 2011 bedarf es einiger Anstrengung, um die aneinander geschnittenen Film- und Spielszenen eines Trailers zum Shooter *Battlefield 3* auseinander zu halten. Die provokante Frage am Ende des Spots „Is it real or is it Battlefield 3?“ ließe sich umformulieren in die Frage: Ist es ein Film oder ist es ein Spiel? Das Material, aus dem die sinnlich wahrnehmbare Form des Films und des Computerspiels zusammengesetzt wird, hat sich mittlerweile so stark angenähert, dass die Unterschiede verschwimmen. Filme verwenden computeranimierte Szenen und Spiele erreichen nahezu die fotorealistische Qualität von Filmen. In Hinblick auf die Gestaltung des sinnlich Wahrnehmbaren ist die Konvergenz zwischen Filmen und Computerspielen unverkennbar. Bleibende Unterschiede des Klangmaterials im Computerspiel, wie die Bildung von Klanggruppen zur Kaschierung von Repetitionen, oder die Berechnung akustischer Effekte in Echtzeit, bleiben für die User und Zuschauer verborgen.

Funktionalisierung adaptierter Vertonungsprinzipien

Die Ähnlichkeiten der Vertonungen von Computerspielen und Filmen sind nicht nur in Hinblick auf das Material, sondern auch den Einsatz des Materials zu beobachten. Computerspiele adaptieren von der Leitmotivik bis hin zur kontrastreichen Verwendung von Lärm und Stille diverse Vertonungsprinzipien, die im Film mehrfach erprobt worden sind und sich mitunter zu Konventionen verdichtet haben. Trotz der oberflächlichen Ähnlichkeiten läuft die Einbindung des Materials im Computerspiel unter anderen Vorzeichen ab. Sie folgt den in Kapitel 3 und Kapitel 4 dargelegten Integrationsstrategien. Das akustische Material wird dabei an Zustände, Ereignisse oder Orte gebunden. Im Zuge dieser Kopplung treten die Samples in einen neuen Organisationskontext ein, der sich drastisch von den synchronisierten Szenen des Films unterscheidet. Sie sind in eine Metastruktur eingebunden, die die Bedingungen definiert, unter denen ein Sample abgespielt wird. Eine charakteristische Anpassung von Musik bei der Einbindung in diesen Organisationskontext besteht z. B. darin, sie in eine repetitive Struktur zu überführen, die der zeitlichen Ungewißheit des Spielablaufs Rechnung trägt.

Durch die Kopplung an Spielereignisse und -zustände erhält das akustische Material Funktionen im Rahmen der Interaktion. Es dient nicht nur der Darstellung, sondern leistet gleichzeitig die Sonifikation von Informationen, auf deren Grundlage die Spieler ihre Entscheidungen treffen. Das akustische Material tritt in den von Suggestion und Feedback geprägten Funktionszusammenhang der Mensch-Computer-Schnittstelle.

Dieser Vorgang der Funktionalisierung lässt sich deutlich an der Adaption der Leitmotiv-Technik beobachten. Sowohl Leitmotive, die bereits aus Filmen bekannt sind, als auch leitmotivisch orientierte Neukompositionen für bestimmte Spiele treten in den Funktionszusammenhang des regelbasierten Systems ein. Der erste Fall ist bereits in dem Atari 2600 Spiel *Star Wars: The Empire Strikes Back* enthalten:

When the snowspeeder endures for two minutes in *Star Wars: The Empire Strikes Back*, it wins the temporary invulnerability of „the Force“ and the *Star Wars* theme plays, as it does when the cartridge first starts up. This rare musical treat effectively draws a connection to the *Star Wars*

movies and also works effectively in the game, making the period of invulnerability even more heightened.⁴

Einerseits sonifiziert das Star Wars-Thema die 20 sekündige Zeitspanne des Unsterblichkeits-Powerups. Andererseits transportiert es Assoziationen und Werte, die sein leitmotivischer Einsatz in den Filmen nahelegt, in das Spiel hinein. Wer die Filme kennt, versteht das Motiv womöglich als Hoffnungssignal der Rebellion gegen das übermächtige Imperium, das bereits durch die Synchronisation des Motivs mit dem Einleitungstext der Episode IV „A new hope“ etabliert wird. Im Spiel erhält der temporäre Vorteil gegen die endlose Reihe imperialer Walker dadurch ein zusätzliches narratives Gewicht.

Den zweiten Fall einer leitmotivisch ausgelegten Neukomposition bietet die *Left 4 Dead* Reihe. Das gesamte Spielgeschehen ist von wiederkehrenden Motiven durchdrungen, die eine Reihe von Ereignissen begleiten. Um eine kompositorische Kohärenz zu schaffen, basieren alle Motive auf dem gleichen Tonmaterial. Bill Van Buren erläutert das angewandte Konzept folgendermaßen:

All the music in *Left 4 Dead* – themes, motifs, hits and effects – are based on a hybrid scale derived from the main opening menu theme. This theme, representing the near death of humanity as it fades to a distance only to be accessed through a radio, is an incremental melodic modulation collecting up all the notes used along the way. This scale is not absolute, and some notes outside the scale are used occasionally to achieve even greater dissonance; but by using this singular and very chromatic scale, almost any piece written in it will generally dovetail pretty decently with any other. At the very least, the pieces seem to spring from the same musical universe. The resulting tunes are deceptively complex – easy to remember but difficult to sing accurately.⁵

Eine besondere Kategorie der kurzen Motive sind die Ankündigungen der sog. Spezialinfizierten. Jeder Spezialinfizierte besitzt ein eigenes Thema, das abgespielt wird, wenn eine räumliche Nähe zu den Überlebenden besteht. Motive dieser Art sind vor allem im Genre des Horror-

⁴(Montfort u. Bogost, 2009, 132)

⁵Bill Van Buren, *Left 4 Dead* Developers Commentary

films zu finden, übernehmen dort jedoch keine dezidiert leitmotivische Funktion. Statt dessen erzeugen sie durch ihre unaufgelöste Dissonanz Spannung und Unsicherheit in Situationen wie z. B. der Erkundung eines unbekanntes Gebiets. *Left 4 Dead* überträgt diese musikalische Gestaltung, welche in vielen Zombie-Filmen zur Anwendung kommt, in das Spiel und versieht sie mit einer spielerischen Funktion. Anstatt ein unbekanntes Schrecken anzudeuten, verraten sie den Überlebenden, welche Spezialinfizierte auf sie lauern. Das Erlernen dieser musikalischen Codes verschafft spielerische Vorteile, da jeder Spezialinfizierte eine bestimmte Angriffsmechanik besitzt, der unterschiedlich entgegen gewirkt werden kann. Über den Einsatz der Leitmotive besitzen die Spieler keinen direkten Einfluss. Sie sind reine Informationsgeber in der Gestalt eines gebräuchlichen Kompositionsschemas des Films.

Neben den Melodien bekannter Stücke kommen in Computerspielen auch konventionalisierte Instrumentierungen und Arrangements vor, die sich im Film zur Begleitung bestimmter Szenen verfestigt haben. So verwendet *Dead Space* z. B. stechende Streicherspitzen, wie sie in *Psycho* und etlichen weiteren Filmen des Horror-Genres zu hören sind. Im *Left 4 Dead* Spielmodus „Scavenge“ sind lang anhaltende, dissonante Akkorde – ebenfalls ein weit verbreitetes dramaturgisches Mittel des Horror-Sounddesigns – zu finden. Das Ziel des Spielmodus ist es, einen Generator mit möglichst vielen Kanistern zu füllen, die

innerhalb der Karte verstreut sind. Mit jedem Kanister gewinnt die Partei der Überlebenden ein paar wertvolle Sekunden dazu. Sobald der Countdown abgelaufen ist, aber noch mindestens einer der Überlebenden einen Kanister trägt, ertönt jene dissonante Klangfläche. Sie unterstreicht den Umstand, dass das Spiel an einem seidenen Faden hängt. Entweder schaffen es die Überlebenden, den Kanister einzufüllen und das Zeitkontingent zu erhöhen oder die Zombie-Fraktion hindert sie daran, um die Runde zu beenden und mit vertauschten Rollen ein bes-

Boomer:



Hunter:



Smoker:

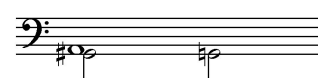


Abbildung 5.1: Leitmotive der Spezialinfizierten aus *Left 4 Dead*

seres Ergebnis zu erzielen. Während dieses dramaturgische Element im Film eine Schrecksekunde einfügt und den Ausgang einer Szene verzögert, begleitet es im Spiel den potentiell entscheidenden Ausgang eines Durchlaufs. Gleichzeitig erbt es aber auch den Schreckensmoment der kinematographischen Inszenierung und überträgt sie auf das Gewinn- und Verlust-Schema des Spielsystems.

Es bleibt festzuhalten, dass heute, stärker als je zuvor, Vertonungsstrategien des Films von Computerspielen adaptiert werden. Gerade in spielbaren Sequenzen handelt es sich jedoch nur um eine oberflächliche Ähnlichkeit, die primär auf das in beiden Medien verwendete Sample-Material zurück zu führen ist. Im Rahmen der Spielerfahrung ergeben sich in Abhängigkeit der Verschaltung akustischer Elemente mit den Spielregeln weitere Funktionszusammenhänge und Bedeutungskomplexe.

Simulation des Kinematographischen Apparats

Computerspiele greifen nicht nur frühere Entwicklungsstufen ihrer eigenen Geschichte auf, wie die 8-Bit Ästhetik, sondern bedienen sich auch vergangener Entwicklungsstufen des Films. Einige Titel simulieren den kinematographischen Apparat, indem sie das Bild mit flackernden Effekten und den Ton mit Knistern und Rauschen unterlegen. Dabei handelt es sich um einen Spezialfall der Adaption kinematographischer Gestaltungsmuster. Anstatt einzelne Muster zu übernehmen, ist die gesamte Darstellung von dieser übergeordneten Design-Richtlinie betroffen. Bild und Ton sind mit Filtern versehen.

Die Simulation des kinematographischen Apparats kommt u. a. als rein stilistischer Effekt zur Anwendung, der im Rahmen des Gameplays keine besondere Funktion übernimmt. In diesen Fällen bleiben die akustischen und grafischen Störartefakte das gesamte Spiel über aktiv. Solche Filtereffekte, die den gesamten Spielablauf begleiten, sind vor allem in Independent Games zu beobachten. Sie sind ein Mittel, um das Spiel durch eine außergewöhnliche Präsentation von der Masse abzuheben. Zwei Beispiele dafür sind die Spiele *Limbo* und *The Misadventures of P.B. Winterbottom*.

In *Limbo* sind die Ränder des flackernden Bilds schwarz gefärbt, als ob ein alter Projektor einen ausgebleichenen Film absputt. Assoziationen an die Filmkunst liegen unmittelbar nahe.

Limbo is presented in grainy black and white, looking like a cross between the animated visions of Tim Burton and the monochromatic art films of Fritz Lang or Ingmar Bergman.⁶

Die akustische Gestaltung orientiert sich an diesem reduzierten Stil und der Unschärfe des Hintergrunds.

The whole concept of the visuals, the horizon is always blurred so you can project your own things into the spaces. I tried to do the same with the sound with noise and textures—you start to hear things that aren't there.⁷

Auf eine andere Ära des Films bezieht sich *The Misadventures of P.B. Winterbottom*:

The Winterbottom Art Style is inspired by the work of writer/illustrator Edward Gorey and silent film greats Harold Lloyd, Buster Keaton, and Charlie Chaplin.⁸

Wie beim Stummfilm untermalt eine Klavierbegleitung das Spielgeschehen.⁹ Sobald die Spieler die zur Lösung der Rätsel nötige Zeitschleife einsetzen, tritt ein Spulgeräusch hinzu - ein weiteres Relikt analoger Medien.

Auch in größeren Produktionen sind solche globalen Filtereffekte einzeln zu finden. *Left 4 Dead* enthält einen Grieseeffekt, dessen Intensität per Menü einstellbar ist. Häufiger unterliegen in solchen Titeln die Rückgriffe auf historische Praktiken und Techniken der Filmkunst den Funktionalisierungen des Spiels, indem sie an Ereignisse und Zustände gekoppelt werden. Anstatt das gesamte Spiel zu begleiten, werden sie nur zu bestimmten Momenten aktiv.

Ein Beispiel hierfür ist die Überblendung in ein Schwarz-Weiß Bild, welche abhängig von dem jeweiligen Spielkontext unterschiedliche Bedeu-

⁶<http://www.theage.com.au/digital-life/games/blogs/screenplay/split-screen-looking-good-20100730-10yc8.html> (21.03.2012)

⁷(Thomsen, 2010)

⁸<http://www.winterbottomgame.com/> (21.03.2012)

⁹Manche Level enthalten zur Abwechslung auch orchestrale Arrangements.

tungen annehmen kann. Verliert in *Left 4 Dead* oder *The Witcher* das Bild seine Farbe, schwebt das Leben des Avatars in höchster Gefahr. Die Ausblendung der Farbinformation symbolisiert die schwindenden Sinne der Spielfigur. In *Splinter Cell Conviction* geht das Bild in Schwarz-Weiß über, wenn der Avatar in Deckung liegt und nicht von gegnerischen Figuren gesehen wird. Die Reduktion der Farbwerte bedeutet hier das Gegenteil, nämlich Sicherheit und Schutz an Stelle von Gefahr. *Fall-out 3* enthält eine komplette Quest, die in Schwarz-Weiß gehalten ist. Sobald der Avatar in Vault 112 die virtuelle Realität „Tranquility Lane“ betritt, wird die gesamte Umgebung in Schwarz-Weiß dargestellt. Die Reduktion der Farbwerte symbolisiert zum Einen die Künstlichkeit der virtuellen Welt innerhalb der virtuellen Welt des Spiels. Zum Anderen etabliert sie gerade durch die Nutzung einer veralteten kinematographischen Technik eine Referenz auf die zurückliegende Epoche, in der die virtuelle Realität einer unberührten Wohnsiedlung vor dem Ausbruch des Kriegs angesiedelt ist. Das musikalische Thema innerhalb der Simulation verstärkt den Gedanken an vergangene Zeiten, da es das Gleiche ist, das die Spieler während des anfänglichen Tutorials in der Rolle eines Kleinkinds zu hören bekommen, während sie ihre ersten Schritte lernen.

Exkurs: Bullet-Time Aus der Übertragung markanter Filmszenen in das Computerspiel entwickeln sich nicht selten neue Spielmechaniken. Ein prominentes Beispiel stellt die Verarbeitung der Zeitlupe-Sequenzen aus dem Film *Matrix* in dem Spiel *Max Payne* dar. Die längste Szene, die in *Matrix* durch den massiven Einsatz der Zeitlupe heraussticht, findet in der von Marmorsäulen gesäumten Lobby eines Hochhauses statt. Neo, der unter seinem schwarzen Mantel ein regelrechtes Waffenarsenal verbirgt, bahnt sich durch den Hagel auf ihn herabregnender Kugeln und zerberstender Wände in einer akrobatischen Choreographie den Weg zum Aufzug am anderen Ende der Lobby. Auch der Held Max Payne aus dem gleichnamigen Spiel ist komplett in schwarz gekleidet und zeigt während des zweihändigen Waffengebrauchs ähnliche Gesten wie Keanu Reeves in seiner Rolle als Neo. Der Zeitlupeneffekt, der im Film als stilistisches Mittel fungiert, wurde im Spiel in ein Gameplay-Feature umgewandelt. Die Spieler sind diejenigen, die den Effekt auslösen, sobald in einem Feuergefecht eine brenzlige Situation entsteht. Der nur begrenzt nutzbare Effekt verlangsamt alle Bewegungen in der

fiktiven Spielwelt, behält jedoch die ursprüngliche Zielgeschwindigkeit des Spielers bei.

Das Spiel transformiert das stilistische Mittel in ein strategisches. Die Bullet-Time in Max Payne kombiniert die Ästhetik des Zeitlupeneffekts mit der spielerischen Notwendigkeit einer Gameplay-Funktion, die den Spielern durch die Verlangsamung des Spielgeschehens in entscheidenden Situationen einen Vorteil verschafft. Dadurch wird einerseits eine Assoziation mit Filmen wie Matrix hergestellt, aus denen der Effekt seine Popularität erhielt. Darüber hinaus erweitert sich der Bedeutungshorizont des Effekts durch die Einbindung in ein interaktives System und die ihm im Zuge dessen zugeschriebene Funktionalität. Der Zeitlupeneffekt ist im Kontext des Spiels eine Ressource, welche die Spieler verbrauchen und wieder aufladen können. Ähnliche Funktionalisierungen und Bedeutungsverschiebungen, die bei der Adaption diskreter Elemente der Bildsprache des Films innerhalb des Spiels auftreten, sind auch in Hinblick auf die Tonsprache zu beobachten.

5.2 Szenengestaltung

Eine der größten Herausforderungen bei der Adaption kinematographischer Formen im Rahmen des Computerspiels besteht in der Szenengestaltung. Die Stärke des Films liegt gerade in der sorgfältigen Komposition einer audiovisuellen Sequenz. Jedes kleinste Detail in dieser zeitlich und räumlich fixierten Ordnung – der Montage und der Mise en scène – leistet seinen Beitrag zu der Gesamtheit einer Szene. Inwiefern ließe sich ein solches Gefüge in ein Computerspiel übertragen? Eine offensichtliche Schwierigkeit besteht in den Ungewißheiten der Interaktion, die eine präzise Planung der Entfaltung einer Szene unterläuft. Der Begriff des Audiovisuellen gerät im Kontext des Computerspiels ins Wanken, wenn man darunter das fixierte Zusammenspiel einer statischen Bild- und Tonspur versteht. Unterschiedliche Entscheidungen enden in unterschiedlichen Ausgängen.

Interaktionsspielräume und die darin realisierbaren Formen kinematographischer Inszenierung lassen sich schwer auf die akustische Architektur des Spiels reduzieren. Statt dessen rückt die gesamte grafische und akustische Ausgabe in den Mittelpunkt der Betrachtung.

Filmische Gestaltung in unterschiedlichen Interaktionsspielräumen

Offensichtlich besteht eine Korrelation zwischen der Planbarkeit und dem Interaktionsspielraum, der den Spielern offensteht. Je kleiner der Interaktionsspielraum ausfällt, desto einfacher lässt sich im Computerspiel der konkrete Ablauf einer Szene planen. Die folgenden Ausführungen schreiten etablierte Strukturkomponenten mit jeweils steigenden Interaktionsspielräumen ab und beleuchten die Möglichkeiten, Grenzen und gängigen Praktiken einer kinematographisch orientierten Szenengestaltung im Rahmen des Computerspiels.

Cutscenes Die größte strukturelle Ähnlichkeit zu Filmszenen ist in den Zwischensequenzen (engl. Cutscenes) des Spiels zu finden. So übernimmt bereits eine der ersten Zwischensequenzen des Computerspiels überhaupt – nämlich das Intermezzo aus *Pacman* – mit der dargestellten Verfolgungsszene eine Trope des Zeichentrickfilms.¹⁰ Zwischensequenzen sind in ihrer reinsten Ausprägung fixierte audiovisuelle Sequenzen, in denen keine Eingriffsmöglichkeiten von Seiten der Spieler vorgesehen sind. Als kurze Abschnitte innerhalb des Spiels folgen sie von der Mise-en-scène bis zur Montage prinzipiell den Strukturprinzipien des Films. Durch die Einbettung in ein interaktives Medium können sie jedoch nicht losgesprochen werden von den grundsätzlichen Anlagen der interaktiven Situation. Als Abschnitte des interaktiven Mediums besitzen auch sie einen Interaktionsmodus: das Warten. Cutscenes unterbrechen den Spielablauf. Ob und wie sich die Spieler mit den audiovisuellen Versatzstücken auseinandersetzen, bleibt ihnen überlassen. Sie können als Belohnungen oder willkommene Pausen geschätzt, oder als störende Unterbrechungen des Spielgeschehens missachtet werden. Mittlerweile gehört es zum Standard, Zwischensequenzen überspringen zu können.

Zwischensequenzen unterscheiden sich voneinander durch das eingesetzte Material. So wurden sie in der Frühzeit des Computerspiels auf Basis der Spielgrafik erstellt und mit Hilfe der Soundchips akustisch untermalt. Obwohl es sich um audiovisuelle Sequenzen handelte, ließ der hohe Abstraktionsgrad des Materials die Zwischensequenzen kaum mehr als Karikaturen kinematographischer Inszenierungen erscheinen.

¹⁰(Howells, 2002, 111)

Nachdem mit dem Vertrieb von Computerspielen auf CD-Rom die Speicherkapazitäten für Filmaufnahmen ausreichten, integrierten ab Mitte der 1990er Jahre zahlreiche Spiele wie *Wing Commander III* (1994), *Battle Isle III* (1995) und *Command & Conquer - Der Tiberiumkonflikt* (1995) digitalisiertes Filmmaterial. Die kurzen Filmabschnitte standen in dieser Zeit jedoch in starkem Kontrast zu der Darstellung des Spiels, die auf berechneter Computergrafik und repetitiven Samples basierte. Heute kommen überwiegend animierte Filme sowie gerenderte oder auf Basis der Game-Engine erstellte Computeranimationen zum Einsatz. Gerenderte Szenen bieten in der Regel eine höhere Grafikqualität und erlauben damit eine spektakulärere Inszenierung, während Cut-Scenes auf Basis der Game-Engine nahtlose Übergänge in spielbare Abschnitte ermöglichen. Einige Entwicklungsstudios setzen Zwischensequenzen gezielt als Markenzeichen ein. So erstellt Blizzard Entertainment für seine Spiele aufwändige gerenderte Cut-Scenes, die schon vor der Veröffentlichung der jeweiligen Titel als Teaser erscheinen, während in der *Command & Conquer*-Reihe trashige Filmaufnahmen einen festen Platz einnehmen. Selbst Filmmaterial dieser Art verstößt in einigen Fällen gegen die Konventionen des Films. So blicken die Schauspieler während des Missionsbriefings in *Command & Conquer* direkt in die Kamera und adressieren die Spieler.

Die klassische Form der Zwischensequenz, die auf jegliche Einflußnahme der Spieler verzichtet, ist in den vergangenen Jahren aus der Mode geraten. Zwischensequenzen sind immer seltener nur lineare audiovisuelle Sequenzen. Da sie zunehmend mit Hilfe von Game-Engines realisiert werden, halten zusätzliche Einflußmöglichkeiten Einzug. So lässt sich in vielen Zwischensequenzen der Bildschirmausschnitt leicht verschieben, während das Rütteln des Controllers markante Momente begleitet. Die Berechnung zur Laufzeit ermöglicht darüber hinaus auch eine stärkere Variantenbildung, ohne dass ein direkter Einfluss während der Sequenz stattfindet. In *The Witcher 2* zeigen die Cut-Scenes z. B. den Hexer Geralt mit denjenigen Ausrüstungsgegenständen, die ihm die Spieler zuvor zugewiesen hatten. Die Spieler bestimmen damit im Vorhinein die Kleidung des Protagonisten. Zwischensequenzen berücksichtigen auf diese Weise variable Aspekte des Spielverlaufs und zeigen Dinge, die auf Entscheidungen der Spieler zurückzuführen sind.

Quick Time Events Ein weiterer Strukturabschnitt, der über einen ähnlich eingeschränkten Interaktionsspielraum verfügt wie Zwischense-

quenzen, sind Quick Time Events. Dabei handelt es sich um Filmsequenzen mit mehreren Abzweigungen, an denen die Spieler im richtigen Moment die auf dem Bildschirm angezeigten Tasten drücken müssen. In der Regel gibt es nur einen erfolgreichen Ausgang der Sequenz und mehrere Möglichkeiten, zu scheitern und neu beginnen zu müssen. Für die Bewältigung einer Quick Time Sequenz ist in erster Linie das Timing entscheidend. Im Gegensatz zu der Zwischensequenz und dem Interaktionsmodus des Wartens ist eine hohe Konzentration erforderlich. Quick Time Events können nicht ignoriert oder übersprungen, aber auch nicht als Verschnaufspause oder atemberaubende Inszenierung genossen werden. Die Gestaltung folgt ebenfalls den Strukturprinzipien des Films, da die einzelnen Abschnitte der Sequenz als audiovisuelle Versatzstücke vorliegen. Während Cut-Scenes vom Action-Adventure bis zum Strategiespiel in diversen Genres vorkommen, treten Quick Time Events ebenso wie die nachfolgend beschriebenen Script-Events überwiegend in Computerspielen auf, die eine durchgehend filmisch orientierte Inszenierung anstreben.

Script-Events Sowohl Zwischensequenzen als auch Quick Time Events sind als eigenständige Strukturabschnitte vom eigentlichen Spiel und dessen Interaktionsfreiheit abgekoppelt. Um auch in spielbaren Abschnitten mit größeren Handlungsspielräumen filmische Szenen nahtlos zu integrieren, verwenden die Entwickler Script-Ereignisse. Sobald alle Voraussetzungen für das Abspielen des Scripts erfüllt sind, wird es ausgeführt. Als Triggerpunkte dienen unsichtbare Schalter, die umgelegt werden, wenn die Spieler bestimmte Bereiche der Spielwelt betreten oder an ausgewählten Orten bestimmte Aktionen ausführen. In *Batman Arkham Asylum* starten viele der nur wenige Sekunden andauernden Skriptsequenzen durch die ort- und kontextbezogene Verwendung der Benutzen-Taste. Für einen kurzen Moment verlieren die Spieler jegliche Kontrolle über ihren Avatar und erlangen sie nach dem Ablauf der Sequenz unmittelbar zurück. Die dargestellten Szenen sind eine Art spezielles, nicht repetitives Feedback auf die Aktionen der Spieler. Analog dazu fungieren gescriptete Kamerafahrten, die beim Betreten eines neuen Raumabschnitts den zu bewältigenden Weg zeigen, als filmisch inszenierte Suggestion.

Abgesehen von Scriptsequenzen, welche als audiovisuelle Abschnitte vorliegen und die Spieler ihres Handlungsfreiraums kurzzeitig berauben, kommen auch Scriptsequenzen zum Einsatz, in denen die Eingriffs-

möglichkeiten durch die Spieler erhalten bleiben. Hierzu gehören z. B. einstürzende Wände, platzende Rohre, plötzlich anspringende oder versagende Maschinen, sowie das Aufwachen totgeglaubter Figuren. In all diesen überraschend eintretenden Ereignissen, die der Umgebung des Avatars entspringen, erfüllen die ihnen zugeordneten Soundeffekte eine wichtige Warnungsfunktion. Welchen Ausschnitt der virtuellen Welt die Spieler in den Blick genommen haben, während das Ereignis eintritt, ist meist unklar. Die akustische Gestaltung des Ereignisses weist die Spieler möglicherweise erst auf die gerade geschehende Veränderung hin und suggeriert ihnen, ihre Aufmerksamkeit darauf zu richten. Auch hier liegt die hohe Bedeutung des akustischen Signals in dem eingeschränkten Sehvermögen und dem deutlich weiterreichenden Hörvermögen in dreidimensionalen Räumen begründet. Gescriptete Ereignisse können zwar übersehen, aber nicht überhört werden.

Script-Sequenzen, die auf Umweltveränderungen der virtuellen Welt zurück zu führen sind, erfüllen entweder rein narrative Funktionen als dramaturgisches Mittel oder bereichern zusätzlich das Spiel mit neuen Herausforderungen. Sie können zustandsvariant oder -invariant sein. Wenn z. B. ein Rohr platzt und in einer Animation plätschernden Wassers mündet, ist die Sequenz zustandsinvariant. Sofern das geplatzte Rohr jedoch einen reißenden Strom verursacht, vor dem die Spieler ihre Avatare retten müssen, ist die Sequenz zustandsvariant.

Script-Events leiten kinematographisch inszenierbare Strukturabschnitte ein, die stärker im Dienst der Erzählung stehen, als dass sie dem Spiel als regelbasiertem System zu Gute kommen. Sie verlieren ihren Reiz, sobald sie verstärkt der Repetition ausgesetzt sind. Außerplanmäßige Ereignisse, die beim ersten Spieldurchlauf für Überraschung sorgten, sind beim zweiten Spieldurchlauf bekannt. Script-Events erhöhen nicht den Wiederspielwert. In Spielpassagen, die den Spielern mehrere Anläufe abverlangen, wirken sie im Gegenteil eher störend.

Szenische Inszenierung von Spielabschnitten

In Abschnitten, die den Spielern die Kameraführung und Positionierung ihres Avatars gewähren, ist eine minutiöse Planung des in einer Szene Hörbaren und Sichtbaren nicht möglich. Die Szenengestaltung besteht deshalb zu einem großen Teil darin, den Rahmen der Szene festzulegen

und einzugrenzen. An Stelle einer finalen Form muß ein Möglichkeitsraum gestaltet werden, der in den meisten Fällen in einer passablen Form mündet. Die folgende Beschreibung des Music Directors¹¹ aus *Left 4 Dead* berührt genau den Kern dieses Designs der Möglichkeiten, das auch die Szenengestaltung als Ganzes durchdringt.

A lot of overreaching dynamic music systems go to great lengths to organize and control a very expressive art form, often to the point of making the results perfectly 'controlled' but also perfectly boring. Our Music Director aims for 'planned serendipity.' By designing the music and rule sets to increase the probability of beautiful happenstance, and to minimize the probability of inappropriate mistakes, we end up with the highest percentage of musical events working as planned, a nice mid percentage of acceptably artful mistakes, and very few actual poor moments. If you over-design the music and rule sets, there are no surprises; and without surprises, listeners are quickly bored. Ironically, by keeping things simple, the music seems planned; greater complexity just leads to greater randomness and many more poor moments.¹²

Die Szenengestaltung besteht entsprechend nicht nur aus Montage und Mise en scène, sondern auch aus Stochastik. Die Wahrscheinlichkeiten, zu denen gewünschte, ästhetisch akzeptable oder ungewünschte Ausgänge entstehen, werden sorgfältig gegeneinander abgewogen. Ungewißheiten entspringen einerseits den implementierten Zufallsoperationen und andererseits den Eingaben der Spieler. Um den Möglichkeitsraum der Eingaben innerhalb abgesteckter Grenzen vorhersehbar zu machen, sind mehrere Techniken in Gebrauch. Sie alle leisten eine Einschränkung des Interaktionsspielraums. Die folgend beschriebenen Inszenierungsstrategien treten dabei selten isoliert, sondern häufig in Kombination miteinander auf.

Untersagte Aktionen Zu den gebräuchlichsten Einschränkungen des Interaktionsspielraums gehört das dauerhafte Verbot oder der tempo-

¹¹Der *Music Director* ist eine Komponente des *AI Directors*, dessen algorithmische Inszenierungsfunktionen Valve als „Structured Unpredictability“ bezeichnet (Booth, 2009).

¹²Mike Morasky, *Left 4 Dead* Developers Commentary

räre Verlust durchführbarer Aktionen. Wenn Aktionen dauerhaft untersagt bleiben, werden sie von den Spielern in der Regel gar nicht als Einschränkungen wahrgenommen. Ein Beispiel dafür ist die Aktion des Kletterns in *Left 4 Dead*. Während Zombie-NPCs über Zäune steigen und an Hauswänden hinaufklettern, sind die von den Spielern gesteuerten Überlebenden dazu nicht in der Lage. Sie stoßen immer wieder auf unüberwindbare Hindernisse, die sie anderweitig umgehen müssen. Durch diese Einschränkung sind die einschlagbaren Wege der Überlebenden bekannt, was die in Echtzeit vollzogene Inszenierung stark begünstigt. So lässt sich mit Hilfe dieses Wissens z. B. berechnen, wo sinnvolle Spawn-Punkte für Horden oder Spezial-Infizierte liegen.

Das temporäre Deaktivieren von Aktionen ist jeweils an bestimmte Bedingungen geknüpft. In *Rage* ist es z. B. nicht möglich, in Siedlungen auf NPCs zu feuern. Sobald die Mündung auf einen NPC zeigt, senkt der Avatar automatisch die Waffe. Größere temporäre Einschnitte in den Aktionsspielraum sind in inszenierten Spielabschnitten wie Träumen und Rückblenden zu finden. Als einflussreiche Sequenzen dieser Art, die viele Computerspiele inspirierten, gelten die spielbaren Albträume in *Max Payne 1*. Im Gegensatz zum Rest des Spiels trägt Max Payne in den Albträumen keine Waffen. Die Spieler können ihren Avatar lediglich durch die langen Gänge eines Labyrinths und die daran anschließenden Räume seines Wohnhauses bewegen. Auch die im restlichen Spiel geltenden physikalischen Gesetze sind leicht modifiziert. Die Spielfigur läuft langsamer und schwebt nach einem Sprung nur behäbig wieder in Richtung Boden. Wiederholte Schreie einer Frau und eines Kindes begleiten die bruchstückhaften, surrealen Abschnitte, durch die die Spielfigur schreiten muß. Die akustische Gestaltung ist ebenso verworren und verstörend wie das Leveldesign des Traums und die veränderten physikalischen Gesetze während der Bewegung des Avatars. Die Einschränkung des Aktionsspielraums überträgt die Hilflosigkeit des im Traum gefangenen Max Payne direkt auf die Spieler, indem ihnen ihre mächtigste Aktion – das Feuern – untersagt bleibt.

Max Payne gilt als eines der ersten Computerspiele, das mit solchen Techniken auf filmische Zwischensequenzen verzichtete und dafür spielbare Abschnitte mit den Möglichkeiten des Mediums inszenierte. Das Besondere an den Sequenzen ist, dass die Spieler – wenn auch eingeschränkt – die Kontrolle über den Avatar behalten. Im Gegensatz zu einer Zwischensequenz, welche die Szene vor den Spielern ausbreitet und entwickelt, wird ihnen hier eine exploratorische Neugier abver-

langt. Vor ihnen eröffnet sich ein multimedial gestaltetes Arrangement, das auf unterschiedliche Arten und Weisen durchschritten werden kann. Die Spieler können panisch hindurch eilen, sich verlaufen, oder langsam vortasten und das Arrangement aus einer selbst gewählten Sicht- und Hörperspektive auf sich wirken lassen. Der begehbbare Traum schafft auf diese Weise eine sehr intensive narrative Erfahrung. Es gibt in ihm kein Gewinnen oder Verlieren, sondern nur die erzählerische Progression durch die Erkundung des Raums.

Viele Action-Adventures, die nach *Max Payne 1* erschienen sind, griffen dieses Verfahren auf. In *Batman Arkham Asylum* treten in Reaktion auf Scarecrows Gas in ähnlichem Stil Mischungen aus Träumen und Rückblenden auf. In der zweiten Traumsequenz erstreckt sich vor dem Avatar ein scheinbar endloser Gang. Wenn man ihm lange genug folgt, beginnt es zu regnen. Wie in einem Hörspiel wird die Ermordung von Bruce Waynes Eltern geschildert, während die Spieler Batman den Gang hinunter schreiten lassen. Nach der Ermordung strauchelt Batman, stürzt in einer Skript-Sequenz vor den leblosen Körpern seiner Eltern zu Boden, und steht als kleiner Junge im Schatten der Fledermaus wieder auf. Die Spieler steuern nun den jungen Bruce Wayne, der mit gebücktem Kopf durch die Gasse weiter voran schreitet, während im Tonsegment das Interview des Jungen auf dem Polizeirevier nach der Tat wiedergegeben wird. Auch in dieser spielbaren Inszenierung, ist die Bewegung durch den Raum entscheidend, während weitere Aktionen nicht möglich sind. Das Deaktivieren von Aktionen ist unmittelbar verbunden mit einem gezielten räumlichen Aufbau. Wenn der Avatar bestimmte Segmente des langen Gangs überquert, löst er unsichtbare Schalter aus, die Umgebungsveränderungen wie den Regen herbeiführen und die hörbuchartige Erzählung starten. Die Gestaltung des Raums bietet weitere Ansatzpunkte zur Einschränkung des Interaktionsspielraums.

Einsperren Wenn an einem Ort der virtuellen Spielwelt eine längere Skriptsequenz für die Inszenierung einer Szene eingesetzt wird, muss sichergestellt werden, dass der Avatar diesen Ort währenddessen nicht verlässt. In der zuvor geschilderten Traumsequenz aus *Batman Arkham Asylum* wurde dieses Problem umgangen, indem der endlos lang wirkende Gang keine Abbiegmöglichkeiten bot und die über den Ton vermittelte Erzählung in Batmans Erinnerung verortet wurde. Wenn der Ton allerdings von einer Position der Spielwelt ausgeht oder ortsgebundene grafische Darstellungen hinzutreten, sind Einschränkungen

des Bewegungsspielraums nötig. Typischerweise werden die Avatare innerhalb des betroffenen Gebiets eingesperrt, indem sich Türen schließen oder sonstige Barrieren das Fortkommen verhindern. *Bioshock* greift in dem Level Fort Frolic z. B. auf diese Lösung zurück, um den verrückt gewordenen Künstler Sander Cohen einzuführen. Kurz bevor die Spieler in Fort Frolic die Tauchkugel erreichen, um weiter nach Hephaestus zu reisen, begrüßt sie Sander Cohen mit einer Performance. Die Tauchkugel versinkt im Wasser und über ihr steigt eine Maske mit Hasenohren in die Höhe, während an den Seiten des Raums ein roter Vorhang ausgefahren wird. Aus der Metro-Station wird eine Bühne. Begleitet von Musik und den grazilen Balletteinlagen herabschwebender Figuren meldet sich Cohen über das Radio zu Wort:

Ah, that's better. Atlas, Ryan, Atlas, Ryan, duh duh duh, duh duh duh. Time was you could get something decent on the radio. The artist has a duty, to seduce the ear and delight the spirit, so say goodbye to those two blowhards, and hello to an evening with Sander Cohen!

Die gesamte Szene über bleibt der Aktionsspielraum uneingeschränkt. Die Spieler können sich bewegen, die Blickrichtung verändern, sowie Springen und Schießen. Es gibt jedoch keine Möglichkeit, den Verlauf der Szene zu beeinflussen oder die Begrüßung von Cohen zu verpassen. Die Türen der Metro-Station bleiben während seiner Ansprache geschlossen. Entsprechend hoch ist die Wahrscheinlichkeit, dass sich die Spieler dem Ballett auf der Bühne zuwenden.¹³ Eine Garantie dafür gibt es jedoch nicht. Während die akustische Gestaltung der Szene – Musik und Ansprache – ortsunabhängig ist, da sie dem Empfangsgerät des Avatars zugeordnet wird, lässt sich die grafische Darbietung ignorieren. Akustische Elemente sind in diesem Sinne planbarer für die Szenengestaltung als die grafische Darstellung, da sich ihnen die Spieler schwerer entziehen können. Auch wenn sie nicht ortsunabhängig übertragen werden, wie hier im Falle des Radios, sind sie im Umkreis eines definierbaren Radius wahrnehmbar. Es ist nicht möglich, sie auszublenden, indem ihnen der Rücken zugekehrt wird.

¹³Bei wiederholtem Durchspielen verlieren Skript-Sequenzen an Reiz. In folgendem Walkthrough-Video ignoriert der Spieler die Bühne und wartet vor der geschlossenen Tür, bis Cohens Ansprache abgespielt wurde: <http://www.youtube.com/watch?v=QCYno1mrbDY> (29.10.2012)

Herausfordern Neben untersagten Aktionen und dem Einsperren der Avatare in bestimmten Arealen können spielerische Herausforderungen für die Szenenplanung verwendet werden. Ein typischer Anwendungsfall dieser Strategie ist das Spawnen von Gegnergruppen. Wenn z. B. in *Mirror's Edge* eine gegnerische Übermacht auftaucht, bleibt den Spielern nur die Flucht. Aufgrund des bewußt hoch angesetzten Schwierigkeitsgrads führen alle Versuche, die Übermacht zu bekämpfen, zum Scheitern und Neubeginn des Spielabschnitts. Das Gamedesign gibt damit ein notwendiges Verhalten vor, das von anderen Gestaltungselementen aufgegriffen werden kann. So unterstreicht ein treibender Elektronik-Soundtrack in hohem Tempo diese Fluchtpassagen, in denen von den Spielern ein schnelles Handeln abverlangt wird. Die Flucht bildet damit eine inszenierte Szene, die vor allem dadurch zustande kommt, dass die Spieler dem drohenden Scheitern entgegensteuern.

Bioshock verwendet das Spawnen von Gegnergruppen, um die Begegnung mit Sander Cohen zu inszenieren. Während die Spieler ihm helfen, sein „Meisterwerk“ zu vollenden, wittert er Zweifel.

That's three of four. What's that look? You don't like it, do you? I don't need to be judged by you, by anyone! Screw you! Screw all you fucking doubters! Here's what I say to all of you!

Cohen sendet anschließend über das Radio Tschaikowskis Blumenwalzer aus dem Ballett *Der Nußknacker* und schreit „Fly away, little moth. Fly!“. Aus mehreren Richtungen erscheinen nun Gegner (sog. Splicer). Sie treffen verzögert ein und setzen die Spieler damit kontinuierlich unter Druck, ohne ein unüberwindbares Hindernis darzustellen. Durch dieses spezielle Spawn-Setup sind die Spieler dazu aufgefordert, ständig in Bewegung zu bleiben. Sie weichen den Angriffen der Splicer aus und rotieren die Sichtperspektive, um neu heranrückende Splicer zu orten. Der Ausbruch von Cohen mündet in einem Kampf, der gleichzeitig auch ein inszenierter Tanz ist. Während der Avatar Pirouetten vollführt, schlagen die Splicer Räder und nähern sich mit akrobatischen Sprüngen. Im Mittelpunkt der Inszenierung steht der Blumenwalzer, der den Spielern mit seinem Dreivierteltakt einen Eingaberhythmus suggeriert und ihnen die Tanz-Assoziation nahelegt.

Whenever this part comes, I get so immersed in the music that I forget I am in a leaky shopping promenade. Instead

I am in an elegant ballroom occupied by kings, aristocrats, and their debutantes. Rather than dodging swings from a lead pipe and returning them in kind with a facefull of monkey wrench, I am making sweeping dance motions along with dance partners that change every two to five seconds because the previous ones have all passed out due to my sheer class and dance prowess.¹⁴

Obwohl ähnliche Kampfsequenzen während des Spielablaufs häufig auftreten, sticht die obige Sequenz durch den Einsatz des Blumenwalzers besonders hervor. Sie konstituiert eine Szene, die den Spielern in Erinnerung bleibt. Im Gegensatz zu vielen anderen Kampfsequenzen ist sie narrativ gerechtfertigt durch den Ausbruch des dem Wahnsinn verfallenen Künstlers.

Das kontrastive Zusammenfallen der unschuldig wirkenden¹⁵ Musik mit den von Gewalt durchdrungenen Bildinhalten ist ein etabliertes Ausdrucksmittel des Films. Chion bezeichnet diese Form der Musikverwendung als „anempathetic music“ und schreibt ihr eine intensivierende Wirkung zu.¹⁶ Anstatt das Gezeigte abzuschwächen, verstärkt sie die dargestellte Brutalität. Die Verwendung dieser filmischen Inszenierungsstrategie in *Bioshock* zeigt damit Parallelen zu Filmszenen auf, die mit der gleichen Strategie operieren. Für manche Spieler erzeugen gerade diese erkannten Bezüge neue Spielerfahrungen.

That part at the end with waltz of the flowers felt just like the end of the First Batman movie, where Batman's fighting the joker's thugs to this very tune. So in a way, I felt like Batman during that sequence, effortlessly taking out those splicers to the beat of the piece.¹⁷

User icepic21 erinnert sich an eine markante Szene aus dem Film *Batman*, der Ende der 1980er Jahre unter der Regie von Tim Burton entstand. Auch wenn die genannte Kampfszene zwischen Batman und Jo-

¹⁴User KaNephren21, http://www.youtube.com/all_comments?v=TBtzIy8AQbQ (29.10.2012)

¹⁵In dem Ballett *Der Nußknacker* begleitet der Blumenwalzer eine Traumsequenz. Die meisten Kompanien verarbeiten die Sequenz, indem sie als Blumen oder Blumenfeen auftreten.

¹⁶(Chion u. a., 1994)

¹⁷User icepic21, http://www.youtube.com/all_comments?v=TBtzIy8AQbQ (29.10.2012)

kers Schergen nicht vom Blumenwalzer, sondern dem „Waltz of Death“ von Danny Elfman begleitet wird, sind die Ähnlichkeiten unverkennbar. In beiden Fällen erfasst die Inszenierung den Kampf eines Helden gegen mehrere Schergen eines wahnsinnigen und exzentrischen Schurken. In beiden Fällen trägt der Walzer zur Charakterisierung des Antagonisten bei. Die Unterschiede ergeben sich aus der interaktiven Situation des Spiels. icepic21 sieht nicht nur die Parallelen beider Szenen. Er fühlt sich in der Rolle von Batman und antizipiert den Rhythmus des Walzers, indem er die von dem Spiel gestellte Herausforderung meistert.

5.3 Gesprochene Sprache

In traditionellen Medien wie dem Film gilt gesprochene Sprache als wichtigstes erzählerisches Mittel, das innerhalb der Tonspur eine herausragende Stellung einnimmt. Chion sieht den Grund dafür in der menschlichen Konstitution:

Sound in film is voco- and verbocentric, above all, because human beings in their habitual behaviour are as well. When in any given sound environment you hear voices, those voices capture and focus your attention before any other sound (wind blowing, music, traffic).¹⁸

Aus dem gleichen Grund begegnen wir auch der Sprache im Computerspiel mit einer hohen Aufmerksamkeit. Die Rezeption von Sprache läuft in beiden Medien ähnlich ab, ebenso wie die Produktion. Sowohl das Casting passender Sprecher, als auch die Vertonung geschriebener Textvorlagen folgt grundsätzlich den gleichen Richtlinien. Im Gegensatz zum Tonfilm ist gesprochene Sprache für das Computerspiel jedoch kein konstitutiver Bestandteil. Viele Computerspiele enthalten keine oder nur sehr geringe Anteile gesprochener Sprache. Darüber hinaus dient sie im Kontext des Spiels nicht nur der Erzählung, sondern übernimmt eine Reihe weiterer Funktionen. Bevor im nächsten Abschnitt untersucht wird, wie Computerspiele die Voice Over Technik des Films adaptieren, sollen deshalb die wichtigsten Funktionen von Sprache im Computerspiel abgeschrieben werden.

¹⁸(Chion u. a., 1994, 6)

Funktionen von Sprache im Computerspiel

Sobald Sprache in Interaktionszusammenhängen auftritt, ist sie mehr als ein Bestandteil der Erzählung. Typische Funktionen bestehen darin, dass sie Anweisungen transportiert, über den Status von Spielzuständen informiert oder das Spielgeschehen kommentiert. Dabei treten diese Funktionen selten isoliert auf, sondern sind in Mischformen zu beobachten.

Erzählende Sprache Erzählende Sprache steht im Dienst der Erzählung und entspricht funktional der Sprachverwendung des Films. Sie tritt hauptsächlich in Dialogen zwischen Avataren und NPCs, oder in Form eines außerhalb des Spielgeschehens stehenden Erzählers auf. Gemessen an der Fülle von Spieltypen und -genres gibt es wenige Computerspiele, die überhaupt Dialoge außerhalb von Cut-Scenes enthalten. Rollenspiele und Adventures sind die einzigen Vertreter, in denen sie zum Standard-Repertoire gehören. Im Gegensatz zum Film ist die Quantität der aufgenommenen Sprachinhalte in diesen Genres jedoch deutlich höher. Die Vertonung geschriebenen Textmaterials gehört dort zu den aufwändigsten Bereichen der Audioproduktion. In Adventures und Rollenspielen, die über eine große Menge von Dialogen verfügen, erreichen die fertigen Aufnahmen mitunter eine Gesamtlänge von mehreren hundert Stunden. So überschreitet die Sprachausgabe des MMOs *Star Wars - The Old Republic* nach Aussage der Entwickler den Umfang von 50 *Star Wars* Romanen.¹⁹ Die großen Mengen aufgenommener Sprache sind primär den Bemühungen um nichtlineare Erzählungen und variable Spielabläufe geschuldet. Durch die Wahl verschiedener Dialogoptionen eröffnet sich für die Spieler die Möglichkeit, unterschiedliche Handlungsverläufe zu erleben. Für die Entwickler resultiert die Bereitstellung dieser Möglichkeit im Worst Case in einem exponentiell steigenden Aufwand im Vergleich zu einer linear erzählten Geschichte. Die schiere Menge an Daten, die bei der Vertonung von Dialogen anfallen, verdeutlichen nochmals die auf Seite 78 angesprochene Abhängigkeit von gesampeltem Sound an die Speicherkapazitäten des Computers. So war die Entwicklung dieser Klangkategorie eng mit der Entwicklung von Kompressionsverfahren und neuen Speichermedien wie der CD-ROM und der DVD verbunden.

¹⁹<http://www.vg247.com/2010/06/20/bioware-swtors-voiceover-content-larger-than-50-star-wars-novels/> (19.06.2012)

Im Mix des Films liegt Sprache an vorderster Stelle. In mehrkanaligen Lautsprecher-Setups ist sie dem Center-Speaker zugeordnet, um eine möglichst gute Sprachverständlichkeit zu gewährleisten. Im Computerspiel, in dem die Spieler die Sicht- und Hörperspektive bestimmen, ist diese Dominanz im Mix nicht gewährleistet. Wenn Spieler in dreidimensionalen virtuellen Räumen den Ausführungen einer Spielfigur lauschen, so hören sie diese aus der Richtung, in der diese Spielfigur in Relation zu der gewählten Perspektive steht. Kehrt der Avatar der sprechenden Figur den Rücken zu, erschallt ihre Stimme aus den hinteren Lautsprechern des mehrkanaligen Setups. Die filmischen Konventionen mehrkanaligen Sounds lassen sich nicht auf das Computerspiel übertragen.

Erzählende Sprache erschöpft sich im Computerspiel selten in ihren narrativen Verweisen. Sie ist darüber hinaus auf unterschiedliche Weise in die Interaktion eingebunden. Eine Gelenkstelle, an der sich die narrativen Inhalte mit spielspezifischen Zielen besonders deutlich verkreuzen, sind *Quests*.²⁰ Sie treten entweder in Dialogen der Avatare mit NPCs in Erscheinung oder werden durch Missionsbriefings an die Spieler herangetragen. Sie verbinden die Erzählung mit dem Spiel, indem sie die Ausführung repetitiver Spielhandlungen wie dem Erbeuten von zehn Wolfsfellen narrativ rechtfertigen und motivieren. Zu diesem Zweck informieren sie die Spieler über die Rahmenbedingungen des Quests und geben ihnen konkrete Aufgaben vor. Gesprochene Sprache dient hier nicht nur der Erzählung, sondern auch der Instruktion und der Information.

Instruierende Sprache Gesprochene Sprache wird in Computerspielen dazu eingesetzt, den Spielern Anweisungen zu erteilen. In geballter Form tritt diese instruktionelle Funktion von Sprache in Tutorials auf, die zentrale Spielprinzipien und die Bedienung des Spiels erklären. Hier werden den Spielern zunächst elementare Spielaktionen wie das Bewegen einer Einheit oder die Steuerung der Kamera abverlangt. Später treten komplexere Instruktionen hinzu, die die Spieler durch die Kombination mehrerer Einzelaktionen bewältigen. Instruktionelle Sprache tritt nicht nur in Tutorials, sondern auch im Hauptspiel auf. Im Gegensatz zu den Tutorial-Stimmen, die den Spielern beibringen, welche

²⁰Eine praxisorientierte Einführung in das Design von Quests gibt Jeff Howard in (Howard, 2008).

Tasten welche Aktionen auslösen, sind Quest-Stimmen innerhalb der fiktiven Welt verortet.

Dass fiktive Spielfiguren den Spielern Befehle erteilen, ist so stark etabliert, dass einige Spiele die gewohnte Gehorsamkeit untergraben. In *Portal* endet beispielsweise das Vertrauen in die KI Glados spätestens dann, wenn sie dem Spieler befiehlt, den Avatar in eine Lava-Grube zu stürzen. Ein ähnlicher Bruch kommt in *Bioshock* vor. Nachdem die Spieler in der Rolle von Jack, dessen Identität zunächst ungeklärt bleibt, Rapture betreten, kontaktiert sie jemand namens Atlas über ein Funkgerät. Er lotst die Spieler durch die zerfallene Unterwasserstadt und legt ihnen eine Reihe von Aufgaben nahe. Im Verlauf des Spiels stellt sich heraus, dass die durch Atlas suggerierten Handlungen der Spieler in der Erzählung als erzwungene Manipulationen zu verstehen sind. Die Phrase „Would you kindly?“ dient als Auslöser, um den spielergesteuerten Avatar Jack gefügig zu machen. Selbst einfache Aktionen wie das Aufheben einer Waffe erscheinen nach diesem narrativen Twist in einer neuen Perspektive.

Would you kindly pick up that short wave radio?

[..] Now, would you kindly find a crowbar or something?“

Ironischerweise haben die Spieler genau wie Jack keine andere Wahl, als den Anweisungen zu folgen. Trotz der großen Bewegungs- und Handlungsfreiräume in den weitläufig gestalteten Levels, sind für die Progression des Spiels jene befohlenen Handlungen zwingend notwendig. Ähnlich wie in *Portal* werden bereits durchgeführte Spielhandlungen rekontextualisiert. Das gebrochene Vertrauen in die instruierenden Stimmen legt den Spielern nahe, über die im Spiel suggerierte Entscheidungsfreiheit nachzudenken.

Informierende Sprache Informierende Sprache transportiert den Status von Spielzuständen oder macht die Spieler auf Ereignisse aufmerksam. Sie geht häufig von Avataren oder Systemstimmen aus. In *Dota 2* meldet sich z. B. eine Stimme zu Wort, sobald Gebäude angegriffen werden oder zerstört worden sind. Wenn die Spieler eine dieser Nachrichten wie „Your top tower is under attack.“ hören, können sie zeitig reagieren und dem angegriffenen Turm zu Hilfe eilen. Informierende Sprache ist in hohem Maße suggestiv. Sie legt konkrete Handlungen nahe und

bietet damit den Spielern in bestimmten Situationen eine Hilfestellung an. So nutzen Game Designer in Adventures bei Rätseln z.B. kurze Sprachsamples, um Hinweise über die Lösung zu geben. Häufig ist der Abspielbefehl des Hinweis-Samples an einen Timer gebunden, der an dem Ort des Rätsels die Dauer der Überlegungen der Spieler mißt und beim Überschreiten einer festgelegten Schwelle den Abspielvorgang auslöst.

In vielen Action-Spielen findet sich eine ähnliche monologartige Einbindung von Sprache. Sie folgt auch dort streng der Interaktionslogik von Feedback und Suggestion im Rahmen des Spielsystems. Einerseits gibt sie den Spielern Rückmeldungen über durchgeführte Aktionen oder den Status bestimmter Spielzustände. Andererseits verweist sie auf Handlungen. Wenn z.B. der Mönch in *Diablo 3* sagt „My strength fails“, dann erfahren die Spieler dadurch, dass die Lebensenergie der Spielfigur auf einen beunruhigend tiefen Wert gesunken ist. Sie werden zur Abhilfe einen Heiltrank benutzen oder versuchen, die Spielfigur aus dem Schlachtgetümmel zurück zu ziehen. Sagt der Mönch hingegen „I need more spirit“, dann war die Geisteskraft als zweite Ressource auf einem zu niedrigem Level, um eine bestimmte Aktion durchzuführen. Das System teilt den Spielern auf diese Weise den Fehlschlag einer versuchten Aktion mit und suggeriert ihnen damit, die folgenden Aktionen möglichst auf die Gewinnung neuer Geisteskraft auszurichten. Sprache ist in solchen Interaktionszusammenhängen ein wichtiger Indikator für den Status des Spielgeschehens. Der Mönch fungiert gewissermaßen als Sprachrohr des Spielsystems, das den Spielern wichtige Informationen zur Bewältigung kritischer Situationen zukommen lässt.

Bei gesprochener Sprache dieser Art tritt die gedoppelte Verweiskfunktion des akustischen Zeichens – als narratives Element und als Darstellung eines Spielzustands oder -ereignisses – sehr deutlich hervor. Die Sprüche der Spielfigur fügen sich nahtlos in das narrative Geschehen, rücken durch ihre Funktionalisierung im Rahmen des Spiels jedoch in eine Zwischenebene. Durch ihre Kopplung an Zustände und Ereignisse sind sie der repetitiven Struktur des Spiels ausgesetzt und fungieren dadurch stärker als vertraute Indikatoren wiederkehrender Spielvorgänge, denn als vorantreibende Elemente einer Erzählung. Spieler sind sich der Sonderfunktion dieser Sprachkategorie bewußt und setzen sich damit u. a. auch in Fan-Art auseinander. So wird in einem Cartoon zum Spiel *Left 4 Dead* ein Charakter, dem ausschließlich die funktionalisier-

ten Sprachfetzen wie „Reloading“ und „I found some ammo“ in den Mund gelegt werden, als psychisch derangiert dargestellt.²¹

Kommentierende Sprache Die kommentierende Sprachfunktion in Computerspielen ist der Berichterstattung einer Sportübertragung nicht unähnlich. In Sportsimulationen wie *Fifa* und *Pro Evolution Soccer* begleiten sogar prominente Sportmoderatoren das Spielgeschehen mit einer Reihe bekannter Phrasen. Der entscheidende Unterschied besteht darin, dass die Spieler das Geschehen nicht aus der Distanz beobachten, sondern aktiv daran beteiligt sind. Die Kommentatoren beurteilen ihre Leistungen, sie bejubeln Erfolge und spotten über Fehlschläge. Kommentare verleihen den möglichen Spieldausgängen dadurch eine höhere Intensität. Sie schaffen eine stärkere Genugtuung, wenn Spielaktionen erfolgreich verlaufen und spornen die Spieler zu besseren Leistungen an, wenn sie dem Spott der Kommentatoren entgehen möchten.

Kommentierte Sprache tritt nicht nur in Sportsimulationen auf, in denen sie sich an anderen medialen Vorbildern wie Fernsehübertragungen von Sportveranstaltungen orientiert, sondern auch in anderen Genres wie dem First Person Shooter. So verwendet die *Unreal Tournament* Reihe z. B. Sprecherinnen und Sprecher²², die erfolgreiche Aktionen sprachlich hervorheben. Erledigen die Spieler eine gegnerische Figur mit einem gezielten Schuss, vermerkt die Stimme dies als „Headshot“. Ein Großteil der Kommentare wie „Double Kill“, „Triple Kill“, „Monster Kill“ und „Rampage“²³ bezieht sich darauf, wie viel Schaden ein Spieler innerhalb eines Lebens oder bestimmter Zeitschranken anrichtet. Sie sind damit eine Art Bewertungskriterium für besondere Leistungen der Spieler.

²¹<http://forums.gametrailers.com/viewtopic.php?f=23&t=556073> (16.12.2012)

²²Je nach Version können sich die Spieler eine Stimme unter mehreren aussuchen, die ihre Aktionen kommentiert.

²³Die *Unreal Tournament*-Kommentierung ist mittlerweile so etwas wie eine Blaupause für kompetitive Multiplayer Titel, in denen die Anzahl genommener und verlorener Leben als Indikator für die Leistung der Spieler herangezogen wird. So adaptierte z. B. auch *Dota 2* die „Killing Spree Announcer“.

Voice Over

Die Voice Over-Technik, bei der außerhalb der Szene stehende Sprecher das Dargestellte kommentieren oder als Erzähler in Erscheinung treten, besitzt im Film große Popularität. Innerhalb der vergangenen Jahre adaptierte das Computerspiel diese Technik auf unterschiedliche Arten und Weisen. Die erzählende Stimme ist dabei in den meisten Fällen nicht den Avataren der Spieler zugeordnet, sondern entweder einer unbekanntem Erzählinstanz oder einem NPC. Während die Voice Over-Technik in Cutscenes genau so wie im Film eingesetzt werden kann, sind in spielbaren Abschnitten neue Konzepte nötig, welche die Interaktion berücksichtigen. Eine Integrationsform, die die Ungewissheiten der interaktiven Situation geschickt umgeht, ist die Einbindung der Voice Over-Stimme zu Beginn eines Spielabschnitts oder in einer Spielsituation, die zum Warten verpflichtet. In dem Märchen Jump'n'Run *Trine* setzt die in der interaktionslosen Eingangssequenz etablierte Stimme beim Start des Levels die Erzählung fort. Sie verbindet auf diese Weise die Zwischensequenz mit dem spielbaren Abschnitt und sorgt damit für eine erzählerische Kontinuität.

Eine ähnliche, etwas erweiterte Strategie kommt in *Prince of Persia - The Two Thrones* zur Anwendung. Die einführende Cutscene zeigt, wie der Prinz und seine Begleiterin Kaileena auf das belagerte Babylon zu steuern. Nachdem ihr Schiff unter Beschuss gerät, verlieren sie sich und stranden an unterschiedlichen Abschnitten der Küste. Die Belagerer nehmen Kaileena in ihre Gewalt, während der Prinz zu ihrer Rettung eilt. An diesem Punkt endet die Cutscene und die Spieler übernehmen die Steuerung des Prinzen. Das Gameplay besteht aus akrobatischen Kletterpartien und gelegentlichen Auseinandersetzungen mit Gruppen von Gegnern. Innerhalb dieser spielbaren Abschnitte erzählen beide Stimmen, die zuvor in der Cutscene etabliert wurden, die Geschichte weiter. Der Prinz kommentiert in Monologen die Situation in der zerfallenen Stadt, während die Stimme der entführten Kaileena als übergeordnete Erzählinstanz seine Motivation schildert. Die Aussagen von beiden Charakteren sind wie im folgenden Beispiel so allgemein gehalten, dass sie nicht mit dem Gameplay in Konflikt geraten.

Kaileena: You should know that it was not love that drove him, but duty. I was his responsibility. He had made a pro-

mise, a promise that was now broken and undone. As with all mistakes he had made, the Prince meant to fix this.

Nicht nur der Prinz steht in der Pflicht, sondern auch die Spieler. Kai-leenas Formulierung lässt sich als direkten Apell an die Spieler vor den Bildschirmen interpretieren, die trotz ihrer kontinuierlichen Eingabefehler beharrlich auf das Weiterkommen pochen. Begünstigt werden die Voice Over-Passagen des Spiels dadurch, dass das Leveldesign jeweils nur einen möglichen Weg vorgibt. Analog dazu besitzt auch die Geschichte keine Abzweigung. Durch die Aneinanderreihung aller aufeinander folgender Sprachsamples lässt sich ein vollständiges Skript erstellen.²⁴ Die einzelnen Aktionen wie Springen, an einer Wand entlanglaufen, oder einen Schwerthieb ausführen, bleiben in der Erzählung außen vor. Sie wiederholen sich ständig und können je nach dem Timing unterschiedliche Ausgänge nehmen. *Prince of Persia - The Two Thrones* vermeidet die erzählerische Antizipation dieser Handlungsfreiräume des Spiels, indem die Erzähler auf einer allgemeineren Ebene davon losgelöst bleiben. Die zeitliche Entfaltung der erzählten Geschichte richtet sich alleine nach dem Tempo, in dem die Spieler das Spiel durchlaufen. Bei jedem der im Raum platzierten Triggerpunkte setzt sich die Geschichte fort. Ihre Progression impliziert gleichzeitig auch eine Progression des Spiels und fungiert damit als Bestätigung einer erfolgreich bewältigten Spielpassage.

Mit *Bastion* erschien 2011 ein Action-Adventure, bei dem die Voice Over-Technik stärker an die Entscheidungen der Spieler gekoppelt ist. Ähnlich wie bei *Prince of Persia - The Two Thrones* wird der Erzähler im Lauf der Handlung als NPC eingeführt, der die Taten des spielergesteuerten Avatars kommentiert.

Während einige Kommentare des Erzählers die fiktive Welt im Allgemeinen beschreiben, dringen andere bis hin zu konkreten Spielaktionen und -ausgängen vor. Das Aufnehmen eines Gegenstands findet ebenso Erwähnung wie der Erfolg oder die Niederlage in einem Gefecht mit gegnerischen Einheiten. Es gibt kaum Aktionen der Spieler, die der Erzähler nicht in die Geschichte einbaut. Selbst wenn die Spielfigur in einen Abgrund stürzt oder mehrere in der Spielwelt platzierte Gegenstände zertrümmert, hat die Voice Over-Stimme einen passenden Spruch pa-

²⁴User „jrdemi“ hat sich tatsächlich die Mühe gemacht, siehe <http://www.gamefaqs.com/ps2/926985-prince-of-persia-the-two-thrones/faqs/47500> (13.05.2013)

rat. Es ist erstaunlich, dass diese stimmlich getragene Erzählstrategie im Computerspiel nicht weiter verbreitet ist. Bereits das 2001 erschienene *Max Payne* machte auf ähnliche Weise Gebrauch von der menschlichen Stimme. Wenn der spielergesteuerte Protagonist Max Payne z. B. Schmerzmittel einsammelt, vermerkt die Stimme seiner Gedanken „The pills would ease the pain“. *Bastion* ist insofern eine Erweiterung dieses Prinzips, als es mehrere Varianten bereit stellt. Wenn die Spieler in *Max Payne* die Pillen liegen lassen, bleibt die Stimme stumm. Es gibt nur ein „entweder“ (if), aber kein „oder“ (else). Der Erzähler in *Bastion* hingegen erwähnt Versäumnisse und Niederlagen ebenso wie Erfolge und Siege. Die Implementierung der Voice Over-Stimme berücksichtigt die durch das Spielsystem gegebenen Abzweigungen. Der Aufwand bei der Umsetzung solch eines adaptiven Erzählers ist deutlich höher als bei linearen Erzählweisen. Neben den erforderlichen Sprachaufnahmen mehrerer alternativer Kommentierungen ist eine Metastruktur nötig, welche im Programm die Abspielprozesse den jeweiligen Ausgängen von Spielsituationen zuordnet. Die Voice Over-Stimme ist damit nicht nur an Trigger-Punkte gebunden, sondern verwendet eine Auswahl von Spielausgängen als Trigger-Entscheidungen.

Einen ähnlichen Ansatz wie *Bastion* verfolgt die *Half Life 2* Mod *The Stanley Parable*. Auch hier dringt der Erzähler bis zu den Aktionen und Entscheidungen der Spieler durch. Der Unterschied liegt jedoch darin, dass der Erzähler in seinen Schilderungen die Entscheidungen der vom Spieler gesteuerten Figur Stanley vorwegnimmt und damit Konflikte erzeugt. Sobald die Spieler gleich zu Beginn der Mod einen Raum mit zwei geschlossenen Türen betreten, bemerkt die Stimme:

When Stanley came to a set of two open doors, he entered the door on his left.

Durch die provokante Vorwegnahme der Entscheidung sehen sich die Spieler vor die Wahl gestellt, ob sie der Stimme gehorchen oder ihre Anweisungen missachten. Alle weiteren Entscheidungen werden auf ähnliche Weise aufgeladen und führen je nach Gehorsam oder Widersetzung zu unterschiedlichen Ergebnissen. In einem der daraus entstehenden Erzählstränge gelangt Stanley in ein Büro mit einer verschlossenen Tür, die nur durch einen Zahlencode zu öffnen ist. Hier demonstriert der Erzähler seine Allwissenheit, indem er nach dem Hinweis, dass Stanley diesen Code unmöglich wissen kann, die Zahlen auflistet und den Spielern das Weiterkommen ermöglicht.

Die Beispiele zeigen, dass die Adaption der Voice Over-Technik in Computerspielen zwar einen höheren Produktionsaufwand erfordert, dafür aber ein sehr ausdrucksstarkes Mittel darstellt, mit dessen Hilfe bislang wenig erprobte Effekte realisierbar sind. So schrieb ein Kritiker über die Stimme des Erzählers in *Bastion*: „[...] it’s frightening to think how much the human voice has done to distance Bastion from the crowd“.²⁵ Auch hier liegt der Schlüssel für einen effektiven Einsatz der Stimme in ihrer Kopplung an Ereignisse des Spielgeschehens - von groben bis hin zu feingranularen Ereignisrastern und von vergangenen bis hin zu zukünftigen potentiellen Entscheidungen der Spieler.

²⁵<http://www.edge-online.com/reviews/bastion-review> (30.05.2013)

6 Audio-Konfiguration

Der Akt des Konfigurierens gehört zu den zentralen operativen Kulturtechniken unserer Zeit. Je mehr Maschinen uns umgeben, denen wir Aufgaben anvertrauen oder über die wir Inhalte konsumieren, desto häufiger sehen wir uns mit der Einstellung variabler Betriebsparameter konfrontiert. Wir finden Konfigurationen im Bereich der Hardware und der Software sowie im Zusammenspiel zwischen ihnen. Bei der Zusammenstellung einzelner PC-Komponenten für ein Komplettsystem spricht man beispielsweise von einer Hardware-Konfiguration. Auch Programme bieten in der Regel eine Reihe von Anpassungsmöglichkeiten, und je stärker Software in alle Bereiche des Lebens vordringt¹, desto häufiger sehen wir uns mit Software-seitigen Konfigurationen konfrontiert. Sound, Grafik und Text sind davon nicht ausgeschlossen. Als Folge der Digitalisierung sind Inhalte nicht mehr an medienspezifische Charakteristika gebunden², sondern als Datensätze in verschiedene Darstellungen transformierbar. Aus Text wird gesprochene Sprache durch die Möglichkeiten der Sprachsynthese und aus Musik entsteht mit Hilfe von Visualisierungsalgorithmen ein multimediales Ensemble.³ Das Konfigurieren eröffnet uns die Möglichkeit, in diese transformativen Prozesse einzugreifen und das Arrangement zu bestimmen, das uns bezüglich persönlicher Vorlieben und dem jeweiligen Rezeptionskontext am geeignetsten erscheint.

Konfigurative Akte - sei es das Zusammenstellen einer Playlist oder die Wahl individueller Komponenten beim Kauf eines Neuwagens - sind

¹Greenfield spricht in Hinblick auf die ubiquitäre Verbreitung von Software auch von *everyware* (Greenfield, 2006).

²Abgesehen von den spezifischen Charakteristika des Digitalen selbst.

³Zur Zeit bestimmen neue Ein- und Ausgabegeräte den Charakter dieser medialen Transformationen. Inhalte werden nicht nur über stationäre Ausgabegeräte wie Bildschirme und Lautsprecher dargestellt, sondern auch über mobile Geräte wie Smartphones, Uhren und Brillen. Als Eingabegeräte fungieren zunehmend die Dinge um uns herum. Mit dem *MaKey MaKey Kit* (<http://makeymakey.com/>), das zwei Studenten des MIT Media Lab entwickelten, wird z. B. eine Banane zum Button und eine Treppe zum Klavier.

mittlerweile so fest im Leben der digitalen Gesellschaft verankert, dass wir sie häufig nicht als solche wahrnehmen. So konstatiert der Medien-theoretiker Alexander Galloway:

In short, to live today is to know how to use menus.⁴

Konfigurative Akte in Computerspielen seien demnach lediglich die Fußnoten einer größeren Transformation.⁵ In Hinblick auf die Rezeption medialer Inhalte wie Text, Sprache oder Musik, lässt sich der Konfigurierbarkeit von Computerspielen hingegen eine größere Bedeutung beimessen, als lediglich die Begleiterscheinung einer umfassenderen Umwälzung zu sein. Sie nimmt gewissermaßen das vorweg, was andere etablierte Medien wie Radio und TV noch erwartet: ein weitaus höherer Anteil der Mitbestimmung der User in Bezug auf die Art der Darstellung. Welche Konfigurationsmöglichkeiten gibt es also in Computerspielen? Galloway identifiziert zwei grundlegende Varianten:

The first is confined to the area of setup. Setup actions exist in all games. They are the interstitial acts of preference setting, game configuration, meta-analysis of gameplay, loading or saving, selecting one player or two, and so on. [...] Yet there exists a second variant [...] around which many of the most significant games have been designed. These are gamic actions in which the act of configuration itself *is the very site of gameplay*. These are games oriented around understanding and executing specific algorithms.⁶

Beide Varianten berühren auch die akustische Dimension des Computerspiels. So ist das Audio-Setup seit Jahrzehnten ein fester Bestandteil des Optionsmenüs von PC-Spielen, in dem die Spieler den Rahmen des Gameplays ebenso festlegen wie den Detailgrad und die Bestandteile der Darstellung. Akustische Inhalte sind dadurch ebenso anpassbar wie die Schrift auf einem E-Book-Reader, die nicht auf das Blatt gedruckt ist, sondern auf Wunsch in unterschiedlichen Schriftarten und Größen gesetzt wird. Die Trennung von Informationen und ihren Darstellungsweisen ermöglicht es, Inhalte auf mehrere mögliche Arten abzurufen. Die zweite Variante, die Galloway nennt, zeigt sich auf akustischer Ebene in Spielaktionen, die eigens darauf ausgerichtet sind, die virtuellen

⁴(Galloway, 2006, 17)

⁵ebd.

⁶(Galloway, 2006, 13f.)

Klangräume zu beeinflussen. Eine wichtige Rolle spielen in dieser Hinsicht die Simulationen technischer Medien wie dem Radio. An Stelle des Spiels mit Algorithmen besteht die Ingame-Konfiguration des Hörbaren allerdings aus der Auseinandersetzung mit abrufbaren Inhalten und deren Abspielmöglichkeiten.

In der Architektur-Metaphorik entspricht die Konfiguration der Inneneinrichtung. Der Rohbau steht bereits und lässt sich nicht mehr verändern. Über die Inneneinrichtung lassen sich hingegen Anpassungen vornehmen, die den eigenen Vorlieben entsprechen. Dabei gibt es Anpassungen, die vor dem Einzug bzw. dem eigentlichen Spiel durchgeführt werden und Anpassungen, die erst dann vorgenommen werden, wenn sich die Spieler schon in den virtuellen Räumen bewegen - Setup-Konfigurationen und Gameplay-Konfigurationen. Je nach Spiel reicht die Palette der Einrichtungsmöglichkeiten von wenigen vorgefertigten Formaten bis hin zu einer großen Auswahl individuell einstellbarer Parameter. Die folgende Bestandsaufnahme der Audio-Konfiguration beginnt bei dem Phänomen, dass Spieler bestimmte Klangquellen ausblenden können. Nach der Analyse gebräuchlicher Muster solcher Setup-Konfigurationen rücken Gameplay-Konfigurationen wie die Wahl verschiedener Darstellungsmodi und der Umgang mit simulierten technischen Medien in den Fokus.

6.1 Anpassbare Hörerlebnisse

PC-Spiele und auch viele Konsolentitel enthalten Optionsmenüs, in denen sich bestimmte Rahmenbedingungen des Spiels konfigurieren lassen. Im Folgenden werden die sich daraus ergebenden akustischen Konfigurationsmöglichkeiten untersucht und typische Konfigurationsmuster analysiert.

Musik abschalten

„Any way to turn off the music?“, fragt User LGM1979 im Steam-Forum von *Hotline Miami*.⁷ Ein erster Spieler rät LGM1979 dazu, die Audio-

⁷<http://steamcommunity.com/app/219150/discussions/0/864950398842733548/>
(09.03.2013)

Dateien zu verschieben. Das führt allerdings nicht zum gewünschten Ergebnis. Wenn die Musik im Spiel starten sollte, erscheint eine Fehlermeldung. Ein zweiter Spieler präsentiert eine bessere Lösung, nämlich die Musik-Dateien des Spiels durch leere Audio-Dateien zu ersetzen. Digital kodierte Stille an Stelle von Musik - der Austausch funktioniert.

Dialoge wie diese sind innerhalb der Gaming-Community keine Seltenheit. Die Spieler möchten eine Klangkategorie - häufig Musik - ausblenden, während die übrigen Bestandteile der Klangkulisse hörbar bleiben. Sie tauschen sich untereinander aus, um den Output des Spiels an ihre Bedürfnisse anzupassen. Neben dem Wortwechsel von Spieler zu Spieler sind manche Anfragen der Spieler direkt an die Spielehersteller gerichtet - häufig in Form der Bitte um ein neues Feature. So greift User BuckRaven im Forum von *Battlefield Play4Free* z. B. die anscheinend mehrfach geäußerte Bitte um „Advanced Audio-Settings“ auf, derzufolge unterschiedliche Sound-Kategorien wie Musik und Ambient Sound separat regelbar sein sollten.⁸ Doch lange nicht alle Anfragen an die Hersteller sind sachliche gehaltene Feature Requests. Im *Modern Warfare 3* Forum moniert User Zaled:

Why can't we turn off music? A simple question, for such a simple feature to implemt [sic] in options. The one thing i would LOVE to do most is to turn off the music. Why can't the developeres make an option that allows you to turn on/off the in game sound and a SEPERATE option to turn on/off the MUSIC? I do not understand this. [..]⁹

Zaleds Foren-Eintrag generierte über 20.000 Views und mehr als 100 Antworten. Viele Spieler pflichten ihm bei, während andere die Abwesenheit der Option verteidigen, um für alle Spieler möglichst gleiche Spielbedingungen beizubehalten. Unabhängig davon, ob sie für oder wider die Abschaltmöglichkeit optieren, ist die Vehemenz erstaunlich, mit der die Spieler ihren Standpunkt vertreten. Bei einem Film ist es kaum vorstellbar, dass Zuschauer Proteste für die Änderung der Tonspur einlegen. Woher kommt diese Selbstverständlichkeit der Spieler, die akustischen Ausgaben des Spiels selbst bestimmen zu wollen?

⁸<http://battlefield.play4free.com/de/forum/showthread.php?tid=89189>
(09.03.2013)

⁹<http://www.callofduty.com/thread/200384924> (09.03.2013)

Ein Grund für die Proteste ist der, dass *Modern Warfare 3* gegen eine Konvention aktueller Computerspiele verstößt. Tatsächlich bietet die überwiegende Mehrzahl von PC-Spielen die Möglichkeit, Spielmusik im Verhältnis zu anderen Klangkategorien des Spiels leiser zu drehen oder gleich ganz auszuschalten. Die Audio-Menüs enthalten typischerweise mehrere Lautstärkeregler, mit deren Hilfe bestimmte Klangkategorien den eigenen Vorstellungen entsprechend aufeinander abgestimmt werden können. Jedes Spiel wird zwar mit Voreinstellungen, einem empfohlenen Mix der akustischen Szenerie geliefert, deren Nutzung aber der Wahl der Spieler überlassen bleibt. Sie sind im Endeffekt diejenigen, die die Regler am Spielpult kontrollieren und ihren eigenen Mix austarieren.

Vor allem im Segment der PC-Spiele besitzt die Nutzung von Konfigurationsmenüs eine hohe Selbstverständlichkeit. Der erste Schritt der Spieler besteht häufig darin, das Optionsmenü zu suchen und die Einstellungen zu überprüfen. Viele Spiele rufen es bei dem ersten Start des Programms nach der Installation automatisch auf, damit die Einstellungen eingesehen und bestätigt oder gegebenenfalls angepasst werden können. Auf anderen Plattformen wie Spielkonsolen sind direkte Konfigurationsaufforderungen hingegen nicht üblich, da den Entwicklern die Basis-Spezifikationen der Systeme bekannt sind. Trotzdem deutet die Integration von Konfigurationsmenüs in Konsolen-Spielen an, dass es sich nicht nur um ein Randphänomen der PC-Gamer handelt, sondern eine breitflächige Transformation darstellt.

Ein Beispiel für das Audio Menü eines PC-Spiels (*Starcraft II*) zeigt Abbildung 6.1. Es enthält mit den Unterkategorien „Lautstärke“, „Setup“ und „Qualität“ drei typische Kategorien der Audio-Konfiguration. Unter „Setup“ können die Spieler die zu verwendende Soundkarte und das Lautsprecher-Setup auswählen. Die Kategorie „Qualität“ wirkt sich mit der Wahl der Audioqualität und der verwendeten Soundkanäle¹⁰ am stärksten auf die Performance aus. Den größten Block bilden die Lautstärkeeinstellungen. Hier bietet sich den Spielern die Möglichkeit, den Audio-Mix zu manipulieren. So können ganze Kategorien ausgeblendet, oder die Lautstärkeverhältnisse den eigenen Vorlieben angepasst werden. Eine Besonderheit stellt die Alternative bezüglich der Musikwiedergabe dar. So dürfen die Spieler in *Starcraft II* bestimmen,

¹⁰Die Anzahl der Kanäle legt fest, wie viele Sounds maximal gleichzeitig abgespielt werden.



Abbildung 6.1: Audio-Menü aus *Starcraft II*

ob die Musik ihrer Fraktion dauerhaft abgespielt wird oder in zwischenzeitlichen Ruhephasen pausiert.¹¹

Welche Klangkategorien einen Lautstärkereglern erhalten und an die Spieler weitergereicht werden, variiert von Spiel zu Spiel. Einige Computerspiele enthalten Lautstärkereglern für die drei klassischen Kategorien der filmischen Tonspur: Sprache, Musik und Sound. Andere berücksichtigen vollkommen andere Klangkategorien. In Autorennspielen wie *Dirt II* lässt sich die Lautstärke des Motors festlegen, in Sportspielen wie *FIFA Soccer* der Pegel des Stadion-Getöses. Das Rollenspiel *The Elder Scrolls: Oblivion* enthält einen Lautstärkereglern für die Schritte des Avatars und berücksichtigt damit den Umstand des kontinuierlichen Durchwanderns der großen Spielwelt. Andere Titel enthalten Regler für die Ambient-Kulisse oder Klänge der Benutzeroberfläche. Es gibt keine festen Konventionen, welche Klangkategorien über das Menü konfigurierbar sind. Dennoch ist ein klares Muster zu beobachten: Die mischbaren Klangkategorien entsprechen gerade jenen Sound-Schichten, die

¹¹Das Action-RPG *Titan Quest* ermöglicht sogar eine noch feinere Justierung, indem ein Schieberegler die Häufigkeit der Hintergrundmusik festlegt.

im Spielverlauf immer wieder auftreten und spiegeln damit die oberste Hierarchieebene der internen Organisation akustischer Artefakte wider.

Neben Audio-Menüs enthalten Computerspiele weitere Menü-Sektionen wie Steuerung, Gameplay und Grafik.¹² In jeder Sektion befinden sich eine Reihe von Einstellungen, die über Pulldown-Menüs, Checkboxes und Slider den jeweiligen Wünschen angepasst werden können. Analog zu der Option, Musik ausschalten zu können, sind viele Einstellungsmöglichkeiten in so vielen Titeln zu finden, dass sie für die Spieler selbstverständlich sind. Sobald eine weit verbreitete Option fehlt, hagelt es Proteste aus der Community. So kann es sich heute z. B. kein Entwickler leisten, auf eine frei belegbare Tastenbelegung und die mögliche Invertierung der Y-Achse bei der Maussteuerung zu verzichten. Ein Großteil der Spieler ändert das Mapping der Tasten zu den Aktionen der fiktiven Spielwelt auf eine Konfiguration, die sie aus anderen Spielen gewöhnt sind.

Doch die Konfigurationsmöglichkeiten enden nicht mit den Menüs. Wie das Eingangsbeispiel mit der Musik aus *Hotline Miami* gezeigt hat, konfigurieren die Spieler die akustische Dimension auch auf der Systemebene, indem sie Audio-Dateien austauschen oder auch Skripte für die Audio-Wiedergabe modifizieren. Der Anspruch der Spieler an eine für ihre Bedürfnisse stimmige Klangkulisse ist in diesen Fällen so hoch, dass die Steuerungsmöglichkeiten des Audio-Menüs nicht mehr ausreichen. Hier verläuft ein schmaler Grat zwischen Konfigurationen und Modifikationen. Foreneinträge dokumentieren, dass Spieler über Mods selbst kleine Nuancen der akustischen Gestaltung ihrem Gustus anpassen. Für das Spiel *Civilization V* existieren z. B. Threads, die sich mit der gezielten Ausschaltung oder Ersetzung einzelner Sounds auseinandersetzen. Durch die offene, auf XML-Dateien basierende Architektur des Spiels sind solche Modifikationen auch ohne Expertenkenntnisse realisierbar. Ein weiteres Resultat dieses Anspruchs sind Sound-Mods, welche bestimmte Geräusche oder Musikstücke durch andere ersetzen. In vielen Fällen ist die Abschaltung und Ersetzung auch ohne Eingriffe in den Code des Spiels möglich. Ein weiterer Konfigurationsaspekt ist in Free-to-Play Titeln wie *DOTA 2* zu finden. Der Hersteller Valve kreierte verschiedene Kommentatoren des Spielgeschehens und stellt es

¹²Die größte Komplexität geht in der Regel von den Grafikoptionen aus. Sie fallen mittlerweile dermaßen umfangreich aus, dass PC-Spielemagazine ihnen mehrseitige Artikel widmen, die sich mit den Auswirkungen und Hardware-Anforderungen einzelner Schalter auseinandersetzen.

den Spielern frei, auch eigene Kommentatoren beizusteuern.¹³ Zu Beginn einer Partie können die Spieler wählen, welche Kommentatoren das Spiel begleiten. Darüber hinaus können sie gegenseitig ihre Kommentatoren untereinander austauschen.

Das eingangs erwähnte *Battlefield Play4Free* stellt weder Lautstärkemenüs noch leicht modifizierbare XML-Dateien bereit, was den Unmut vieler Spieler erklärt. Ganz ähnlich gestaltet es sich bei dem Titel *Battlefield Bad Company 2*. Hier gibt es ebenfalls keine Lautstärkeregler, sondern ausschließlich die Wahl bestimmter Presets. Abbildung 6.2 zeigt alle verfügbaren Sound-Presets. Abgesehen von der Einstellung „War Tapes“ beziehen sich alle anderen Optionen auf das physikalische Lautsprecher-Setup der Benutzer. Anstatt das Lautstärkeverhältnis verschiedener Audio-Kategorien beeinflussen zu können, informieren die Spieler die „runtime Master Unit“ lediglich über ihre Abspielhardware. Spiele wie *Bad Company 2*, die wenige oder keine Einflussmöglichkeiten auf den Mix gewähren, verfolgen in der Regel eine kinematographisch orientierte Präsentation.¹⁴ Die Abmischungen aller Klangkategorien sind genaustens geplant und ausgepegelt. Weitere Eingriffe in den intendierten Mix sind nicht vorgesehen. Die geringen Einstellungsmöglichkeiten schaffen für alle Spieler ähnliche Hörbedingungen, was besonders für den Multiplayer-Modus relevant ist. So bleibt es den Spielern versagt, durch die Ausblendung oder die Hervorhebung bestimmter Klangkategorien spielerische Vorteile zu erlangen.

Mit dem Schwerpunkt auf der Abspielhardware verweisen Presets auf den Ursprung der Konfigurationsmöglichkeiten in PC-Spielen. Einen wichtigen Ausgangspunkt der heutigen Menüs bildeten Entwicklungen rund um den IBM-PC Mitte der 1980er Jahre. Durch das Reverse Engineering des BIOS Chips durch die Firma Phoenix Technologies, war es auch anderen Herstellern möglich, einen IBM-kompatiblen PC zusammen zu stellen.¹⁵ Im Gegensatz zu Konsolen und den frühen Heimcomputern entstand aus dieser architektonischen Offenheit eine hohe Variabilität der PC-Hardware. Gemeinsam mit der Tendenz des Computerspiels, die Leistung aller Komponenten auszureizen, gab dieses Vorkommen unterschiedlicher Hardwarekonfigurationen den Aus-

¹³Vorbild dieser Konfigurierbarkeit von Stimmen ist die Spielereihe *Unreal Tournament*, in der die Spieler ebenfalls zwischen verschiedenen Kommentatoren wählen konnten.

¹⁴Ein Sound-Preset in *Bad Company 2* trägt den Titel „Home Cinema“.

¹⁵siehe (Ceruzzi, 1998, 277f)

SETTING YOUR SOUND

Customize your game's soundscape to take full advantage of your listening environment. Select YOUR SOUND SYSTEM from the Audio Options menu and enable the runtime Master Unit to optimize the mix for your audio setup.

Home Cinema	Mixed with full dynamic range for loud playback on large speakers.
Hi-Fi (default)	The default setting with standard dynamic range for mid-sized speakers.
Small speakers	Optimized for soft playback on small speakers with reduced dynamic range.
War Tapes	The most extreme setting for an intense audio experience.
Headphones	Specifically tuned for listening on headphones.

Abbildung 6.2: *Battlefield Bad Company 2* - Sound Presets

schlag, Einstellungsmöglichkeiten an die Spieler zu delegieren. Was sie an das Programm weiterreichten, war das Wissen über ihre Hardware-Konfiguration – allen voran die im PC eingebauten Soundkarten dieser Zeit brachten unterschiedliche qualitative Ergebnisse hervor.¹⁶ Optionen, die das Hardware-Setup betreffen, sind bis heute erhalten geblieben. So können die Spieler weiterhin das Ausgabegerät bestimmen, sofern mehrere Soundkarten zur Auswahl stehen, und festlegen, ob sie auf ein bestimmtes Lautsprecher-Setup oder Kopfhörer zurückgreifen möchten. Weiterhin sind bestimmte hardwaregestützte Features wie EAX ein- und ausschaltbar. In Ergänzung zu diesen basalen Hardware-Setups fanden in den 1990er Jahren weitere Kontrollmöglichkeiten Einzug in die Optionsmenüs. Es entstand ein Zusatzangebot von Einstellungsmöglichkeiten, das in den heutigen Konfigurationsmenüs mündete.

Exkurs: Konfigurierbarkeit auf Konsolen

Die Abschaltung und Ersetzung der Musik des Computerspiels ist ein Thema, das auch in Komponistenkreisen diskutiert wird. Der Diskurs flammte u. a. auf, als Microsoft die *Xbox 360* vorstellte. Eine Richtlinie der neuen Plattform sah vor, dass die gesamte Musik eines Spiels prin-

¹⁶Das Video „Evolution of PC Audio – As told by Secret of Monkey Island“ verschafft einen guten Eindruck, wie groß die Abweichungen verschiedener Soundkarten in Hinblick auf die MIDI-basierte Musikwiedergabe ausfielen: <http://www.youtube.com/watch?v=a324ykKV-7Y> (15.05.2013)

zipiell ersetzbar sei. Spieler erhielten nun die Möglichkeit, Musik ihrer Wahl während des Spielvorgangs laufen zu lassen. Im Grunde genommen war nun auch auf der Konsole ein Konfigurationsschema erlaubt, das in PC-Spielen schon lange etabliert war. Bereits in DOS-basierten Spielen wie den ersten beiden Titeln der *Quake* und *Grand Theft Auto* Reihe war es möglich, die Musik auszutauschen. Während das Spiel lief, konnte die CD-ROM, von der die Audiotracks gelesen wurden, durch eine beliebige andere Audio-CD ersetzt werden. Seit dem Aufkommen von Multitasking Betriebssystemen sind auch die Lautstärken parallel laufender Programme aufeinander abstimmbare. Spiele und Media-Player stehen seitdem in einvernehmlicher Koexistenz. Die Anforderung der potenziellen Ersetzbarkeit bzw. dem parallelen Ablaufen von Musik und Computerspiel der *Xbox 360* lässt sich als Ausdruck ihrer Redefinition als universelles Entertainment-Gerät verstehen. Während die *Xbox* noch als „videogame system“ auf den Markt kam, positionierte Microsoft den Nachfolger explizit als „videogame entertainment system“. Die Konsole rückte damit in die Nähe des PCs und adaptierte das Multitasking und die damit verbundenen Konfigurationsaspekte der Plattform.

Motivationen für den eigenen Mix

Angenommen, es stehen in dem Audio-Menü umfangreiche Konfigurationsmöglichkeiten bereit. Wie werden sie benutzt? Warum möchten einige Spieler beispielsweise die Musik abschalten? Ebenso wie die Spieler einem Computerspiel auf viele verschiedene Arten begegnen¹⁷, gebrauchen sie unterschiedliche Konfigurationsmuster, denen unterschiedliche Motivationen zu Grunde liegen. Der Gedanke liegt nahe, dass zwischen Spielweisen und Konfigurationsweisen eine Korrelation besteht. Die angebotenen Anpassungsweisen kommen den unterschiedlichen Arten entgegen, auf die ein Spiel gespielt werden kann.

Ein typisches Konfigurationsschema besteht darin, alle für das Spiel irrelevanten Klangkategorien auszublenden. Wenn beispielsweise Musik nicht an Ereignisse und Zustände des Spielmodells gekoppelt ist, transportiert sie irrelevante Informationen, die im schlimmsten Fall spielrelevante Geräusche überdecken. Insbesondere kompetitiv ausgerichtete Spieler ziehen daraus die Konsequenz, Musik vollständig abzuschalten.

¹⁷siehe Abschnitt §1.3

Umgekehrt können akustische Elemente, die den Spielern Nutzen bringen, durch konfigurative Eingriffe stärker hervorgehoben werden. Um das Abschalten von Klangschichten zu verhindern, müssten die Entwickler im Grunde genommen vollständig auf dekorative akustische Elemente verzichten, indem alles Hörbare an Zustände und Ereignisse des Spielmodells gebunden wird.

Im Gegensatz zu diesem funktionalen Konfigurationsschema, das mit den Zielen des im ersten Kapitel vorgestellten Spielens nach Zahlen korreliert, sind auch konfigurative Eingriffe denkbar, die nicht nach den Werten hinter den Zeichen suchen. Wenn die Spieler das Erleben einer Geschichte oder die Versenkung in der fiktiven Welt in den Vordergrund rücken, schalten sie möglicherweise potentiell störende Geräusche wie die Klänge der Benutzeroberfläche aus. Nicht die funktionale, sondern die ästhetische Dimension der Darstellungselemente fungiert als Wegweiser dieses entgegengesetzten Konfigurationstypus.

Neben diesen beiden Polen sind viele weitere Szenarien denkbar, die jeweils eigene Konfigurationsschemata mit sich bringen. In vielen Fällen wird das Spiel zu einer Art Meta-Beschäftigung, in deren Rahmen sich weitere Beschäftigungen vollziehen. In Online-Games gehört dazu beispielsweise die Pflege sozialer Kontakte. Solche Meta-Beschäftigungen geben den Ausschlag für Konfigurationen des Hörbaren. Die Klangkulisse des Spiels wird mit Hilfe des Audio-Mixers auf Betriebssystemebene weit in den Hintergrund gedrängt, während Sprachkonferenzprogramme wie *Teamspeak*, *Ventrilo* oder *Skype* auf volle Lautstärke gedreht sind.

Ein ähnliches Multitasking-Phänomen findet sich in Single-Player-Games. So stellen die Spieler die Musik des Spiels aus und speisen statt dessen über einen externen Media-Player andere Musik oder auch Hörspiele ein, die sich mit der restlichen Klangkulisse des Spiels vermischen. Informelle Gespräche mit Spielern ergaben, dass diese Konfigurationen vor allem dann auftreten, wenn die Spieler bereits mehrere hundert Stunden mit dem Spiel verbracht haben. Das Eintauchen in den bekannten Interaktionsfluß und die vertrauten Abläufe fungiert dabei eher als Nebenbeschäftigung, zu der weniger vertraute Inhalte konsumiert werden.

Während der gesamten Nutzungsdauer eines Spiels können sich die Prioritäten und Motivationen der Spieler verschieben. Es gibt also nicht nur unterschiedliche Konfigurationsmuster bei unterschiedlichen Spie-

lern, sondern auch unterschiedliche Muster bei einer einzigen Person über die gesamte Spieldauer hinweg. Die durch die Konfigurationsmenüs gegebenen Möglichkeiten tragen dazu bei, unterschiedlichen Anforderungen gerecht zu werden. Der Grund dafür liegt in der langen Zeit der Nutzung und der damit verbundenen Repetition von Abläufen. Die Anforderungen an das Spiel und dessen akustischer Gestaltung sind womöglich ganz anders, wenn die Spieler es das erste Mal durchspielen oder bereits 300 Stunden damit verbracht haben.

Wie bereits erwähnt wurde, sind Audio-Konfigurationen für den Großteil der Spieler selbstverständlich. In traditionellen Medien sind Eingriffe dieser Art hingegen nicht üblich. So ist es beim Film unvorstellbar, dass der Rezipient das Lautstärkeverhältnis der Klangschichten ändert oder einige Kategorien sogar komplett ausblendet. Der fixierte Mix ist dort ein integraler Bestandteil des Werks. Das Computerspiel bricht in dieser Hinsicht offenbar stark mit den bestehenden Traditionen. Das Medium fördert einen neuen Umgang mit seinen Inhalten und insbesondere mit inszenierten Klangkulissen. Es ermöglicht den Spielern, die Rahmenbedingungen ihrer Erfahrungen selber festzulegen.

Wie die vorherigen Ausführungen um die Abschaltbarkeit von Musik gezeigt haben, ist der Anspruch der Spielegemeinschaften so weit gestiegen, dass Forderungen an Änderungen und Ergänzungen von Computerspielen zur Tagesordnung gehören. Der eigene Sound-Mix ist nur ein Ausdruck dieses Phänomens, dass alle Bereiche der Gaming-Kultur durchdringt.¹⁸ Das Prinzip der Mitbestimmung greift dabei auch auf die Inhalte der Spiele über. So mündeten die Enttäuschungen vieler Spieler über das Ende der *Mass Effect* Trilogie in organisierten Protesten. In einer Petition der Initiative „Retake Mass Effect 3“ heißt es:

We believe: That it is the right of the writers and developers of the Mass Effect series to end that series however they see fit

However, we also believe that the currently available endings to the series: Do not provide the wide range of possible outcomes that we have come to expect from a Mass Effect game¹⁹

¹⁸Eine aktuelle Ausprägung ist das Crowd-Funding, bei dem die Spieler die Entwicklung ausgewählter Titel finanziell unterstützen.

¹⁹<http://retakemasseffect.chipin.com/retake-mass-effect-childs-play> (05.04.2012)

Die Unzufriedenheit über den Abschluss einer Geschichte ist dabei kein genuines Phänomen des Computerspiels. Jenkins stößt auf dieses ernüchternde Moment in Fan-Kulturen, die sich mit der Vorhersage (dem sog. Spoiling) von Fernsehserien beschäftigen.²⁰ Die Auflösung der Geschichte kann mit der Spannung, die im Verfolgen und Ausloten der Potentialitäten liegt, nicht mithalten. Der Unterschied zu den *Mass Effect* Protesten besteht allerdings darin, dass die Spieler nicht nur ihre Enttäuschung ausdrücken, sondern konkrete Änderungen fordern. Sie erwarten von den Entwicklern, dass diese einen Patch nachliefern, der die Kritikpunkte aufnimmt und damit mehr Möglichkeiten des Endes integriert.

Die Entwickler befinden sich damit in einer Position, die sich drastisch von der klassischen Autorschaft unterscheidet. So gilt es als unkonventionell, von dem Autor eines Romans alternative Ausgänge des Plots zu verlangen oder ihn dazu bewegen zu wollen, das „Balancing“ zwischen verschiedenen Figuren zu korrigieren.²¹ Audio-Konfigurationen, wie die Ersetzung von Musik, sind Phänomene, die sich in diesem ständig in Verhandlung befindlichem Spannungsfeld zwischen Produktion und Nutzung entfalten. Nicht übersehen werden sollte dabei, dass neben dem Ausschalten von Musik während des Spielens auch der umgekehrte Fall eintritt. So vermerkt ein Fan des Soundtracks von *Plants vs Zombies*: „As others have said I also leave the game running just to listen to the tunes“.²²

Fallstudie: „Sound Whoring“

Vor allem in kompetitiven Multiplayer-Titeln nutzen Spieler die Konfigurationsmöglichkeiten zur Optimierung ihres Spiels. In grafischer Hinsicht gehört dazu das Herabsetzen der Darstellungsqualität zu Gunsten eines flüssigeren Spielablaufs. Weiterhin ermöglicht die Detailarmut ggf. eine leichtere Differenzierung zwischen zustandsvarianten und -invarianten Darstellungselementen. In manchen Spielen lassen sich durch

²⁰(Jenkins, 2006, 34)

²¹Wir begegnen diesem Phänomen aber beispielsweise beim Barden, der den Wünschen seiner Zuhörer Folge leistet (Murray, 1997, 185ff), sowie im Bereich der Fan-Fiction

²²<http://forums.steampowered.com/forums/showthread.php?t=854924&page=4>
(30.07.2012)

die Helligkeits- und Kontrasteinstellungen Vorteile erzielen, indem dunkle Spielabschnitte besser einsehbar sind. In der *Gothic*-Serie erkunden die Spieler dadurch finstere Gebiete, ohne vorher spezielle Ingame-Items wie Fackeln zu erwerben. Ähnliche Konfigurationsstrategien sind für die auditive Ebene in Gebrauch, wenn auch weniger prominent als im Falle des visuellen Gegenparts..

In Multiplayer First Person Shootern ist die konfigurative Praktik, möglichst optimale Hörbedingungen herzustellen und als spielerische Vorteile auszunutzen so weit verbreitet, dass sich ein eigener, wenn auch abfälliger, Begriff dafür etabliert hat. Beim sogenannten *sound whoring* stützen sich Spieler während des kompetitiven Spiels massiv auf akustische Informationen wie Schrittgeräusche. Sie orten auf diese Weise gegnerische Spielfiguren hinter Wänden und sind weniger anfällig für Angriffe in ihrem Rücken. Spieler, die sich dieser Praktiken bedienen, werden häufig als „Camper“ diffamiert oder des Cheateus bezichtigt, wenn visuell ausgerichtete Spieler auf die Frage „Wie konnte ich gesehen werden?“ keine zufriedenstellende Antwort finden. Mittlerweile besitzt diese aufs gezielte Hören ausgerichtete Spielweise eine treue Schar von Anhängern, die unter anderem dessen pejorative Bezeichnung kritisieren.

Why the need to call it *whoring* when we all have volume controls....and all choose how we want to listen to them?
It's not like it's a hack, a mod, or anything illegitimate. It's *sound*, and we listen to it.²³

Ausgehend von den spielerischen Vorteilen, den das Hören gegnerischer Fußschritte verschafft, bildete sich in der Spielergemeinschaft ein eigener Diskurs darüber heraus. Neben grundsätzlichen Diskussionen über Wirkungskraft und Akzeptanz der Spielweise geben ihre Adepten der Gemeinschaft praktische Tipps, wie man sich die akustische Gestaltung am besten zu Nutze machen kann. In Foren und Videotutorials tauschen sie sich darüber aus, welche Headsets, Einstellungen und Hörstrategien den größten Erfolg versprechen. Typische Expertengespräche diskutieren in diesem Zusammenhang z. B., ob ein Surround-Headset ein differenzierteres Hören als ein Stereo-Headset ermöglicht, oder ein Stereo-Headset die entscheidenden Geräusche deutlicher abbildet. Im Mittelpunkt der Dialoge steht ein Austausch persönlicher Erfahrungs-

²³<http://www.cod4boards.com/forum/showthread.php?243592-Why-is-it-called-sound-WHORING> (03.04.2012)

werte, der häufig auch die Anwendung des „sound whoring“ in verschiedenen Spielen umfasst.

Wie verlässlich diese auditiv ausgerichtete Spielweise spielentscheidende Informationen verschafft, ist stark von der jeweiligen akustischen Ausgestaltung eines Computerspiels abhängig. Je besser die Möglichkeiten der Praktik des „sound whoring“ ausfällt, desto größer ist die Gemeinschaft, die sich dieser Praktik bedient. Spiele wie *Call of Duty* unterstützen sie dezidiert, indem sie akustische Aspekte mit der Spielmechanik verschalten. Ein Beispiel hierfür ist der Modifikator *SitRep*. In den Multiplayer-Modi mehrerer Titel der Reihe können Spieler durch Modifikatoren bestimmte Aspekte ihrer Spielfigur anpassen. So lässt sich beispielsweise die Bewegungsgeschwindigkeit erhöhen, das Nachladen beschleunigen oder der Fallschaden verringern. Der Modifikator *SitRep* bewirkt in der Pro-Variante, dass die Schrittgeräusche der Gegner viermal lauter und die Schrittgeräusche der Teammitglieder viermal leiser werden.²⁴ Für Spieler, die den Modifikator verwenden, ändert sich also der Sound-Mix stark zu Gunsten eines auditiv ausgerichteten Spiels. Um sich vor solchen Spielern zu schützen, bietet das Spiel einen weiteren Modifikator an. Durch die Wahl des *Ninja-Perks* in der Pro-Variante verstummen die eigenen Fußschritte.

Wenn die konzentrierte Wahrnehmung der akustischen Gestaltung spielerische Vorteile verschafft, gewinnt eine auf das Hören ausgerichtete Spielweise deutlich an Attraktivität. Fußschritte bilden bislang die wichtigste Kategorie in dieser Hinsicht. Spiele wie *Left 4 Dead* vermitteln jedoch auch durch weitere Geräusche und die Musik spielrelevante, räumliche Informationen. Immer dann, wenn ein Spezial-Infizierter in die Nähe der Überlebenden tritt, erklingt eine kurze Melodie und ggf. ein Sprachsample des Infizierten. Auch hier tauschen Spieler in Foren ihre Erfahrungen aus und geben Hinweise darüber, was einzelne Geräusche bedeuten. Der User „Darthcaboose“ verweist im folgenden Beitrag z. B. auf die verwendeten Musikbausteine, die sich auf der Betriebssystemebene auch isoliert abspielen lassen.

Head to your left4dead/sound/music folder and check out the “bacteria” and “mob” folders (I’m guessing we all know what the witch and tank music sounds like).

²⁴<http://callofduty.wikia.com/wiki/SitRep> (31.01.2012)

The “bacteria” folder contains the various sound clips the music director will play anytime a particular boss infected is about to spawn. The boisterous boomer, skittish hunter and the minor melodramatic chords of the smoker all give the skilled listener a chance to figure out what’s about to strike. Sometimes, you hear these particular sounds well before the infected boss has a chance to spawn.²⁵

Spieler setzen sich auf diese Weise aktiv mit den Bausteinen des Spiels und den Regeln, nach denen sie zusammengesetzt werden, auseinander. Der Einblick in den Aufbau des Spiels fördert das Verständnis über dessen Mechanismen und eröffnet spielerische Vorteile.

Manchmal sind die Informationen der akustischen Ebene auch über andere Kanäle verfügbar, die von kompetitiven Spielern mit einbezogen werden. So besitzen einige Computerspiele optional zuschaltbare Untertitel, die den Ton des Spiels textuell darstellen. In *Left 4 Dead* erscheinen die eingeblendeten Untertitel offenbar kurz bevor das Geräusch hörbar wird.

Yeah, you can see [Tank growls] before the music starts to kick in, or the ground starts shaking. Or [Witch Moan] before you actually hear the witch cry (unless you got epic headsets).²⁶

Die Texteinblendung ermöglicht insofern einen Informationsvorsprung. Eine Folge davon ist, dass auch körperlich unbeeinträchtigte Spieler Hilfen aktivieren, die eigentlich für Taube und Hörgeschädigte vorgesehen sind. So kommentiert ein User das Untertitel-Feature mit folgenden Worten:

I just turn on full CC²⁷ and can see everything :D. I’m slightly hard of hearing.²⁸

²⁵<http://web.archive.org/web/20090505043300/http://boomercharged.net/2009/01/19/the-other-director/>

²⁶<http://web.archive.org/web/20090505043300/http://boomercharged.net/2009/01/19/the-other-director/>

²⁷Closed Captioning: engl. Bezeichnung für Untertitel, die auch Geräusche textuell darstellen

²⁸<http://web.archive.org/web/20090505043300/http://boomercharged.net/2009/01/19/the-other-director/>

6.2 Multiple Darstellungsmodi

Der Kern eines Films liegt im Arrangement des Hör- und Sichtbaren, der Ton- und der Bildspur. Der Kern eines Computerspiels liegt in den Modellen hinter der Darstellung, und diese Modelle werden zunehmend komplexer. Eine Auswirkung des Mooreschen Gesetzes²⁹ auf das Computerspiel ließe sich darin ausmachen, dass sich immer detailliertere Modelle herausbilden - von physikalischen Simulationen der Schwerkraft bis hin zur Modellierung des Verhaltens künstlicher Intelligenzen. Für die Ausgaben des Computerspiels birgt die gestiegene Informationsdichte der modellierten Entitäten eine große Herausforderung: Die Ausgaben sind Selektionen aus immer größer werdenden Zustandsmengen. Die Auswahl dieser Selektionen wird in einigen Fällen an die Spieler delegiert und damit zu einem weiteren Aspekt der Konfiguration.

Informations-Overflow bei komplexen Modellen

Ab einer bestimmten Größe der Zustandsmenge entsteht das Problem, sie auf die beschränkte Fläche der Ausgabegeräte zu projizieren, ohne die Wahrnehmungskapazitäten der Spieler zu überfordern. Unseren Sinnen sind Grenzen gesetzt. Viele virtuelle Welten enthalten schlichtweg mehr Informationen, als Augen und Ohren auf einem Schlag aufnehmen können. Aus diesem Grund besteht die Ausgabe immer häufiger aus einer Auswahl aller verfügbaren Informationen. Hinter den dargestellten Ausschnitten der jeweiligen Sicht- und Hörperspektiven verbergen sich weitere Darstellungsschichten, die nach Bedarf ein- und ausgeblendet werden.

Eine Herausforderung des interaktiven Designs besteht darin, Kompromisse zwischen einer ästhetisch ansprechenden und einer funktionalen Darstellung zu finden. Vor allem in Open World-Spielen ist es absehbar, dass die von den Spielern verfolgten Spielweisen abschnittsweise unterschiedliche Anforderungen an die Darstellung stellen. So wird es für die meisten Spieler Phasen geben, in denen die Versenkung in der fiktiven Welt im Vordergrund steht. Wenn sie in *Assassin's Creed 2* eine Sightseeing-Tour durch das Venedig der Renaissance einlegen oder in

²⁹Gordon Moore, einer der Gründer des Halbleiterherstellers Intel, prognostizierte 1965, dass sich die Anzahl der Schaltkreiskomponenten auf einem Chip in regelmäßigen Abständen verdoppelt.

Grand Theft Auto 4 durch simulierte Freizeitaktivitäten ihre sozialen Kontakte mit virtuellen Bekannten pflegen, sind viele Informationen der Spielwelt, die in kritischen Spielsituationen unverzichtbar sind, nicht von Belang. Es bleibt zu vermuten, dass die Einblendung oder Hervorhebung von dezidiert spielrelevanten Informationen in diesem Modus eher störend wirkt. In entscheidenden Spielschnitten wie einer Verfolgungsjagd, die von den Spielern rasche Entscheidungen fordert, sind diese funktionalen Darstellungen hingegen elementar.



Abbildung 6.3: Optionale Informationen

eines Schalters innerhalb des Optionsmenüs ist das Ein- und Ausblenden solcher Informationen ein konfigurativer Akt, der die Darstellung direkt während des Spiels beeinflusst. Abbildung 6.3 zeigt die einblendbaren Zusatzinformationen auf der Weltkarte von *Civilization V*. In der unteren Grafik sind die Konturen der Hexagone sowie Icons für die Ressourcen der jeweiligen Felder hinzugegetreten.

Wie in Abschnitt 1.1 demonstriert wurde, sind die Darstellungen vieler modellierter Objekte bereits mit zusätzlichen Informationen angereichert, die dauerhaft angezeigt oder kontextabhängig einblendet werden. Manche Objekte oder Situationen treten durch visuelle und akustische Hinweise wie farbliche Umrandungen und suggestive Soundeffekte hervor. Einige dieser Zusatzinformationen blendet das Programm automatisch ein, sobald bestimmte Bedingungen erfüllt sind. Andere Zustandsinformationen werden durch den Befehl der Spieler einblendet. Ähnlich wie das Umlegen



Abbildung 6.4: Standardansicht und Detektivmodus (*Batman Arkham City*)

Einen Schritt weiter geht nun die Bereitstellung mehrerer Darstellungsmodi, die die Spieler nach Bedarf wechseln können. Dabei handelt es sich um narrativ gerechtfertigte Darstellungsweisen, die durch jeweils angedeutete Wahrnehmungsveränderungen der Avatare motiviert sind. In *Assassins's Creed* können die Spieler beispielsweise den Modus Eagle View aktivieren. Die opulenten grafischen Details der Spielwelt verschwinden in den Schatten, während Figuren und benutzbare Objekte hervortreten. Die Erzählung des Spiels erklärt Eagle View als eine den Assassinen angeborene Fähigkeit, die ihnen eine besonders geschärfte Wahrnehmung verleiht. Für die Spieler treten dadurch für bestimmte Spielabschnitte relevante Informationen in den Fokus. So sind beispielsweise die Charaktere der Spielwelt farblich markiert: Gegnerische Wachen und Patrouillen sind rot, und Zielpersonen schimmern golden. Auch in einer größeren Menschenmenge ist eine Zielperson dadurch leicht identifizierbar. In dem dritten Teil der Ezio-Trilogie sind in Ergänzung zu der farblichen Markierung die geplanten Laufwege der Wachen durch rote Schlieren gekennzeichnet. Die Spieler können diese Zusatzinformation in ihre Planungen einbeziehen.

Neben der Hervorhebung relevanter Informationen werden durch Eagle View auch Dinge sichtbar, die vorher verdeckt waren. An den Wänden bedeutender Monumente wurden dem Avatar Botschaften hinterlassen. Dort, wo im normalen Darstellungsmodus eine dekorative Textur die Wand beplasterte, sticht in Eagle View eine rot schimmernde Nachricht hervor. Die virtuelle Welt besteht aus mehreren Schichten, die durch den Wechsel des Darstellungsmodus separat fokussierbar sind.

Eine ähnliche Funktionalität wie Eagle View in *Assassin's Creed* stellt der Detektivmodus in *Batman Arkham Asylum* und *Arkham City* bereit. Auf Knopfdruck verschwinden die Details der Spielwelt und machen einer schematischen Ansicht Platz, die spielrelevante Objekte und Akteure hervorhebt. Die Screenshots in Abbildung 6.4 zeigen, dass der leuchtende Schriftzug „Gotham City Olympus“ beim Wechsel in den Detektivmodus als irrelevante Dekoration zurücktritt. Statt dessen wird die bewußtlose Person am Boden markiert, die in der Standardansicht kaum zu erkennen war. Darüber hinaus ist in der Mitte des Bildes eine einbrechbare Wand hervorgehoben, die Batman mit der richtigen Ausrüstung zum Einsturz bringen kann. Ebenso wie Eagle View reduziert der Detektivmodus die Spielansicht auf das Wesentliche und entlastet damit die Spieler in dem Identifizierungsakt relevanter Spielobjekte. Der Detektivmodus trifft eine audiovisuelle Vorauswahl zustandsvarianter Darstellungselemente.

Im Gegensatz zu der grafischen Ausgestaltung wurde die akustische Dimension im Design alternativer Darstellungsmodi bislang kaum funktional eingesetzt. Dabei wäre ein analoges Hervorhebungsprinzip auch mit akustischen Elementen realisierbar. So könnten z. B. Umgebungsgeräusche im Mix zurücktreten und spielrelevante Sounds wie Schrittgereusche stärker in den Vordergrund rücken.

Darstellungsalternativen

Neben multiplen Darstellungsmodi, die durch die Spieler umschaltbar sind und damit einen Aspekt der Darstellungskonfiguration bilden, kommen wechselnde Darstellungen auch als rein narrative und ästhetische Mittel zum Einsatz. Nicht nur die Modelle des virtuellen Raums werden durch steigende Ressourcen wie Rechenpower und Speicherkapazität umfangreicher, sondern auch die Möglichkeiten der Darstellung vervielfältigen sich. So kann eine Zustandsmenge in einem Spiel auf unterschiedliche Arten dargestellt werden.

Ein Beispiel für den Einsatz im Rahmen der Erzählung, in dem die gesamte Umgebung grafisch und akustisch auf zwei unterschiedliche Arten dargestellt wird, bietet *Bioshock 2*. In einer Spielszene schlüpfen die Spieler für kurze Zeit in die Rolle einer „Little Sister“ und bekommen die Spielwelt aus ihrer subjektiven Perspektive präsentiert. Die

gleichen Räume, die zuvor von dem Zerfall der Unterwasserstadt Rapture zeugten, sind in ihren Augen unversehrt. Die Leichen der verrückt gewordenen Splicer sind für sie friedlich ruhende Engel in einem Bett aus Rosen, um das keine Fliegen, sondern Schmetterlinge kreisen. Sowohl die grafische Gestaltung als auch das Sounddesign liegen für diesen Levelabschnitt doppelt vor. Die eine Fassung zeigt Rapture in der Wahrnehmung des Protagonisten Jack, die andere Fassung zeigt die Unterwasserstadt aus Perspektive einer „Little Sister“.

Ein weiteres Beispiel bietet das Jump'n'Run *Giana Sisters: Twisted Dreams*, in dem die Spieler jederzeit zwischen zwei Parallelwelten hin- und herschalten können. Die eine Welt zeigt eine blühende, die andere Welt eine verdorrte Vegetation. Die Hintergrundmusik wechselt nahtlos von Chris Hülsbecks und Fabian Del Priores Kompositionen und Arrangements in Heavy Metal Versionen der schwedischen Band Machinae Supremacy. Anders als in *Bioshock 2* erfüllt die doppelte Darstellung weniger einen narrativen Zweck als einen spielerisch-ästhetischen. Die einsammelbaren Diamanten sind über beide Welten verteilt, so dass die Spieler geschickt zwischen ihnen wechseln müssen, um eine hohe Punktzahl zu erreichen.

6.3 Die Simulation technischer Medien

Die Konfigurationsmöglichkeiten der Audio-Menüs sind nicht die einzigen Schalter, mit denen die Spieler den akustischen Rahmen ihrer Spielerfahrungen festlegen. Auch innerhalb der fiktiven Spielräume eröffnen sich Optionen, mit denen sie die Hörräume, in denen sie sich bewegen, gestalten. Eine herausragende Stellung nehmen in dieser Hinsicht die Simulationen technischer Medien wie dem Radio ein. Im Gegensatz zu den Einstellungen der Menüs treten diese ergänzenden Konfigurationen des Hörbaren auch als spielerische und narrative Handlungen in Erscheinung. Sie sind dementsprechend in dem Geflecht von Spiel und Erzählung zu verorten, das von den Entwicklern designt und arrangiert wird, und das die Spieler auf unterschiedliche Weisen durchschreiten.

Welche technischen Medien in einem Titel simuliert werden, richtet sich in der Regel nach der fiktiven Spielwelt. Die Einbindung setzt ein Setting voraus, in dem diese Medien plausibel erscheinen. Ein Radio in einem Mittelalter-/Fantasy-Setting bedürfte einiger Erklärung, um

die Stimmigkeit der fiktiven Welt zu wahren.³⁰ In Computerspielen, die in der Zukunft spielen, sind hingegen auch konfigurative Eingriffe leicht zu rechtfertigen, die für uns noch im Bereich der Utopie liegen. So ist es in dem in *Watch Dogs* modellierten Chicago der Zukunft möglich, Telefonate mitzuhören und Elemente der Infrastruktur wie Ampelschaltungen zu manipulieren. Im Folgenden werden einige typische Simulationen technischer Medien abgeschrieben, die den Spielern die Konfiguration des akustischen Raums der Spielwelt ermöglichen.

Radio

Simulationen des Radios sind am häufigsten in Computerspielen vorzufinden, die umfangreiche virtuelle Welten entwerfen und den Spielern deren Erkundung mit Hilfe eines Avatars ermöglichen. Im postnuklearen Ödland von *Fallout 3* empfangen die Spieler z. B. neben dem unabhängigen Ein-Mann-Sender „Galaxy News Radio“, in dem sowohl kritische Berichte über das Weltgeschehen als auch Musik der 40er Jahre laufen, die Station „Enclave Radio“ mit Marschmusik und Propagandareden des selbsternannten Präsidenten. Neben Rollenspielen wie *Fallout* greifen aber auch andere Genres auf simulierte Radios zurück. So enthält z. B. die Wirtschaftssimulation *Tropico 3* den Radiosender „Tropico News Today“, welcher die Spieler sowohl mit Fertigstellungsnachrichten von Gebäuden über den Spielfortschritt auf dem Laufenden hält, als auch mit lateinamerikanischer Musik unterhält. Die größte Popularität erlangten Radiosimulationen im Rahmen der *Grand Theft Auto* Reihe. Das Radio bildet dort den Schlüssel zu einer ausdifferenzierten urbanen Musikkultur, die innerhalb der fiktiven Städte über die Empfangsgeräte zugänglich ist. Bereits der erste, 1997 erschienene Teil verknüpfte die Abspielmomente der Tracks einer Audio-CD mit dem Kapern eines Kraftfahrzeugs. Sobald der Avatar in ein Auto einstieg, startete das simulierte Radio des Wagens. Alle nachfolgenden *Grand Theft Auto* Titel bauten dieses Musikintegrationskonzept sukzessive aus, indem sie die Anzahl und das Angebot der Radiosender kontinuierlich erweiterten. So bietet der 2008 erschienene Serienteil *Grand Theft Auto IV* insgesamt 20 verschiedene Sender an. Der Sender „Independent FM“

³⁰ Meistens wird in solchen Fällen auf die dem Film entlehnte extradiegetische Musikkonzeption mit adaptiven Elementen oder Live-Darbietungen innerhalb der Spielwelt, z. B. durch Musiker auf dem Marktplatz, zurückgegriffen.

erlaubt es den Spielern darüber hinaus, beliebige Stücke der eigenen Musiksammlung in das Spiel zu integrieren.

Welche Bedeutungen die Radiosender für die Spieler entwickeln, hängt nicht nur von ihren Inhalten, sondern auch von ihren Implementierungen ab. So sind z. B. in den ersten Teilen der *Grand Theft Auto* Reihe die Sender mit Fahrzeugklassen verknüpft. Stiehlt der Spieler einen Truck, so ist der Country-Sender eingestellt, während ein schwarzer Sportwagen den Rocksender als Vorgabe besitzt. Der Default-Wert spiegelt die stereotypen Musikpräferenzen der Wageninhaber wider. Die in Teil 2 eingeführte Möglichkeit, in geklauten Fahrzeugen den voreingestellten Sender zu verändern ermöglicht neue Akte der Rekonfiguration in einer perfekt konfigurierten Welt. Durch das Ändern der Radiosender bauen sich die Spieler dadurch innerhalb des gestohlenen Wagens ein Stück Privatsphäre wieder auf, die sie bei den gewaltsam beraubten NPCs zuvor durchbrochen haben. Im Gegensatz zu allen anderen motorisierten Verkehrsteilnehmern, die ihrem vorprogrammiertem Musikgeschmack folgen, treffen die Spieler ihre eigene Wahl. Doch selbst in *Grand Theft Auto IV* und dessen Fülle von Radiosendern bleibt die Wahl des Senders als Ausdruck von Individualität eine Illusion. Im Gegensatz zu realen Autoradios lässt das mit dem Fahrzeug verschaltete Radiomodell eine wichtige Funktion vermissen, nämlich das Abschalten. Da für viele Missionen und die Progression des Spiels die Nutzung von Fahrzeugen erforderlich ist, begeben sich die Spieler zwangsweise in die Situation der Dauerberieselung. Ein weiteres Implementierungsdetail, der uniforme Empfang in allen Bereichen der Stadt, untermauert diese Deutung.

In *Fallout 3* und *Fallout New Vegas* vermittelt das Radio durch ein anderes Implementierungsmodell andere Bedeutungen. Nicht überall im postnuklearen Ödland sind die spärlichen Sender erreichbar. Sobald sich der Avatar der Spieler außerhalb der Reichweite des eingeschalteten Senders befindet, hören sie nur noch statisches Rauschen. Das Radiomodell berücksichtigt hier also auch den Senderadius und simuliert dadurch den Vorgang der Übertragung selbst. Im Quest-Verlauf des Spiels trifft der Avatar auf Radiomoderatoren und beteiligt sich an dem Wiederaufbau der Radiolandschaft durch die Reparatur von Sendetürmen. Symbolisierte das nicht abschaltbare Empfangsgerät in *Grand Theft Auto IV* noch die Zwänge einer von Massenmedien beherrschten Konsumgesellschaft, so steht es in *Fallout* eher für den Hoffnungsschimmer einer beinahe untergegangenen Menschheit auf ihr zukünftiges Wieder-

auferstehen. Die Radiosender begleiten die Spieler bei ihren einsamen Wanderungen durch das Ödland als kreisförmige Inseln der Zivilisation.

Wie in Kapitel 3 bereits behandelt wurde, ist die musikalische Untermalung von Computerspielen je nach dem verfolgten Konzept sehr aufwändig. Neben den bausteinartig organisierten musikalischen Strukturen müssen Abspielprozesse und Transitionspunkte festgelegt werden. Insbesondere eine geplante stimmige Musikbegleitung ist dabei aufgrund der zeitlichen Ungewißheit des Handlungsverlaufs schwer zu realisieren. Mit der Simulation technischer Medien gelangte ab den 1990er Jahren eine neue Integrationsform von Audio-Inhalten in das Entwickler-Repertoire, die insbesondere für die Einbindung lizenzierter Musik geeignet ist. Sie ermöglicht die Integration linearer Musik in ein nicht-lineares Medium. Während lizenzierte Musik ansonsten hauptsächlich in Menüs untergebracht ist, wie u. a. bei den Sportspielen von Electronic Arts oder den Titeln der *Dirt*-Reihe, ermöglichen die Simulationen technischer Medien den Einsatz lizenzierter Musik innerhalb des Spielgeschehens. Indem die Steuerung der Abspielprozesse von simulierten Radios und sonstigen Geräten an die Spieler weitergereicht wird, ist das Programm in dieser Hinsicht entlastet. Die Verwendung von Medien zur Musikwiedergabe befreit die Entwickler von der Notwendigkeit, Regeln für den Musikverlauf definieren zu müssen. Die Konfiguration technischer Medien durch die Benutzer tritt an die Stelle der Konstruktion plausibler Fortläufe der akustischen Ausgestaltung. Im Gegensatz zu adaptiver Musik, die durch ein komplexes Ereignis-Setup an das Spielgeschehen gekoppelt wird, stellt die Musikwiedergabe über das Radio einen konfigurativen Akt der Spieler dar.

Die eingangs angesprochene Abhängigkeit technischer Medien von dem jeweiligen Setting bestimmt, zu welchem Grad die darüber integrierte Musik andere Musikkonzepte ersetzen kann. Das Setting der modernen Großstadt in der *Grand Theft Auto* Reihe bietet durch die Autoradios und viele weitere beschallte Orte wie Discos und Internet-Cafés so viele Plätze, an denen die Spieler Musik ausgesetzt sind, dass die Entwickler es dabei beließen. In *Fallout New Vegas* ist hingegen zusätzlich zu den Radioinhalten auch extradiegetische Musik integriert. Sie springt erst dann ein, wenn die Spieler das Radio ausgeschaltet haben. Die Entscheidung liegt bei ihnen, ob sie lieber Frank Sinatra und den Ink Spots im Radio lauschen, oder den aus früheren *Fallout*-Teilen übernommenen Ambient-Soundtrack von Mark Morgan in bestimmten Regionen zu hören bekommen.

Die über das Radio integrierte Musik erfüllt dabei andere Funktionen als Musikkonzepte, die eine extradiegetischen Untermalung vorsehen. Die Radio-Musik unterstreicht nicht wie im Film bestimmte Szenen und Momente, sondern ist ein konstitutiver Bestandteil der fiktiven Welt. In *Tropico 3* trägt der Latin-Soundtrack zur Etablierung der von den Spielern als El Presidente verwalteten Bananenrepublik bei. Im postnuklearen Ödland der *Fallout* Reihe fungieren die Stücke von den Ink Spots, Billie Holiday und weiteren Interpreten als Reminiszenzen an vergangene Zeiten und konstituieren damit das retro-futuristische Setting.

Ein interessanter Effekt der Musikauswahl zur Charakterisierung bestimmter Orte oder Zeiten besteht darin, dass die Spieler eine Verbindung zu Musik aufbauen, die ihnen ansonsten womöglich verschlossen geblieben wäre. Ein Beispiel dafür ist die wieder aufgeflammete Popularität der Ink Spots als Aushängeschild der *Fallout*-Reihe. Sie zeigt sich in den zahlreichen Kommentaren zu dem auf youtube eingestellten Titelsong von *Fallout 3*, „I don't want to set the world on fire“³¹:

Im 13 and I absolutely love the Ink Spots (acursedpaladin)

Modern music pales in comparison to the fantastic art provided by the ink spots and I for one am thankful to the fallout series for introducing me to this amazing band. (AdmiralNesence)

you know whats funny after playing fallout 3 I went to my grandmas house and was singing the song and she was like how do you know that song. (slosshexo)

When I found out my dead Grandma had an autograph from the ink spots, I flipped shit. (LAWLCthulhu)

Als Zeiteugnis der 1940er Jahre leistet die Aufnahme der Ink Spots im Kontext des Spiels eine sehr eindringliche Verbindung zwischen Realität und Fiktion. Dass die Großmutter das Stück aus ihrer Jugend noch kennt, macht das im Spiel suggerierte Vergangene zu einem gewissen Grad authentisch. Die Verknüpfung der durch einen Nuklearkrieg in Schutt und Asche gelegten Welt mit der Zeile „I don't want to set the world on fire“ und dem hoffnungsgeschwängerten Gesang der Vokal-

³¹http://www.youtube.com/all_comments?v=6l6vqPUM_FE (04.10.2012)

gruppe sorgt hingegen für eine ironische Brechung. Das Resultat daraus ist eine verklärte Nostalgie, die ebenso auf eine Zeit verweist, in der noch „gute Musik“ gespielt wurde, als auch eine mit dem Spielerlebnis in Verbindung gebrachte Sentimentalität.

Jukebox

Neben dem Radio gehört die Jukebox zu den am häufigsten anzutreffenden Abspielgeräten in virtuellen Welten. Sie bietet sich ebenfalls dazu an, lineare Musik in das Computerspiel zu schleusen. Sprachbeiträge sind hingegen unüblich. In *Bioshock* sind über die Geräte insgesamt 30 Songs aus den 1940er und 50er Jahren zu hören, die in Kombination mit dem grafischen Art déco Stil maßgeblich das retro-futuristische Setting der Unterwasserstadt Rapture charakterisieren. *Fallout 3* bietet ebenfalls die Jukebox als Abspielgerät, verwendet sie jedoch lediglich als stationäres Empfangsgerät für den fiktiven Radiosender „Galaxy News Radio“, anstatt ein eigenes Lieder-Repertoire anzubieten.

Anders als das portable, in der Regel nicht sichtbare Radio, das die Spieler auf ihren Streifzügen begleitet, ist die Jukebox als unbewegliches Objekt an einem zuvor festgelegten Ort platziert. Sie leistet die akustische Konfigurierbarkeit eines Raumes, der im Rahmen des Leveldesigns mit ihr ausgestattet wurde. Ihr Sound entspringt einer oder potentiell auch mehreren Quellen, falls ihre Simulation Lautsprecher enthält, die an verschiedenen Koordinaten im virtuellen Raum positioniert sind. Die Größe der Jukebox erleichtert ihr Auffinden durch die Spieler und ist sicher ein Grund für ihre weite Verbreitung. So greifen selbst Spiele mit einem Science Fiction Setting wie *Starcraft II* auf die Jukebox als Abspielgerät zurück. Einen weiteren Grund für ihren häufigen Einsatz bildet ihr natürliches Auftreten in öffentlichen Räumen, das keiner narrativen Rechtfertigung bedarf.

Technische Medien wie die Jukebox gehören primär zur Klasse optional durchführbarer Spielaktionen. Sie erfüllen in den meisten Fällen keine unmittelbaren ludischen Funktionen. Für den Ausgang einer Spielsituation, das Gewinnen oder Verlieren, ist der Senderwechsel eines Radios oder das Einschalten einer Jukebox in der Regel irrelevant. Stattdessen fungieren die Geräte als wichtige narrative Funktionsträger, mit deren Hilfe die Spieler ihre Spielerfahrungen weiter formen.

Trotz des dekorativen Charakters sind Implementierungen der Jukebox denkbar, in denen die darüber abgespielte Musik Spielereignisse auslöst. In der benutzergenerierten Kampagne „Escape from Toronto“ für *Left 4 Dead 2* löst das Benutzen einer Jukebox ein Crescendo-Event aus, das von *System of a downs* Song B.Y.O.B. begleitet wird.³²

Auch bei der Jukebox stellt sich die Herausforderung des Settings. In einer mittelalterlichen Welt wäre ein großer erzählerischer Aufwand nötig, um eine von sich aus spielende Musikmaschine zu rechtfertigen. Eine Jukebox im Fantasy-Dungeon verstößt gegen die Erwartungshaltung der Spieler. Die Einbindung musizierender NPCs bietet in diesen Fällen eine Möglichkeit, um die Funktionalität der Jukebox nichtsdestotrotz anzubieten. In *The Elder Scrolls: Skyrim* befinden sich an vielen Orten der Spielwelt Barden, von denen sich die Spieler Stücke wünschen können. Ihre Playlist enthält Titel wie „The Age of Oppression“, „Ragnar The Red“ und „The Dragonborn Comes“. Mit dem Addon *Hearthfire* wurde es darüber hinaus möglich, einen Barden für das virtuelle Eigenheim des Avatars anzuheuern. Der Barde ist somit gewissermaßen die Jukebox des Mittelalter-Settings. Genau so wie sie verschafft sie den Spielern die Möglichkeit, einen bestimmten Ort auf Anfrage mit Musik zu beschallen. Die Verortung der Musik innerhalb der virtuellen Räume trägt darüber hinaus dazu bei, die fiktive Welt mit einer erlebbaren Musikkultur anzureichern.

Funkgerät, Fernseher, Telefon, und Spiel

Neben Radio und Jukebox, die vorrangig zur Wiedergabe von Musik eingesetzt werden, enthalten Computerspiele eine Reihe weiterer simulierter Medien, die primär im Dienste der Erzählung stehen. So berichten Zeitungen und Fernsehbeiträge in *Deus Ex Human Revolution* über die politischen Entwicklungen und Ereignisse der fiktiven Spielwelt, in welche der vom Spieler gesteuerte Adam Jensen verstrickt ist. In *BioShock* treiben die über ein Funkgerät empfangenen Befehle und Monologe die Handlung maßgeblich voran. Sobald der Avatar unsichtbare Triggerflächen im virtuellen Raum überschreitet oder Quest-bezogene Aktionen ausführt, werden die Sprachsamples abgespielt.

³²<http://www.youtube.com/watch?v=q7Pj14uOX38> (03.08.2012)

Das Mobiltelefon in *Grand Theft Auto IV* dient schließlich nicht nur als Kommunikationszentrale für die Aufträge und sozialen Kontakte des Protagonisten Niko Bellic. Über den Telefonservice ZIT erfährt der Spieler auch den Interpreten und Titel des Musikstücks, das gerade im Radio läuft. Die Informationen werden an die Plattform Rockstar Social Club weitergeleitet, über welche die Spieler den Song in der Realität erwerben können. Alternativ ist auch der Kauf über Amazon MP3 möglich. Das Mobiltelefon etabliert über diesen Mechanismus eine Brücke zwischen den Inhalten der fiktiven Welt und der realen Musikindustrie. Verglichen mit virtuellen Welten, in denen sich komplexe Wirtschaftskreisläufe herausgebildet haben³³, handelt es sich um eine rudimentäre Verknüpfung. Trotzdem ist der Kauf lizenzierter Spielmusik auf diese Weise ein Novum. Die technische Grundlage für diese zusätzliche Verwertung musikalischer Inhalte bildet die Geschlossenheit des Software-Systems. Im Gegensatz zu den ersten beiden Teilen der *Grand Theft Auto* Reihe, in denen die Musik auf Audio-CDs untergebracht war, verwendet *Grand Theft Auto IV* ein exotisches Container-Format für die Audio-Daten. Spieler, die sich nicht näher damit auseinandersetzen, sehen sich zum Kauf gezwungen, wenn sie ein Stück außerhalb des Spiels hören möchten. Ein ähnlicher Trend der Zweitverwertung, der die Hörer durch künstlich eingeführte technische Hürden zum doppelten Kauf verleiten will, wurde bereits in Hinblick auf Handy-Klingeltöne festgestellt.³⁴ Um dennoch an die Daten zu gelangen, entwickelten technisch versierte Fans des Spiels das Programm *Open IV*, das die Extraktion aller Audio-Daten ermöglicht. Neben der Musik lassen sich damit auch alle anderen Radiobeiträge außerhalb des Spiels hören. Geschäftsmodelle wie der Verkauf von Musik aus Computerspielen, die auf programmtechnisch implementierten Restriktionen basieren - in diesem Fall die Unterbringung der Musik in einem proprietären Containerformat - rufen Gegenbewegungen auf den Plan. Die akustische Architektur des Computerspiels kann sich in diesem Sinne auch zu einem Spannungsfeld aus monetären Interessen und freien Zugangsweisen zu Spielinhalten entwickeln.

Virtuelle Welten enthalten nicht nur technische Medien, sondern auch weitere Spiele. In einigen Fällen handelt es sich dabei um Karten- oder Würfelspiele, die den Charakteren des Spiels als Einnahmequelle dienen. In anderen Fällen sind die integrierten Spiele selber Computerspiele

³³(Castronova, 2003)

³⁴(Drescher, 2008, 53)

le eines anderen Genres oder einer anderen Epoche. Die Bereitstellung von Spielen in Spielen ist ein weiterer Ausdruck der Maturität des Mediums, der nicht zuletzt seine rasante technische Entwicklung betont. Im Jargon der Spielekultur hat sich dafür der Begriff *Minigame* etabliert. Ebenso wie Radio und Jukebox eine Musikkultur innerhalb einer fiktiven Spielwelt etablieren, transportieren Minigames ein Freizeitangebot in die virtuellen Gefilde.

Eines der bekanntesten Minigames, das der Bezeichnung zum Trotz sehr umfangreich ausfällt, befindet sich in dem 1993 erschienenem Grafikadventure *Day Of The Tentacle*. Im Zimmer des NPCs Eddy findet der von den Spielern gesteuerte Bernhard einen Computer vor. Wenn die Spieler Bernhard dazu auffordern, den Computer zu benutzen, wird er selbst zum Spieler. Eddys Zimmer verschwindet und es erscheint der Startbildschirm von *Maniac Mansion*, dem Vorgänger von *Day Of The Tentacle* aus dem Jahr 1987. In der Rolle von Bernhard können die Spieler sich nun beliebig lang mit dem komplett enthaltenen Computerspiel auseinandersetzen. Wenn sie es durchspielen, während des Spielverlaufs scheitern oder es frühzeitig beenden, kehren sie als Bernhard in das Zimmer von Eddy zurück. Die Einbindung von *Maniac Mansion* in *Day Of The Tentacle* ist eine eindrucksvolle Demonstration der gestiegenen Rechenkapazitäten und Ressourcen des Computers. Während *Maniac Mansion* 1987 die damals gegebenen technischen Möglichkeiten weitestgehend ausreizte, sind seine Anforderungen in Hinblick auf die 1993 verbreiteten Computersysteme so gering, dass es komplett innerhalb des Nachfolgers platziert werden kann. Während die Einbindung umfangreicher Spiele wie *Maniac Mansion* selten vorkommt, sind kleinere Spiele oder Ausschnitte aus Spielen häufiger anzutreffen. So enthielt die 1991 veröffentlichte *Commander Keen*-Episode „Goodbye Galaxy“ ein Remake von *Pong*, und in dem Shooter *Rage* sind Räume aus älteren Spielen von id Software – *Wolfenstein 3D*, *Doom* und *Quake* – versteckt.

Glitch

Technische Medien lenken Aufmerksamkeit auf sich selber, wenn sie von ihrem erwarteten Verhalten abweichen. Fehlfunktionen und das kurzzeitige Aussetzen des Normalbetriebs sind auch in Computerspielen vorhanden. Als ungewollte Störungen zeugen sie von der Komple-

xität des Software-Systems, in dem durch unzureichend einkalkulierte Konstellationen immer wieder Bugs, Exploits und Glitches zu Tage treten. Software besitzt aus diesem Grund selten eine finale Form. Sobald Unzulänglichkeiten bekannt werden, erscheinen Patches und Bugfixes, um die entdeckten Risse zu flicken und auszuglätten. Bei Computerspielen gehören dazu neben der Behebung technischer Probleme auch Anpassungen des Regelsystems. Insbesondere das Balancing – die Feinabstimmung gegeneinander ausspielbarer Einheiten, Gegenstände und Strategien – gibt den Ausschlag für jede Menge Änderungen, die von den Spielern berücksichtigt werden müssen. Patches erscheinen nicht ausschließlich von den Entwicklern des Spiels. Titel, die eine große Fan-Gemeinschaft um sich scharen, werden gelegentlich auch von einigen technisch versierten Fans weiterentwickelt. Ein imposantes Beispiel hierfür ist das Engagement der Fan-Gemeinschaft des Rollenspiels *Vampire: The Masquerade - Bloodlines*. So zeugt der 2012 veröffentlichte Unofficial Patch 8.2 selbst acht Jahre nach Erscheinen des Spiels von den Bemühungen der Community, das Spiel zu verbessern.

Im Gegensatz zu ungewollten Störungen, die auf ihre Behebung drängen, werden gewollte Störungen technischer Medien gezielt als ästhetische, narrative und spielerische Mittel eingesetzt. Ein bereits angesprochenes Beispiel dieser Art sind die simulierten Radiosender in *Fallout 3*, die außerhalb ihrer Sendereichweite nur als unbestimmbares Rauschen zu hören sind. Der gestörte Empfang markiert dabei die bislang unerreichten Gebiete der sich im Wiederaufbau befindlichen Zivilisation. Ein weiteres Beispiel bietet *Deus Ex Human Revolution*, in dem die Spieler in die Rolle des mit technischen Augmentierungen versehenen Agenten Adam Jensen schlüpfen. Während Jensen in Hengsha darauf hinarbeitet, die Vorsitzende des Konzerns Tai Jong Medical zu konfrontieren, treten Glitches in seiner Software auf. Bild und Ton fallen kurzzeitig aus und berauben die Spieler des Sicht- und Hörvermögens. Folgen die Spieler der Aufforderung, ein Chip-Upgrade in der nächsten LIMB-Klinik vornehmen zu lassen, sind die Probleme kurzzeitig behoben. Als sich herausstellt, dass die neuen Chips der Teil eines Komplotts sind, müssen die Spieler in der Rolle von Jensen einen kompletten Boss-Kampf mit den eingeschränkten Wahrnehmungsmöglichkeiten überstehen. Die defekten Augmentierungen der Spielfigur Jensen werden durch die visuellen Störungseffekte direkt an die Spieler weitergereicht. Sowohl in der First- als auch in der Third-Person-Perspektive blinken die Infor-

mationselemente des Head-up-Displays, während Lichter und Konturen der Spielwelt bläulich flackern.

Die technischen Augmentierungsdefekte in *Deus Ex Human Revolution* sind direkt in die Erzählung eingebettet. In *Batman Arkham Asylum* und *Arkham City* treten ähnliche Bildstörungen auf, wenn der Avatar in Kämpfen Treffer einsteckt. Der kurzzeitige Wahrnehmungsverlust des strauchelnden, angeschlagenen Avatars trifft auf diese Weise auch die Spieler. In diesem Fall ist die Störung keinem innerhalb der virtuellen Welt etablierten technischen Medium zuzuordnen, sondern der Wahrnehmung der Spielfigur. Das Programm übersetzt die eingesteckten Schläge des Avatars für die Spieler in Bild- und Tonstörungen. Inszenierte Glitches können aber auch einen Schritt weitergehen, wie die folgende Spielszene demonstriert.

„Did anyone catch the game last night?“ Die Ansage über die in der Irrenanstalt Arkham Asylum installierten Lautsprecher und das Husten von Batman hätten ein größeres Mißtrauen wecken können. Und doch erwischt es die Spieler eiskalt, als sich das Bild in bunt gerasterten Blöcken verliert und der Sound nach einem unkontrollierten Rauschen und einer zerhackten Stimme verstummt. Sie überkommt eine Furcht, die sich im Bangen um die Funktionstüchtigkeit ihrer Hardware ganz realen Dingen zuwendet.

First reaction:...Aw fuck I broke my TV!³⁵

That scared me. I was thinking that my video card died :O³⁶

oh my god when i was playing this nd when tha pink glitchy shit came up i was liek OH NOEZ WHAT'S BROKEN???? XBOX OR TV???³⁷

Und der Alptraum geht noch weiter. Kurz nachdem der zerhackte Bildschirm schwarz wird, wiederholt sich das Intro des Spiels. Das Batmobil fegt durch eine mit Laub bedeckte Allee in Richtung von Arkham Asylum. Vielleicht sind Fernseher, Playstation und Grafikkarte doch noch

³⁵http://www.youtube.com/all_comments?v=x3DyZll5GwA (08.10.2012), User SweeneyTodd98

³⁶ebd., User PleaseXit

³⁷ebd., User izeatstoomuch

unversehrt. Es dauert einen Moment, bis sich die nächste Irritation einstellt. Am Steuer sitzt nicht Batman, sondern der Joker. Schnitt. Aus der gewohnten Third-Person-Perspektive steuern die Spieler nun auch den Joker. Unter seiner Begleitung wird der gefesselte Batman in die Irrenanstalt eingeliefert. Schnitt. Die Spieler haben nun wieder die Perspektive von Batman eingenommen. Der Joker zückt eine Schußwaffe, zielt in Richtung der Spieler und drückt ab. Schnitt. Das Spielmenü wird angezeigt. Mission failed, „Retry“ oder „Quit“? Dazu der belehrende Hinweis „Use the middle stick to dodge Joker’s gun fire.“. Das Gamepad besitzt allerdings keinen „middle stick“. Schnitt. Ein Grabstein mit der Aufschrift „R.I.P. Bruce Wayne“. Batman steigt aus dem Grab und tritt dem Bossfight mit Scarecrow entgegen. Der Nebel der Verwirrung lichtet sich.

Der inszenierte Glitch im Rahmen dieser Traumsequenz leistet etwas, das nur wenigen kulturellen Artefakten gelingt. Er evoziert eine außerordentlich eindringliche Verbindung zwischen Fiktion und Realität, welche die Rezipienten beeindruckt, aber auch verstört zurück lässt. Mehrere tausend Kommentare auf *youtube*³⁸ und in weiteren Kanälen der Netzgemeinschaft zeugen von der Verarbeitung dieses regelrecht traumatischen Erlebnisses. Das Oszillieren zwischen dem Virtuellen und der Realität, technisch vermittelten fiktiven Inhalten und dem Versagen der für diese Vermittlung notwendigen Technik stellt das Vertrauen und die Versenkung in das virtuelle Geschehen grundsätzlich in Frage. Es reißt uns gewaltsam aus den simulierten Abläufen, nur um uns mit einer größeren Portion Mißtrauen wieder dorthin zurück zu befördern. Ähnlich wie in der eingangs erwähnten *Stanley Parable*, in der die Spieler nur gewinnen können, indem sie das Programm beenden, legt die Szene aus *Arkham Asylum* den Spielern die Reflektion ihrer Rolle als User nahe.

³⁸www.youtube.com/all_comments?v=x3DyZll5GwA (08.10.2012)

Ausblick

Von der Stadt über den Wolken geht es in die Stadt unter Wasser. Plötzlich befinden wir uns nicht mehr in Columbia, dem Schauplatz von *Bioshock Infinite*, sondern in Rapture, dem Schauplatz der beiden vorangegangenen Teile der *Bioshock*-Serie. Während wir Django Reinhardts Version des Chansons *La Mer* hören und über den Sprung in die bekannte Unterwasserstadt rätseln, ist auch unser Avatar sichtlich verwirrt. Auf seine Frage, was das hier sei, antwortet seine Begleiterin Elisabeth lediglich: „Ein Durchgang. Einer von vielen.“ Dann leitet sie uns zu einer Tauchkugel, die zurück zu dem Leuchtturm führt, an dem alles begann. Doch sobald wir die Tür des Leuchtturms passieren, erstreckt sich vor uns ein Meer von Leuchttürmen.

Es gibt eine Million Millionen Welten. Alle verschieden und alle ähnlich. Konstanten und Variablen. Da ist immer ein Leuchtturm. Da ist immer ein Mann, immer eine Stadt.³⁹

Elizabeth kommentiert das Bild und beschreibt damit nicht nur die in der Erzählung des Spiels auftretenden Paralleluniversen, sondern auf einer Metaebene auch Prinzipien des Gamedesigns. Einerseits lässt sich der Kommentar auf die *Bioshock*-Reihe und die Struktur von Sequels beziehen. Ob in Rapture oder Columbia, der grundlegende Aufbau ist identisch. Der Leuchtturm birgt das Tor in eine dystopische Stadt, in der unser Avatar eine Reihe von Herausforderungen bestehen muss. Ein Leuchtturm, ein Mann, eine Stadt. Andererseits verweist der Kommentar auf einen Aspekt des Aufbaus von Computerspielen und Software an sich und damit auch deren akustischer Architektur: Konstanten und Variablen. Dieser Aspekt wird nicht nur in der Schlusssequenz von *Bioshock Infinite* aufgegriffen. So begegnen uns während des Spielverlaufs immer wieder die Physiker Rosalind und Robert Lutece, die sich damit auseinandersetzen. An einer Stelle lassen sie uns eine Münze werfen.

³⁹ *Bioshock Infinite*, 2013

Das Ergebnis ist immer das Gleiche, der Münzwurf ist vorprogrammiert, konstant. In einem späteren Abschnitt stellen sie uns vor die Wahl zwischen zwei verschiedenen Schmuckstücken für Elizabeth. Der Ausgang ist variabel. So wie Rosalind und Robert Lutece die Konstanten und Variablen der Paralleluniversen ergründen, loten Spieler ihre Handlungsfreiheit aus. Beim Münzwurf können sie nur einen Knopf auf dem Gamepad drücken, das Ergebnis liegt nicht in ihrer Hand. Bei der Wahl des Schmuckstücks sind es zwei Knöpfe. Hier treffen die Spieler eine Entscheidung. Konstanten und Variablen durchdringen sämtliche Ebenen des Computerspiels, bis hin zu Vertonungsprozessen. Ein Ereignis kann z. B. konstant vertont sein, indem ihm ein einziges akustisches Element zugeordnet wird. Mit Hilfe einer Klanggruppe oder der Modifikation akustischer Parameter entstehen variable Vertonungen. Wie die Million Millionen Welten ist die akustische Gestaltung mehrerer Spieldurchläufe ebenfalls immer verschieden und immer ähnlich.

Aufgrund dieser Komplexität - den Million Millionen Welten verschiedener Spieldurchläufe - müssen neue Zugangsweisen und kritische Werkzeuge entwickelt werden, die das Computerspiel nicht nur in der Tradition des Films, Romans oder Gesellschaftspiels verorten, sondern es als Ausdruck einer Software-Kultur verstehen. Zwar treten im Computerspiel Darstellungstypen und Ausdrucksweisen aus unterschiedlichen Zeiten und Traditionen zusammen. Dazu gehören Erzählweisen, Spielkonzepte, Animationen und Grafiken sowie aufwändig produzierte Klangräume, Soundeffekte und Musik. Hergebrachte Erklärungsmodelle einzelner Komponenten sind jedoch nicht hinreichend, um das gesamte Möglichkeitsfeld zu fassen. Aus dem Amalgam verschiedenartiger kultureller Artefakte entstehen ständig neue Formen. Es bedarf neuer Differenzierungsmerkmale, um diese Resultate einer kontinuierlich fortschreitenden Medienhybridisierung weiter zu untersuchen. Ein grundlegendes Differenzierungsmerkmal dieser Art ist die Unterscheidung zwischen Konstanz und Variabilität. Dass *Bioshock Infinite* diesen Aspekt selber thematisiert und auslotet, ist ein weiteres Beispiel für die Fähigkeit des Computerspiels zur Selbstreflexion.

Die in dieser Arbeit skizzierte akustische Architektur des Computerspiels bietet eine Reihe von Anknüpfungspunkten, um die Ausdrucksfähigkeit akustischer Elemente in diesen Formen weiter zu untersuchen. Einige dieser Anknüpfungspunkte docken an die Logik des Spiels an, andere an die digitalen Repräsentations- und multimedialen Ausgabeformen des Computers. Beide spielen auch außerhalb des Computer-

spiels eine zentrale Rolle. Einerseits durchdringt das Spiel immer mehr Bereiche. So gilt z. B. das *Newsgaming*, die Aufbereitung aktueller Geschehnisse in einer Art spielbaren Reportage, als vielversprechendes Werkzeug journalistischer Praxis.⁴⁰ Andererseits steigt die Zahl der verfügbaren Daten in allen Lebensbereichen an⁴¹, wodurch Fragen rund um die Repräsentation und Darstellung dieser Daten eine größere Bedeutung annehmen.

Der Architektur-Ansatz ist aber nicht nur für die Berücksichtigung der akustischen Dimension nützlich. Vielmehr lassen sich durch die Betrachtung des Spiels als etwas im Bau befindliches, und den Verknüpfungen zwischen Darstellungselementen und den darunter liegenden Modellen, wahrnehmbare und verborgene Konstituenten gemeinsam in den Blick nehmen. Die gesamte Darstellung des Computerspiels - Sound, Grafik, Text und Vibrationen durch Force Feedback - ist in einem Spannungsfeld zwischen Ästhetik und Funktionalität sowie Randomisierung und Determiniertheit verortet, das sich vor dem Hintergrund der Prozeduralität ausführbaren Codes und der interaktiven Spielsituation entfaltet. Künstlerische Gestaltungsprozesse und die damit verbundenen gestalterischen Freiheiten begegnen dem möglichst eindeutigen, auf Konventionen beruhenden Interaktionsdesign der Mensch-Computer-Schnittstelle. In den Game Studies sind diese Aspekte bisher unterrepräsentiert.⁴²

Eine Vorgehensweise, die Rolle von Darstellungselementen in Software zu erfassen, besteht darin, in einem ersten Schritt ihre Kopplung an Ereignisse und Zustände des Programms nachzuvollziehen. Die Unterscheidung zwischen zustandsvarianten und -invarianten Elementen schafft dabei ein grundlegendes Verständnis für die Gestaltung im Interaktionskontext. Kapitel 3 und Kapitel 4 zeigen exemplarisch, wie sich für den zeitlichen und räumlichen Aufbau des Spiels ein Zugang zu dem Arrangement aus Zeichen und den gegebenenfalls dahinter verborgenen Zahlen schaffen lässt. So entstehen beispielsweise in zeitkritischen Spielsituationen vollkommen andere Voraussetzungen für die funktionalen

⁴⁰Siehe (Bogost u. a., 2012) für eine Einführung. Weitere Ansätze, mit Hilfe von Spielprinzipien Anreize für die Nutzung eines Angebots zu schaffen, sind weit verbreitet und werden zur Zeit unter dem Begriff *Gamification* diskutiert.

⁴¹Der weltweite Trend zu immer umfangreicheren Datenmengen wird derzeit unter dem Begriff *Big Data* diskutiert.

⁴²Siehe (Picard, 2009, 334). Sowohl in narrativ orientierten als auch in ludisch orientierten Ansätzen werden die Darstellung und damit verbundene ästhetische Fragestellungen größtenteils außen vor gelassen.

Anbindungsmöglichkeiten akustischer Elemente als in Spielsituationen, die den Spielern beliebig viel Zeit einräumen. Akustische Elemente sonifizieren hier Zeitpunkte, Zeitspannen und Geschwindigkeiten, fungieren aber auch als Brücken zu Aspekten der Erzählung. Darüber hinaus sind unterschiedliche Formen der Auseinandersetzung mit diesen Arrangements zu berücksichtigen. Computerspiele schaffen Angebote, die auf mehrere Arten und Weisen genutzt werden können - von dem möglichst effizienten Umgang mit einem regelbasierten System bis hin zur Versenkung in einer detailliert gestalteten fiktiven Welt.

Während der Schwerpunkt dieser Arbeit auf akustischen Elementen als Teil der Ausgabe lag, operieren Computerspiele mittlerweile auch vermehrt mit akustischen Eingaben - seien es Gesangsspuren in Musikspielen, Sprachbefehle, stimmliche Artikulationen, oder Atemgeräusche.⁴³ Auch Eingaben lassen sich als Bestandteil der akustischen Architektur des Computerspiels verstehen. Ähnlich wie bei akustischen Bausteinen der Ausgabe lässt sich ihre Kopplung mit Ereignissen und Zuständen des Spielmodells analysieren. Im Gegensatz zu Tastatureingaben oder der Bewegung einer Maus sind akustische Eingaben jedoch komplexer. Während der Tastendruck ein klar definiertes digitales Signal an die Maschine sendet, muss die akustische Eingabe durch Verfahren wie das Sampling erst in eine digitale Form überführt werden. Anschließend stellt sich die Frage, welche Aspekte der akustischen Eingabe das Programm erfasst und im Kontext des Spielmodells weiterverarbeitet, und welche Aspekte es ausblendet.

Bei einem Musikspiel wie *Singstar* ist es beispielsweise ein entscheidender Unterschied, ob das Programm lediglich die Tonhöhe der eingesungenen Spuren berücksichtigt, oder ob es auch eine Spracherkennung durchführt und die Textsicherheit der Spieler überprüft. In beiden Fällen bewertet das Programm die akustischen Eingaben unterschiedlich und legt dadurch unterschiedliche Spielweisen nahe. Wenn das Spiel auf den eingehenden Audiostrom lediglich eine schnelle Fouriertransformation anwendet und die daraus gewonnenen Werte mit hinterlegten Werten - dem korrekten Tonhöhenverlauf der Gesangsspur - vergleicht, so können die Spieler den Text vernachlässigen. Anstatt das Stück in

⁴³Einige Nintendo DS Spiele nutzen das eingebaute Front-Mikrofon, um die Spieler mit ihrer Atemkraft Gegenstände bewegen zu lassen. So können sie z. B. gegen grafisch dargestellte Ballons pusten, die gemäß der physikalischen Simulation des Programms ihre Position verändern. Das Programm misst dafür die Intensität des Rauschens, das durch das Pusten ins Mikrofon entsteht.

Karaoke-Manier zu singen, summen sie vielleicht nur die Melodie in das Mikrofon und erhalten trotzdem die volle Punktzahl. Gerade in kompetitiven Spielsituationen könnte diese Spielweise für Kontroversen sorgen. Besteht das Ziel des Spiels darin, den jeweiligen Song möglichst akkurat nachzusingen oder darin, die enger gestellte Aufgabe zu erfüllen, d.h. die Tonhöhen exakt zu reproduzieren?

Das Beispiel zeigt: Bereits die Selektion der Algorithmen, die auf akustische Eingaben angewandt werden, kann weitreichende Folgen auf den Umgang mit dem Spiel haben.⁴⁴ Ebenso wie die Spieler zwischen zustandsinvarianten und zustandsvarianten Darstellungselementen differenzieren, unterscheiden sie spielrelevante Aspekte der Eingabe von Aspekten, die für den Spielausgang unerheblich sind. Das Programm favorisiert bestimmte Eingabeformen und -muster, indem das Spielsystem Bewertungskriterien beinhaltet und die Eingaben in Punkte, Erfolge und Niederlagen translatiert. Welche Eingaben das Programm besser, schlechter, oder überhaupt bewertet, loten die Spieler während des Spiels kontinuierlich aus.

Akustische Eingaben bringen aber nicht nur neue Herausforderungen für Entwickler und Spieler mit sich. Sie prägen auch das Verständnis des Computerspiels als Ganzes. Je stärker das Computerspiel mit Eingaben wie stimmlichen Artikulationen operiert, desto offensichtlicher positioniert es sich als interaktives Medium. Lässt sich ein Spiel, das durch Maus, Tastatur oder Gamepad bedient wird, von außenstehenden Beobachtern noch fälschlicherweise auf seinen audiovisuellen Part reduzieren, wird die immer noch weit verbreitete Deutung des Computerspiels als eine Art Film durch akustische Eingaben vollends unplausibel. Anstelle der audiovisuellen Präsentation über Bildschirm und Lautsprecher rückt der Dialog zwischen Spielern und Spiel in den Fokus. Spielern ist dieser Dialog zwar selbstverständlich, doch außenstehenden Beobachtern bleibt er häufig verschlossen.

Ebenso wie akustische Eingaben die Bedeutung der Präsentation des Spiels relativieren, legen sie ein umfassenderes Verständnis des Spielraums nahe. Nicht nur die Spielwelt, die sich auf dem Bildschirm eröff-

⁴⁴Selbst ohne Spielsystem im Hintergrund legen Programme, die akustische Eingaben weiterverarbeiten, einen spielerischen Umgang nahe. Apps zur automatischen Erkennung von Songs und Interpreten laden die Nutzer z. B. dazu ein, zu experimentieren. Welche Stücke erkennen die Dienste auf welche Weise und welche erkennen sie nicht? Genau die Neugierde darüber, wozu das technische System in der Lage ist, prägt auch den Umgang mit Musikspielen wie *Singstar*.

net, ist relevant. Vielmehr treten virtueller und realer Raum in ein Verhältnis, das durch gegenseitige Wechselbeziehungen geprägt ist. Schallereignisse im realen Raum führen potentiell zu Veränderungen des virtuellen Raums. Die im virtuellen Raum dargestellte Spielsituation motiviert die Spieler dazu, auf bestimmte Weise innerhalb des realen Raums zu agieren. Je stärker sich Szenarien einer umfassenderen Computerisierung wie das von Mark Weiser skizzierte *ubiquitous computing*⁴⁵ durchsetzen, desto wichtiger werden die räumlichen Konstellationen, in denen sich das Spiel vollzieht. Das Computerspiel der Zukunft würde sich nicht mehr vorrangig auf Projektionsflächen wie Bildschirmen und Lautsprechern abspielen, die klar als Tore in virtuelle Welten erkennbar sind. Es wäre vielmehr eine besondere Spielart in komplett computerisierten Umgebungen, die sich aus unzähligen Mikrointeraktionen zusammensetzt.

Weiterführende Studien könnten den wachsenden Bestand neuer Ein- und Ausgabegeräte in den Blick nehmen. So versprechen vor allem die Korrelationen zwischen Eingaben und Ausgaben unterschiedlicher Darstellungsweisen - ein gesprochenes Wort führt zu einer Animation, eine Geste verändert akustische Parameter - Einblicke in die Interaktionszusammenhänge und Transformationsmöglichkeiten bausteinartiger Ausdrucksmittel. Die Transformation der Eingaben in anders geartete Ausgaben verweist auf ein Phänomen digitaler Ausdrucksmittel, das nicht nur im Kontext des Computerspiels Bestand hat, hier aber in besonders geballter Form hervorsticht. Wie in Kapitel 6 angerissen, büßen Ausdrucksformen im Digitalen ihre ursprünglich medienspezifischen Charakteristika ein. Sie sind nicht mehr an ihre materiellen Produktionsmittel gebunden. Text erfordert keine auf Papier gedruckten Lettern, Bewegtbild keine belichteten Filmstreifen, gesprochene Sprache keine Stimme. Stattdessen schafft der technische Apparat des Computers eine universelle Plattform, auf der mediale Inhalte in Form von digitalen Repräsentationen vorliegen. An die Stelle ursprünglich physischer Bedingungen treten Verfahren der Informationsverarbeitung wie Dateiformate, Codecs und Filter. Sämtliche Daten und Anweisungen rund um die Aufbereitung und Wiedergabe dieser Daten fließen potentiell ineinander. Sie sind offen für Querverbindungen und Transformationen. Besteht die Eingabe aus ein paar gesprochenen Sätzen, könnte das Programm das Gesagte beispielsweise beibehalten und mit einer anderen, synthetisch erzeugten Stimme ausgeben. Es könnte auch einzelne

⁴⁵siehe (Weiser, 1991)

Wörter austauschen oder für bestimmte Schlüsselwörter Bilder einblenden oder Soundeffekte abspielen. Der Übersetzungsprozess ist dabei einerseits von den Regeln für den Austausch abhängig und andererseits durch die Wissensbasis des Programms bzw. dessen Rückgriffe auf externe Datenbanken geprägt. Diejenigen Transformationen, die von den Spielern kontrollierbar sind, fallen in den Bereich der Konfiguration.

Trotz der nahezu grenzenlosen Potenzialität medialer Transformationen und Hybridbildungen sind die ursprünglichen Vorbilder digitaler Repräsentationen im Gros der heutigen Medienlandschaft eindeutig erkennbar - seien es Bücher, Karten, Filme oder Alben. Digitale Repräsentationen erfordern und ermöglichen zwar andere Formen des Zugriffs, der Distribution und der Anreicherung mit ergänzenden Inhalten – man vergleiche beispielsweise die klassische Übertragung eines Fußballspiels auf einem Fernseher mit einem Live-Stream auf einem mobilen Endgerät, der durch Kommentare aus Social Media Plattformen flankiert ist. Dennoch blieben die Repräsentationen akustischer, visueller und textueller Inhalte aus struktureller Sicht ihren historischen Vorgängern verpflichtet. E-Books und MP3-Alben sind lineare Datensätze, die Büchern und CDs nachempfunden sind. Algorithmische Kompositionen hingegen, die bei jedem Abspielvorgang neue Varianten hervorbringen, haben sich als eigenständige Ausdrucksformen kaum in bemerkenswertem Umfang durchgesetzt - obwohl heutige MP3-Player und Smartphones ohne Weiteres dazu in der Lage wären, Zufallsprozesse einzubeziehen. Gerade hier zeigt sich einer der bedeutendsten Aspekte des Computerspiels als Ausdrucksform. Es bündelt von algorithmischen Kompositionen bis hin zu dreidimensionalen, in Echtzeit simulierten Szenarien die Erzeugnisse einer Software-basierten Kulturproduktion – und stellt die daraus entstehenden interaktiven Arrangements unter der Ägide des Spiels einer breiten Masse an Usern zur Verfügung.

Literaturverzeichnis

- [Aarseth 1997] AARSETH, Espen J.: *Cybertext: Perspectives on Ergodic Literature*. Johns Hopkins University Press, 1997
- [Bagnall 2011] BAGNALL, Brian: *Commodore: A Company on the Edge*. Revised. Variant Press, 2011
- [BIU 2011] BIU, e.V.: *Gamer in Deutschland*. 2011
- [Boer 2002] BOER, James: *Game Audio Programming*. Charles River Media, 2002
- [Bogost 2007] BOGOST, Ian: *Persuasive Games: The Expressive Power of Videogames*. MIT Press, 2007
- [Bogost u. a. 2012] BOGOST, Ian ; FERRARI, Simon ; SCHWEIZER, Bobby: *Newsgames: Journalism at Play*. MIT Press, 2012
- [Böhn 2010] BÖHN, Andreas (Hrsg.): *Techniknostalgie und Retrotechnologie*. Kit Scientific Publishing, 2010
- [Booth 2009] BOOTH, Michael: *The AI Systems of Left 4 Dead*. http://www.valvesoftware.com/publications/2009/ai_systems_of_l4d_mike_booth.pdf. Version: 2009. – abgerufen am 07.08.2013
- [BPjM 2011] BPjM: Doom aus der Liste der jugendgefährdenden Medien gestrichen - BPjM-Entscheidung Nr. 5847 vom 4.8.2011 bekannt gemacht im Bundesanzeiger Nr. 131 vom 31.8.2011. In: *BPjM Aktuell* (2011), 3
- [Brandon 2007] BRANDON, Alexander: Audio Middleware - The Essential Link from Studio to Game Design. (2007). http://mixonline.com/basics/education/audio_audio_middleware/. – abgerufen am 13.02.2012

- [Bridgett 2010] BRIDGETT, Rob: *From the Shadows of Film Sound: Cinematic Production & Creative Process in Video Game Audio*. blurb.com, 2010
- [Castronova 2003] CASTRONOVA, Edward: On Virtual Economies. In: *Game Studies* 3 (2003), Nr. 2
- [Ceruzzi 1998] CERUZZI, Paul E.: *A History of Modern Computing*. MIT Press, 1998
- [Chion u. a. 1994] CHION, Michel ; GORBMAN, Claudia ; MURCH, Walter: *Audio-Vision*. Columbia University Press, 1994
- [Christian Höferle 2008] CHRISTIAN HÖFERLE: Musikgames werden zum Standbein für Urheber. (2008). <https://www.gema.de/nl/092008/branchennews/musikgames.html>. – abgerufen am 07.07.2013
- [Chun 2011] CHUN, Wendy Hui K.: *Programmed Visions: Software and Memory*. MIT Press, 2011
- [Clark 2007] CLARK, Andrew: Defining Adaptive Music. (2007). http://www.gamasutra.com/view/feature/1567/defining_adaptive_music.php. – abgerufen am 07.08.2013
- [Commodore] COMMODORE: *C64 Bedienungshandbuch*, http://homepage.hispeed.ch/commodore/files/c64_de.pdf. – abgerufen am 30.05.2013
- [Crawford 2002] CRAWFORD, Chris: *The Art of Interactive Design: A Euphonious and Illuminating Guide to Building Successful Software*. No Starch Press, 2002
- [Crawford 2003] CRAWFORD, Chris: *Chris Crawford on Game Design*. New Riders Games, 2003
- [Dittbrenner 2007] DITTBRENNER, Nils: *Soundchip-Musik: Computer- und Videospielemusik von 1977-1994*. Universität Osnabrück, 2007
- [Dodgson u. a.] DODGSON, Harry ; WAGNER, Larry ; BENSEMA, Nick ; WILLIAMS, Roger: *AtariAge - Definitive Combat Dissassembly*. http://www.atariage.com/2600/archives/combat_asm/. – abgerufen am 07.08.2013

- [Drescher 2008] DRESCHER, Peter: Could Ringtones be more annoying?
In: COLLINS, Karen (Hrsg.): *From Pac-Man to Pop Music: Interactive Audio in Games and New Media*. Ashgate, 2008
- [Dyer-Witthford u. de Peuter 2009] DYER-WITHEFORD, Nick ; PEUTER, Greig de: *Games of Empire: Global Capitalism and Video Games (Electronic Mediations)*. University of Minnesota Press, 2009
- [Farnell 2007] FARNELL, Andy: An introduction to procedural audio and its application in computer games. (2007). <http://obiwannabe.co.uk/html/papers/proc-audio/proc-audio.pdf>. – abgerufen am 07.08.2013
- [Farnell 2010] FARNELL, Andy: *Designing Sound*. MIT Press, 2010
- [Fildes 2008] FILDES, Jonathan: *BBC NEWS | Technology | 'Oldest' computer music unveiled*. <http://news.bbc.co.uk/2/hi/7458479.stm>. Version: 2008. – abgerufen am 07.08.2013
- [Fuller 2008] FULLER, Matthew (Hrsg.): *Software Studies: A Lexicon*. MIT Press, 2008
- [Galloway 2006] GALLOWAY, Alexander R.: *Gaming: Essays On Algorithmic Culture*. University Of Minnesota Press, 2006
- [Gamasutra 2007] GAMASUTRA: *Road To The IGF: Audiosurf Making Waves With Music*. http://www.gamasutra.com/php-bin/news_index.php?story=16549#.UFAQJPKFL8s. Version: 2007. – abgerufen am 07.08.2013
- [Gamma u. a. 1994] GAMMA, Erich ; HELM, Richard ; JOHNSON, Ralph ; VLISSIDES, John: *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley Professional, 1994
- [Gee 2003] GEE, James P.: *What Video Games Have to Teach Us About Learning and Literacy*. Palgrave Macmillan, 2003
- [Görig 2009] GÖRIG, Carsten: Das kriegen wir schon klein. In: *GEE* (2009), Nr. 43, S. 74–77
- [Greenfield 2006] GREENFIELD, Adam: *Everyware: The Dawning Age of Ubiquitous Computing*. New Riders Publishing, 2006

- [Greening] GREENING, Chris: *Interview: The Orchestral Music and Interactive Sounds of Civilization V*. <http://www.squareenixmusic.com/features/interviews/michaelcurran.shtml>. – abgerufen am 26.03.2012
- [Grimshaw 2007] GRIMSHAW, Mark: *The Acoustic Ecology of the First-Person Shooter*, University of Waikato, Unpublished PhD thesis, 2007. <http://wikindx.com/mainsite/phd.html>
- [Hanson 2011] HANSON, Ben: *War Tapes: The Sounds Of Battlefield 3 - Features - www.GameInformer.com*. <http://www.gameinformer.com/b/features/archive/2011/02/28/war-tapes-the-sounds-of-battlefield-3.aspx>. Version: 2011. – abgerufen am 07.08.2013
- [Hiebert u. a. 2007] HIEBERT, Garin ; HARRISON, Peter ; PEACOCK, Daniel: *OpenAL Programmers Guide*. 2007. – abgerufen am 14.10.2012
- [von Hilgers 2000] HILGERS, Philipp von: Eine Anleitung zur Anleitung. Das taktische Kriegsspiel 1812-1824. In: *International Journal for the Study of Board Games cnws* (2000), Nr. 3, 59-77. <http://www.boardgamestudies.info/pdf/issue3/BGS3Hilgers.pdf>
- [Howard 2008] HOWARD, Jeff: *Quests: Design, Theory, and History in Games and Narratives*. A K Peters, 2008
- [Howells 2002] HOWELLS, Sacha A.: Watching a Game, Playing a Movie: When Media Collide. In: KING, Geoff (Hrsg.) ; KRZYWINSKA, Tanya (Hrsg.): *ScreenPlay*. Wallflower Press, 2002
- [Hsu 2010] HSU, Jeremy: *A Renaissance Scholar Helps Build Virtual Rome*. <http://www.livescience.com/8945-renaissance-scholar-helps-build-virtual-rome.html>. Version: 2010. – abgerufen am 20.06.2012
- [Jenkins 2006] JENKINS, Henry: *Convergence Culture: Where Old and New Media Collide*. NYU Press, 2006
- [Jones 2008] JONES, Steven E.: *The Meaning of Video Games: Gaming and Textual Strategies*. Routledge, 2008
- [Juul 2005] JUUL, Jesper: *Half-Real: Video Games Between Real Rules and Fictional Worlds*. MIT Press, 2005

- [Kastbauer 2011] KASTBAUER, Damian: *Limbo - Exclusive Interview with Martin Stig Andersen*. <http://designingsound.org/2011/08/limbo-exclusive-interview-with-martin-stig-andersen/>. Version: 2011. – abgerufen am 07.08.2013
- [Kent 2001] KENT, Steven: *The Ultimate History of Video Games*. Three Rivers Press, 2001
- [King u. Krzywinska 2006] KING, Geoff ; KRZYWINSKA, Tanya: *Tomb Raiders and Space Invaders: Videogame Forms and Contexts: Video Games in the 21st Century*. Tauris I B, 2006
- [Kirn 2009] KIRN, Peter: *Music, Physics, Space in Perfect Fusion: Interview, Creators of Game Osmos*. <http://createdigitalmusic.com/2009/09/24/>. Version: 2009. – abgerufen am 07.08.2013
- [Kitchin u. Dodge 2011] KITCHIN, Rob ; DODGE, Martin: *Code/Space: Software and Everyday Life*. The MIT Press, 2011
- [Kittler 1997] KITTLER, Friedrich: There is no software. In: JOHNSTON, J.H. (Hrsg.): *Literature, Media, Information Systems*. G+B Arts International, 1997 (Critical Voices Series)
- [Knuth 1974] KNUTH, Donald E.: Computer Programming as an Art. In: *Commun. ACM* 17 (1974), Nr. 12, S. 667–673
- [Kramer u. a. 1997] KRAMER, G. ; WALKER, B. ; BONEBRIGHT, T. ; COOK, P. ; FLOWERS, J. ; MINER, N. ; NEUHOFF: *Sonification report: Status of the field and research agenda*. 1997
- [Kushner 2004] KUSHNER, David: *Masters of Doom: How Two Guys Created an Empire and Transformed Pop Culture*. Reprint. Random House Trade, 2004
- [Langston 1989] LANGSTON, Peter: Six Techniques for Algorithmic Music Composition. In: *International Computer Music Conference (ICMC)* International Computer Music Association, 1989
- [Leonard 2003] LEONARD, Tom: *Gamasutra - Features - Building an AI Sensory System: Examining The Design of Thief: The Dark Project*. http://www.gamasutra.com/view/feature/2888/building_an_ai_sensory_system_.php?print=1. Version: 2003. – abgerufen am 19.01.2012

- [Levy 1984] LEVY, Steven: *Hackers: Heroes of the Computer Revolution*. Anchor Press Doubleday, 1984
- [Livingston 2007] LIVINGSTON, Jessica: *Founders at Work: Stories of Startups' Early Days*. Apress, 2007
- [Mackenzie 2006] MACKENZIE, Adrian: *Cutting Code*. Peter Lang Publishing, 2006
- [Manovich 2001] MANOVICH, Lev: *The Language of New Media*. The MIT Press, 2001
- [Manovich 2008] MANOVICH, Lev: *Software Takes Command (book draft)*. <http://www.softwarestudies.com/softbook>. Version: 2008. – Draft-Version vom 20. November 2008
- [Marks 2008] MARKS, Aaron: *The Complete Guide to Game Audio: For Composers, Musicians, Sound Designers, Game Developers*. 2. Auflage. Focal Press, 2008
- [Mäyrä 2009] MÄYRÄ, Frans: Getting into the Game - Doing Multidisciplinary Game Studies. In: WOLF, Mark J. P. (Hrsg.) ; PERRON, Bernard (Hrsg.): *The Video Game Theory Reader 2*. Routledge, 2009
- [McGonigal 2012] MCGONIGAL, Jane: *Reality is Broken*. Vintage, 2012
- [Miller 2009] MILLER, Kiri: Schizophonic Performance: Guitar Hero, Rock Band, and Virtual Virtuosity. In: *Journal of the Society for American Music* 4 (2009), S. 395–429
- [Montfort u. Bogost 2009] MONTFORT, Nick ; BOGOST, Ian: *Racing the Beam: The Atari Video Computer System (Platform Studies)*. MIT Press, 2009
- [Murray 1997] MURRAY, Janet H.: *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. Free Press, 1997
- [Myers 2010] MYERS, David: *Play Redux: The Form of Computer Games (Digital Culture Books)*. University of Michigan Press, 2010
- [Neeb 2009] NEEB, Christian: Reportage: Schuss-Wechsel. In: *GEE* (2009), Nr. 60, S. 68–72
- [von Neumann u. Morgenstern 2007] NEUMANN, John von ; MORGENSTERN, Oskar: *Theory of Games and Economic Behavior*. Commemorative edition. Princeton University Press, 2007

- [Newell 2008] NEWELL, Gabe: *Gabe Newell Writes for Edge*. <http://www.edge-online.com/opinion/gabe-newell-writes-edge>.
Version: December 2008. – abgerufen am 09.03.2012
- [Nishikado 2004] NISHIKADO, Tomohiro: Nishikado-san speaks. In: *Retro Gamer* 3 (2004)
- [Pfleger 2004] PFLEGER, Helmut: *Schach Zug um Zug*. 6. Auflage. Bassermann Verlag, 2004
- [Pias 2010] PIAS, Claus: *Computer Spiel Welten*. 2. Auflage. diaphanes, 2010
- [Picard 2009] PICARD, Martin: Video Games through Theories and Disciplines - Art and Aesthetics. In: WOLF, Mark J. P. (Hrsg.) ; PERRON, Bernard (Hrsg.): *The Video Game Theory Reader 2*. Routledge, 2009
- [Riemann 1884] RIEMANN, Hugo: *Musikalische Dynamik und Agogik*. D.Rather, 1884 http://openlibrary.org/books/OL24446531M/Musikalische_Dynamik_und_Agogik
- [Robinett 2006] ROBINETT, Warren: Adventure as a Video Game: Adventure for the Atari 2600. In: SALEN, Katie (Hrsg.) ; ZIMMERMAN, Eric (Hrsg.): *The Game Design Reader: A Rules of Play Anthology*. MIT Press, 2006
- [Slocum 2003] SLOCUM, Paul: *Atari 2600 Music And Sound Programming Guide*. http://www.qotile.net/files/2600_music_guide.txt. Version: 2003. – abgerufen am 07.08.2013
- [Spitzer 2002] SPITZER, Manfred: *Musik im Kopf: Hören, Musizieren, Verstehen und Erleben im Neuronalen Netzwerk*. Schattauer, 2002
- [Sterling 1993] STERLING, Bruce: *War is Virtual Hell*. http://www.wired.com/wired/archive/1.01/virthell_pr.html.
Version: 1993. – abgerufen am 07.08.2013
- [Thomsen 2010] THOMSEN, Michael: *How Limbo Came To Life*. <http://xboxlive.ign.com/articles/112/1120477p2.html>.
Version: Juli 2010. – abgerufen am 21.03.2012
- [Wardrip-Fruin 2009] WARDRIP-FRUIIN, Noah: *Expressive Processing: Digital Fictions, Computer Games, and Software Studies (Software Studies)*. MIT Press, 2009

- [Wark 2007] WARK, McKenzie: *Gamer Theory*. Harvard University Press, 2007
- [Weiser 1991] WEISER, Mark: The computer for the 21st century. In: *Scientific American* 265 (1991), Nr. 3, S. 66–75
- [Welsch 1998] WELSCH, Wolfgang: Wirklich: Bedeutungsvarianten - Modelle - Wirklichkeit und Virtualität. In: KRÄMER, Sybille (Hrsg.): *Medien, Computer, Realität: Wirklichkeitsvorstellungen und Neue Medien*. Suhrkamp, 1998
- [Weske 2000] WESKE, Jörg: *Digital Sound and Music in Computer Games*. <http://3daudio.info/gamesound/history.html>. Version: 2000. – Eingesehen am 01.02.2012
- [Wolf 2002] WOLF, Mark J. P.: *The Medium of the Video Game*. University of Texas Press, 2002
- [Wolf 2003] WOLF, Mark J. P.: Abstraction in the Video Game. In: WOLF, Mark J. P. (Hrsg.) ; PERRON, Bernard (Hrsg.): *The Video Game Theory Reader*. Routledge, 2003
- [www.scai.fraunhofer.de 2011] WWW.SCAI.FRAUNHOFER.DE: *Pressemitteilung: Rechnen auf mehreren Grafikkarten beschleunigt numerische Simulation um Größenordnungen*. [http://www.scai.fraunhofer.de/index.php?id=1568&no_cache=1&tx_ttnews\[year\]=2011&tx_ttnews\[month\]=01&tx_ttnews\[day\]=03&tx_ttnews\[tt_news\]=222&cHash=85beb93a7152958241b9f673df67d09c](http://www.scai.fraunhofer.de/index.php?id=1568&no_cache=1&tx_ttnews[year]=2011&tx_ttnews[month]=01&tx_ttnews[day]=03&tx_ttnews[tt_news]=222&cHash=85beb93a7152958241b9f673df67d09c). Version: 2011. – abgerufen am 07.08.2013
- [Yee 2006] YEE, Nick: The Labor of Fun. In: *Games and Culture* 1 (2006), Januar, Nr. 1, S. 68–71