

## 2. METAL GEAR SOLID und die sinnliche Ästhetik

---

Wenn die Spielenden in METAL GEAR SOLID V: THE PHANTOM PAIN (Konami 2015; im Folgenden zu MGS verkürzt) erstmals Spielfigur Snake<sup>1</sup> in die Nähe eines Gegners steuern, meldet sich Mentor Revolver Ocelot über Funk und warnt, Snake müsse sich aller Sinne bewusst sein, um unbemerkt voranzukommen. Ein ähnlicher Appell scheint für die Analyse des Computerspiels angebracht: Aufgrund ihrer multimedialen Komposition verlangen Computerspiele eine Herangehensweise, die verschiedene Sinneskanäle berücksichtigt und in der Interpretation integriert. Die Sinne strukturieren die menschliche Wahrnehmung und damit die ästhetische Komponente des Computerspiels. Die Bedeutung der Ästhetik aufgrund ihrer Brückefunktion zur Verbindung der Sphären Erzählung und Spiel macht eine genauere Betrachtung ihrer Funktionsweise notwendig. Der Begriff wird dabei in seiner griechischen Grundbedeutung von *aisthesis*, »Wahrnehmung«, verstanden (vgl. Nünning 2008, 4) und drückt damit eher die Phänomene einer generellen Betrachtung als die durch eine spezifische Theorie gefilterte Begriffsauffassung aus. Als Entsprechung zum »Wie« der Narratologie (vgl. Martinez und Scheffel 2005, 20) ist die Ästhetik eine formale Kategorie, die nach den Mitteln der Darstellung eines Inhalts fragt.<sup>2</sup> Sie strukturiert, wie Ideen und Konzepte über die Wahrnehmung transportiert werden. Die Wahrnehmung eines Computerspiels erfolgt, wie jede Wahrnehmung überhaupt, immer über die Sinne des Menschen. Es ist daher hilfreich, sich für die Strukturierung von Analyse und Interpretation auf

- 
- 1 Die Namensgebung für die Protagonisten in den verschiedenen Teilen der METAL GEAR Serie ist mitunter sehr verwirrend, auch weil die Erzählungen stets an verschiedenen Punkten der Gesamtchronologie stattfinden und die Verbindung zwischen Protagonisten der verschiedenen Teile durch Verwandtschafts- bzw. Klonungsverhältnisse bestimmt ist. Einigermassen konstant ist »Snake« als Teil des Codenamens, jeweils kombiniert mit einem Adjektiv zur genaueren Bestimmung (z.B. »Solid Snake« in METAL GEAR SOLID von 1998 und »Naked Snake« in METAL GEAR SOLID 3: SNAKE EATER von 2004). Angesprochen wird die Spielfigur in METAL GEAR SOLID V: THE PHANTOM PAIN (wie auch in anderen Teilen) einfach als »Snake«. Dieser Name wird daher auch hier verwendet.
  - 2 Ein für diese Arbeit entscheidendes allgemeines ästhetisches Merkmal ist die paradoxe Situation zwischen Vereinfachung und Erweiterung, die unter dem Stichwort »Miniaturisierung der Welterfahrung« an späterer Stelle erläutert wird.

die Sinneskanäle zu stützen, die das ansonsten unüberschaubar komplexe Gebilde Computerspiel in leichter zu erfassende Einheiten zerlegt. Verschiedene Medientypen nutzen verschiedene Sinneskanäle zur Vermittlung ihrer Informationen, und je mehr Sinne in die mediale Kommunikation involviert sind, desto komplexer ist das Unterfangen, die Art der Vermittlung analytisch zu strukturieren. Oft unternehmen Interpretierende automatisch eine Einteilung nach Sinneskanälen und betrachten systematisch beispielsweise das Bild und den Ton beim Film, um anschließend die Teilerkenntnisse zusammenzuführen. Das Computerspiel involviert nochmals weitere Sinneswahrnehmungen und steigert die Komplexität daher zusätzlich. Entsprechend hilfreich ist die systematische Gliederung in unterschiedliche Kanäle, auch als Modalitäten (im Englischen: »modes«) bezeichnet, die das Bewusstsein für die Bedeutung einzelner Komponenten eines Werks schärfen und eine genaue Identifikation von bedeutsamen Momenten erleichtern.

Für die meisten traditionellen Medien ist die Zuordnung zu bestimmten Sinneskanälen unproblematisch. Text beispielsweise ist dem Sehsinn zugehörig und wird ausschließlich über diesen konsumiert. Experimentelle Literatur kann durch andere Sinneseindrücke erweitert werden, aber der Kern bleibt ein optisches Medium.<sup>3</sup> Das akustische Gegenstück dazu ist die Musik, oder Tonkunst, wenn man noch allgemeiner bleiben möchte. Grundsätzlich ist Musik ein Medium, das ausschließlich über den Hörsinn wahrgenommen wird, obwohl je nach Argumentation auch hier beispielsweise der Tastsinn eine Rolle spielen kann (wenn Bässe durch ihre Vibrationen zu einem haptischen Erlebnis führen). Eine Musikdarbietung ist davon wiederum getrennt zu betrachten, hier kommt zusätzlich zur Musik ein optisches Element hinzu, wenn die Musiker oder Interpreten auf der Bühne stehen und sich auch visuell in Szene setzen.

Mit Hören, Sehen und Tasten sind auch bereits die drei medial wichtigsten Sinne beschrieben. Sinne wie Geruch oder Geschmack werden von Medien kaum bespielt, sofern man bei einem vergleichsweise engen Medienbegriff bleibt, der nicht zum Beispiel auch Speisen und Getränke einschließt. Es hat immer wieder Versuche gegeben, diese Sinne in andere Medienerfahrungen einzubeziehen, beispielsweise durch CDs oder Booklets, die mit einem Geruchslack beschichtet sind und ein von den Machern als passend empfundenes Aroma verströmen.<sup>4</sup> Solche Versuche waren aber bestenfalls experimentell, schlimmstenfalls bedeutungslos und üblicherweise ohne tiefere Verknüpfung mit dem zugehörigen Inhalt; eine Formspielerei ohne Funktion. Damit bleiben lediglich sinnliche Assoziationen wie beispielsweise der Geruch von Ölfarbe oder eines Holzrahmens bei der Malerei – bei-

3 Dass Text nicht wirklich dem Sehsinn, sondern einem anderen Kanal zugeordnet wird, wird an späterer Stelle im Detail entwickelt.

4 Das Album *TEENAGE DREAM* (Capitol/Virgin 2010) von Katy Perry beispielsweise duftet nach Zuckerwatte. (Vgl. Anderson 2010)

des Sinneseindrücke, die durch das gewählte Medium bedingt und im Normalfall nicht Teil des Werkes, sondern lediglich Eigenschaften des verwendeten Materials sind. Ganz ohne Relevanz sind diese Sinne für die Medienerfahrung allerdings nicht, können die assoziativen Kräfte des Menschen doch ebenfalls gewinnbringend eingesetzt werden, wie an späterer Stelle noch genauer auseinandergesetzt wird.

Die klassische Aufteilung in fünf oder sechs Sinne greift zu kurz für die Strukturierung einer Analyse, die einem modernen Medium wie dem Computerspiel gerecht werden möchte, zumal wenn zwei dieser Sinne nicht medial ausgewertet werden. Überhaupt gilt die Ansicht, der Mensch verfüge über lediglich fünf Sinne, längst als überholt. (Vgl. Bohlen 2017, 7) Die gemeinhin bekannten fünf Sinne Sehen, Hören, Tasten, Riechen und Schmecken werden in ihrer Existenz nicht bestritten, der Katalog der menschlichen Wahrnehmungskanäle wurde aber schon im 19. Jahrhundert immerhin um den Gleichgewichtssinn erweitert. (Vgl. Klinke 1995, 251) Ein endgültiger Konsens über die Anzahl der tatsächlich dem Menschen zur Verfügung stehenden Sinne ist bisher nicht erzielt worden. Uneinigkeit herrscht insbesondere zwischen der klinischen Psychologie und alternativen, esoterischen Ansätzen. Für die hier vorgestellte Anwendung zur Strukturierung von medialen Erlebnissen bieten allerdings beide Seiten nützliche Kategorien, weshalb die unterschiedlichen Vorschläge im Folgenden zusammengeführt und vereint werden. Als Quellen für die insgesamt vierzehn Sinneskanäle dienen dabei vorwiegend einerseits das Standardwerk der medizinischen Psychologie von Gerrig und Zimbardo (2008), sowie andererseits die Abhandlung über **DIE ZWÖLF SINNE DES MENSCHEN** von Rudolf Steiner (2017).<sup>5</sup>

Dabei besteht zwischen den Sinnen, die aus der medizinisch-psychologischen Literatur entnommen sind, und jenen, die auf dem esoterischen Ansatz Steiners basieren, ein wichtiger physiologischer Unterschied: Die medizinisch-psychologischen Sinne sind jeweils eigenen physiologischen Rezeptoren zugeordnet, während die zusätzlichen von Steiner beschriebenen Sinne sich aus Informationen dieser physiologischen Sinne speisen und sie weiterverarbeiten. Entsprechend trennt die Kategorisierung im Folgenden zwischen organischen und kognitiven Sinnen, oder auch zwischen primären und sekundären Sinneseindrücken. Diese Aufteilung findet sich in der Wahrnehmungsforschung gespiegelt in der Unterscheidung zwischen Empfindung und Perzeption:

Perception is different from sensation. Whereas sensation can be defined as the detection and processing of sensory information, perception is the act of interpret-

---

5 Die esoterischen Ausführungen von Steiner sind sicherlich medizinisch-psychologisch zumindest nach aktuellem Stand schwer haltbar. Sie beschreiben aber immerhin Phänomene der Wahrnehmung und können damit für die Analyse von Medienerfahrung hilfreich sein.

ting and organizing this sensory information to produce a meaningful experience of the world and of oneself.

(De Ridder et al. 2014, 5)

Primärsinne sind dann diejenigen Wahrnehmungskanäle, die physiologisch fixierbar und bestimmten Organen zuzuordnen sind: Sehen, Hören, Riechen, Schmecken, Gleichgewicht, Kinästhesie, Schmerz, Berührung und Temperatur. (Vgl. Gerrig und Zimbardo 2008, 108ff) Es sind diese Sinne, die selbständig Informationen von der Außenwelt in die Innenwelt des Individuums hineinragen und interpretierbar machen. Alle diese Sinne verfügen außerdem über eine integrierte Möglichkeit zu ebendieser Interpretation der Informationen, um den aufgegriffenen Reizen Bedeutung zuzuweisen; jedem sensorischen Kanal entspricht also ein perzeptueller Kanal.

Die Wahrnehmung der Welt erschöpft sich für den Menschen aber nicht in diesen neun Sinnen. Über die organischen Sinne hinaus existieren Interpretationen der Informationen aus den Sensoren der Primärsinne, die ebenfalls Teil der Perzeption sind, ohne einen eigenen sensorischen Apparat zur Verfügung zu haben. Daher röhrt die Hierarchisierung in der Bezeichnung der unterschiedlichen Sinnesarten: Die Sekundärsinne sind vollständig abhängig vom sensorischen Input aus den Primärsinnen und verfügen nicht selbst über einen Zugang zur Außenwelt. Gleichzeitig präsentieren sie sich dem inneren Empfinden gegenüber aber als Eindrücke von außen, sodass trotz fehlender Verbindung zum Außen doch der Eindruck einer Sinnesreizung entsteht. Diese Sekundärsinne werden in der medizinischen Psychologie als perzeptuelle Mechanismen beschrieben, während Steiner sie als eigene Sinne betrachtet (und ihnen eigene sensorische Kanäle zugesteht). Steiner identifiziert zwölf Sinneskanäle: Sehsinn, Geschmackssinn, Wärmesinn, Hörsinn, Riechsinn, Wortesinn, Tastsinn, Gleichgewichtssinn, Gedankensinn, Bewegungssinn, Ichsinn, Lebenssinn. (2017, 9ff) Eine Reihe von Sinnen in Steiners Liste findet sich auch im medizinischen Kontext: Sehen, Hören, Schmecken, Riechen, Tasten, Wärme, Gleichgewicht und Bewegung/Kinästhesie führen auch Gerrig und Zimbardo zumindest in ähnlicher Form, wenngleich anders kontextualisiert. (2008, 107ff) Damit bleiben vier von Steiners Sinneskanälen, die über das Medizinische hinausgehen: Wortesinn, Gedankensinn, Ichsinn und Lebenssinn. Der so entstehenden Gesamtheit aus neun medizinischen und vier esoterischen Sinneskanälen fügt das hier vorgestellte Modell noch zwei weitere Dimensionen hinzu: Erstens einen Zeitsinn, der die temporale Orientierung in der Realität ermöglicht und sowohl für erzählerische wie für spielerische Vorgänge eine zentrale Rolle spielt. Und zweitens einen Aktivitätssinn, der die Bewusstheit des Individuums umschreibt, eine bestimmte Reaktion der Außenwelt verursacht zu haben – eine Form der Wahrnehmung, die besonders für die Analyse von Computerspielen

mit ihren interaktiven Potentialen von ganz entscheidender Bedeutung ist. Damit wächst der Katalog auf vierzehn Sinne an.

*Tabelle: Sinnesmodalitäten im Überblick*

	Gerrig und Zimbardo	Steiner	Computerspiele
Primärsinn	Sehen	Sehsinn	Sehen
	Hören	Hörsinn	Hören
	Tasten	Tastsinn	Tasten
	Riechen	Riechsinn	Riechen
	Schmecken	Geschmackssinn	Schmecken
	Gleichgewicht	Gleichgewichtssinn	Gleichgewicht
	Kinästhesie	Bewegungssinn	Kinästhesie
	Schmerz		Schmerz
	Temperatur	Wärmesinn	Temperatur
Sekundärsinne		Wortesinn	Sprachsinn
		Ichsinn	} Identität
		Gedankensinn	
		Lebenssinn	Innensinn
			Zeit
			Aktivität

## 2.1 Das Tor zur Welt: Primärsinne

Neun Sinne bilden nach aktuellem Stand der Wissenschaft die Basis für die menschliche Wahrnehmung und damit für das Bild der Realität. Diese Sinne sind mit entsprechenden physiologischen Organen verbunden, die verschiedene äußere Reize wie Licht oder Luftdruck in elektrische Impulse umsetzen. Diese Impulse wiederum kann das Gehirn deuten und verarbeiten, um Informationen für die Wahrnehmung aufzubereiten. Die neun Primärsinne sind allerdings nicht gleichwertig für die Orientierung des Menschen in der Welt, sondern haben unterschiedliche Bedeutung und sind damit in ihrer Wichtigkeit auch für die Analyse und Auswertung von Medien unterschiedlich.

Der dominante Sinn in der menschlichen Orientierung ist der Gesichtssinn. (Vgl. Rock und Victor 1964, 594, sowie Heller et al. 1999) Der größte Teil der menschlichen Wahrnehmung (»über 90 % aller sensorischen Informationen« laut Wittling 1976, 9) läuft über den visuellen Sinneskanal ab, von der Navigation um Hindernis-

se herum bis hin zum Erkennen von vertrauten Menschen. Insofern verwundert es nicht, dass sich sowohl die Evolution als auch die Forschung besonders mit diesem Sinn auseinandersetzen, wie Gerrig und Zimbardo erläutern:

Die Sehfähigkeit ist die komplexeste und am höchsten entwickelte Sinnesmodalität beim Menschen, wie auch bei den meisten anderen mobilen Landwesen. [...] Dem Menschen hilft die Sehfähigkeit, Merkmalsveränderungen in der physikalischen Umwelt wahrzunehmen und das Verhalten entsprechend anzupassen. Die Sehfähigkeit ist die am besten untersuchte Sinnesmodalität.

(Gerrig und Zimbardo 2008, 119)

Die Dominanz des visuellen Kanals lässt sich bei Computerspielen deutlich beobachten: Der grafischen Umsetzung kommt in Spielrezensionen große Bedeutung zu, mitunter scheint sie einer der Hauptpunkte bei der Bewertung von neu erschienenen Spielen zu sein. (Vgl. Newman 2002 und Wolf 2003, 53) Das Computerspiel ist ein dominant grafisches Medium, weil Menschen dominant visuelle Lebewesen sind. In den meisten Fällen kommt daher auch in der Analyse von und kritischen Auseinandersetzung mit Computerspielen den optischen Aspekten eine besondere Bedeutung zu.

Die visuelle Gestaltung gliedert sich im Wesentlichen in zwei Aspekte: Perspektive und Stil. Die Perspektive beschreibt die Position der Spielenden in der virtuellen Welt, oft in Relation zur zentralen Spielfigur, während der Stil die vielen unterschiedlichen Möglichkeiten umfasst, diese virtuelle Welt zu gestalten. In der englischsprachigen Fachpresse hat sich für den Stil der Begriff des »artstyle« als Bezeichnung für die Gesamtheit der visuellen ästhetischen Entscheidungen durchgesetzt. Der zusammenfassenden Natur des Wortes gemäß gehört eine ganze Reihe an sich durchaus unterschiedlicher Elemente in diese Kategorie, wie beispielsweise Farbe, Licht, Proportion, Stilismus, Auflösung, Dimension. Sie alle beschreiben einen Teil dessen, was wir als komplettes Bild wahrnehmen und in den meisten Fällen auch intuitiv einordnen können. Unter Farbe beispielsweise fallen der allgemeine Farbeindruck sowie der Umfang des Spektrums. Manche Spiele nutzen satte, kräftige Farben mit einer reichhaltigen Palette unterschiedlicher Töne, manche Spiele beschränken sich auf einfache Primärfarbenkontraste, wieder andere gebrauchen nur stark eingeschränkte Paletten mit geringer Sättigung – die Militär-Shooter der 2000er Jahre kommen in den Sinn, die über Entwicklergrenzen hinweg erstaunlich einheitlich auf blasses Braun- und Grautöne setzen. (Vgl. Shepherd 2017) Für den Eingang in die Analyse muss der Aspekt der Farbe dabei nicht unbedingt einen Fixpunkt bilden; Farbinterpretation ist ohnehin ein schwieriges Feld, umso mehr, wenn neben der ästhetischen eventuell auch eine funktionale Ebene zu berücksichtigen ist. Eine Bedeutungszuschreibung an bestimmte Farben ist immer mit Vorsicht zu genießen, sind die Assoziationen mit unterschiedlichen Farben doch grundsätzlich von vielen Faktoren abhängig und vor allem kulturell geprägt. (Vgl.

Faulstich 2002, 152) Viel aufschlussreicher als der Absolutwert einer Farbe ist ohnehin ihr mimetischer Status, mithin die Aussage, die eine Farbwahl über das Verhältnis zwischen virtueller Welt und Vorlage trifft. Die eben erwähnte Dominanz von ausgewaschenem Braun bei Militärspielen wie CALL OF DUTY (Activision 2003) oder BATTLEFIELD (Electronic Arts 2002)<sup>6</sup> beispielsweise trifft kaum eine künstlerische Aussage, sondern begründet sich vor allem im außergewöhnlichen Erfolg des Genres und dem Wunsch der Entwickler, durch ästhetische Ähnlichkeit Spielende vom Kauf zu überzeugen. Die Mimesis wird hier zwischen verschiedenen Titeln eines Genres angesetzt; sie trifft keine Aussage darüber hinaus.

Die Farbpalette von MGS zeigt sich sehr ähnlich dem Modern Military Shooter, obwohl sich die generischen Überschneidungen in dieser optischen Ähnlichkeit erschöpfen. MGS ist ein »Stealth Action Game«, also ein Spiel, bei dem die Spielfigur möglichst ungesehen in Echtzeit durch eine von Feinden beherrschte Welt navigieren und Missionsziele erfüllen muss. Das Spiel ähnelt damit HITMAN, das ebenfalls oft mit diesem Genre in Verbindung gebracht wird; MGS geht aber nicht so weit, seine Level als Puzzleaufgaben zu konzipieren, sondern bleibt dem zeitkritischen Echtzeit-Gedanken enger verbunden. Trotz der mitunter bizarren Figuren und Handlungsstränge des METAL-GEAR-Universums vermittelt das Spiel eine weniger ludische und mehr an der Story orientierte Gewichtung und operiert mit einem insgesamt realistischeren Duktus, der sich eben in der Farbgebung realisiert findet. Durchbrochen wird die grundsätzlich realistische visuelle Gestaltung durch spielbezogene farbliche Markierungen, die Missionsziele, wichtige Ausrüstungsgegenstände wie das eigene Pferd oder eine mitgebrachte Waffe sowie die über ein Fernglas markierten Gegner hervorheben. Diese Informationen werden im Stile heutiger Augmented-Reality-Anwendungen (AR) als Folie über die virtuelle Welt gelegt und erweitern die »natürlichen« Sehfähigkeiten der Spielfigur entscheidend.

In HITMAN dagegen zeigen sich die Farben im Vergleich zwar immer noch recht nah an dem, was man in der realweltlichen Vorlage erwarten würde, allerdings stets etwas übersättigt, zu kräftig, zu klar. Damit unterstreicht das Spiel seine Ausrichtung als überlebensgroße Romantisierung eines an sich hässlichen und schmutzigen Geschäfts. Selbst, wenn 47 schmutzige Kleidung trägt, ist die Verschmutzung klar, deutlich, konturiert. Sie unterscheidet sich von der Realität, wie sich auch die Grenzen zwischen Gut und Böse in HITMAN viel deutlicher darstellen als in der Lebenswirklichkeit der meisten Menschen. Die Stilentscheidungen in der Gestaltung

---

6 In diesem Fall ist die Rede weniger von konkreten Iterationen dieser Spiele, sondern von der Serie (»franchise«) in ihrer Gesamtheit. Aus diesem Grund sind auch nicht die stetig wechselnden Entwickler angegeben, sondern die Publisher als Eigentümer der Namensrechte. Die Jahreszahlen beziehen sich auf die jeweils ersten Titel der Serien.

von HITMAN führen zu einer Distanzierung vom Geschehen, zu einem nicht aufdringlichen, aber stets präsenten Abstand von der Realität, der die Taten der Spielenden in der virtuellen Welt als unecht und überzogen markiert und in gewisser Weise moralisch relativiert. HITMAN ist kein Mordsimulator, sondern ein Spiel; es dient nicht der exakten Replik des Lebens eines Auftragsmörders, sondern einer sicherlich makabren, aber letztlich harmlosen Unterhaltung. Spielerische Funktion erhält die Farbgebung vor allem bei aktiverter »Intuition«: Die Spielwelt wird statt farbig in blassen Graustufen dargestellt und Wände werden durchsichtig. Missionsziele leuchten rot, Feinde, die eine Verkleidung durchschaut haben, gelb. Diese Darstellung entspricht weitgehend den AR-Informationen in MGS.

Es gibt selbstverständlich andere, visuell bedeutend auffallendere Spiele. So arbeitet beispielsweise HEAVY RAIN (Quantic Dream 2010) mit einem geradezu plakativen Wechsel der Farben zur Markierung einer emotionalen Veränderung in der Personenkonstellation der Handlung. Die Erzählung beginnt mit einer Familienidylle, die in kräftigen und abwechslungsreichen Farben schillert. Dann kommt ein Sohn von Spielfigur Ethan bei einem Unfall ums Leben, und die gesamte visuelle Gestaltung reflektiert diesen Einschnitt. Die Farbpalette schrumpft deutlich, die Sättigung der Farben schwindet. Der Gesamteindruck des Spiels wird grau, leb- und freudlos. Eine erneute Änderung gibt es erst am Ende des Spiels und auch dann nur unter bestimmten Umständen: Wenn die Spielenden durch ihre Entscheidungen im Spielverlauf das »gute« Ende erreichen, kehren Wärme und Sättigung in die Schlusssequenz zurück. Mit der Farbe ändert sich auch die Beleuchtung der Sequenzen: Aus den hell erleuchteten und durch mitunter markanten Schattenwurf als »sonnig« markierten Eröffnungssequenzen fällt die Handlung in einen entsättigten Farbbrei, in dem auch Licht und Schatten enger zusammenrücken. Die Gleichgültigkeit der depressiven Stimmung der Hauptfigur spiegelt sich in der angenäherten Äquivalenz von hellen und dunklen Stellen der Spielwelt.

Die Perspektive in MGS bietet zunächst keine besonderen Überraschungen und entspricht dem *third-person*-Standard, bei dem die virtuelle Kamera der Spielfigur folgt und ihr stets über die Schulter blickt.<sup>7</sup> Durch die externe Kamera ist es den Spielenden möglich, mehr von ihrer Umwelt wahrzunehmen, als es direkt durch die Augen der Spielfigur möglich wäre. Dadurch gleicht das Spiel Beschränkungen in den Feinheiten der Bewegung aus: Während ein echter Agent vorsichtig um eine Ecke oder über ein paar Mauertrümmer blicken könnte, bietet die Steuerung für Snake nur die Optionen Stehen, Hocken oder Liegen und damit im Zweifel nur die Wahl zwischen Deckung, aber keiner Sicht oder voller Sicht, aber ganz ohne Deckung. Die externe Kamera ermöglicht den vorsichtigen Blick um Hindernisse

<sup>7</sup> Zum Zusammenhang zwischen der Spielperspektive und dem Verhältnis zwischen Spielfigur und Spielenden folgen Ausführungen an anderer Stelle.

herum, ohne dabei die Spielfigur gleich vollständig den wachenden Blicken der feindlichen Figuren auszusetzen.

In HITMAN bietet die grundsätzlich ähnliche Perspektive ebenfalls einen spielerischen Vorteil, der sich inhaltlich aufgreifen lässt: Wenn Agent 47 sich beispielsweise an eine Wand drückt, können die Spielenden – wie bei MGS – aufgrund des externalisierten Blicks um die Ecke sehen, während diese Blickrichtung für die Spielfigur selbst völlig unmöglich ist. Das hat vorwiegend spielerische Bewandtnis, denn der Fokus auf einem lautlosen Vorgehen macht es immer wieder nötig, Wachen aus einem Hinterhalt heraus anzugreifen. Waren die Blicke von Spielfigur und Spielenden identisch, könnten die Spielenden eine sich nähernde Figur erst sehen, wenn diese den Blick erwidern kann – ein Alarm wäre fast unausweichlich. Durch den Trick der Perspektive dagegen können die Spielenden die Bewegungen der Spielfigur so exakt abstimmen, dass die feindliche Figur ausgeschaltet wird, ohne dass andere Figuren es bemerken.

Diese Fähigkeit findet sich in diesem Fall innerhalb der Welt von HITMAN auch erzählerisch legitimiert. Agent 47 ist, wie bereits in Kapitel 2 erläutert, ein genetisch manipulierter Klon, dessen körperliche Fähigkeiten über das hinausgehen, was normale Menschen vermögen. Bei der Gestaltung eines Spiels mit einer solchen Hauptfigur stehen Entwickler vor dem Problem, diese übersinnlichen Fähigkeiten einem Publikum von Spielenden zu vermitteln, die selbst höchstwahrscheinlich keine solchen Eigenschaften besitzen. Dieser kleine perspektivische Kunstgriff ist damit eine konsequente Weiterführung des Gedankens, den Fagerholt und Lorentzon zum Thema visuelles Interface entwickeln:

The most significant loss of perception that occurs when transferred into a virtual world is the broken perceptual link to the internal states of the player avatar. There is no possibility for players to make use of the perceptual mechanisms normally used to estimate ones [sic!] status of health, exhaustion etc. Further, basic internal mechanisms such as sense of speed or sense of height are reduced to interpreting information conveyed visually, auditory and haptically.

In order to create a more visceral game experience and hence strengthen the player immersion, UI designers must make up for this perceptual loss. Here, meta-perception (as previously presented) should be an appropriate tool in order to connect the player more closely to his or her in-game avatar.

(Fagerholt und Lorentzon 2009, 88)

Die Perspektive in HITMAN und METAL GEAR SOLID erledigt die Aufgabe der »Meta-Perzeption« ganz nebenbei, indem sie den Spielenden einen übernatürlichen Blick um die Ecke gestattet, der sich als die besondere Kraft einer gesteigerten Sinnesleistung und hochspezialisierten Intuition darstellt. Der Agent hört das Nahen der Schritte, ihre Vibration im Boden, er hört den Atem der feindlichen Figur und kann aufgrund seiner Ausbildung genau einschätzen, wann die Person genau in der rich-

tigen Position ist. Die Spielenden haben diese Fähigkeiten nicht, aber aufgrund der Wahl der Perspektive können sie trotzdem richtig reagieren, ganz so, als hätten sie selbst die übernatürliche Wahrnehmung. Gleches gilt für die Spielmechanik »Intuition«, mit der Agent 47 in HITMAN durch Wände blicken und die Position seiner Zielperson genau bestimmen kann. Diese Funktion ist nicht mimetisch im Sinne der Nachahmung einer wirklichen Wahrnehmungserfahrung der Spielenden; sie überträgt stattdessen die rein visuell nicht darstellbaren Fähigkeiten des Protagonisten auf die Spielenden.

Neben dem Schulterblick in METAL GEAR SOLID und HITMAN gibt es eine ganze Reihe anderer Möglichkeiten. Berühmt ist die Ich-Perspektive bei den entsprechend benannten Ego-Shootern, bei der die Blicke von Figur und Spielenden zusammenfallen.<sup>8</sup> Das Gegenstück im Hinblick auf die Figurenfixierung ist die Vogelperspektive, die zumeist in Spielen zum Tragen kommt, die auch spielerisch und inhaltlich nicht auf eine bestimmte Figur bezogen sind: Aufbau-Simulationen und Strategiespiele. Abgesehen von solchen weit verbreiteten Formen gibt es immer wieder auch experimentelle Perspektiven, die vor allem mit bestimmten Spielen assoziiert werden. So erzeugen fest eingestellte Kameras in den Räumen der Spielwelt das Gefühl, tatsächlich das Geschehen über fest installierte Kameras zu verfolgen. RESIDENT EVIL (Capcom 1996) ist berühmt für diese festen Perspektiven, die mit der Bewegung der Spielfigur auf die jeweils nächste Kamera umschneiden und durch ihre oft unvorteilhaften Blickwinkel zeitweise mehr verdecken als zeigen. (Vgl. Alvarez Igarzábal 2017, 224) In RESIDENT EVIL wird diese Technik wirkungsvoll eingesetzt, um den klaustrophobischen, bedrückenden Charakter der Situation (eingeschlossen in einem unheimlichen Herrenhaus voller todbringender Monster) und die zu diesem Inhalt passende B-Movie-Atmosphäre zu unterstreichen. Wie Horrorfilme, die den Zuschauer oft über Einzelheiten im Unklaren lassen und damit den Schauer verstärken, spielt auch RESIDENT EVIL auf diese Weise mit dem Grauen des Verborgenen. (Vgl. Burke 1989, 93)

Ein Nebeneffekt der ungewöhnlichen Blickwinkel in RESIDENT EVIL ist eine verstärkte Bedeutung des anderen dominanten Wahrnehmungskanals: des Ge-

---

8 Strittig ist, ob die Ego-Perspektive wirklich zu einer stärkeren Teilhabe am Geschehen auf dem Bildschirm führt. Auer schreibt: »Die Third-Person-Perspektive erzeugt eine gewisse Distanz zum Avatar« (2015, 21), Denisova und Cairns bestätigen, »that first person POV is more immersive than its third person counterpart« (2015, 147). Jeong et al. dagegen befinden: »However, there was no significant effect of the player's perspective on arousal.« (2008, 195) Die Validität solcher Studien ist immer ein schwieriges Thema, weil die Wirkung der Perspektive unter Umständen subtil ist und sich erst im Zusammenspiel mit anderen Faktoren zeigt. Denisova und Cairns bspw. nutzen für ihre Studie THE ELDER SCROLLS V: SKYRIM (Bethesda 2011), bei dem die Spielenden zwischen *first-person*- und *third-person*-Ansicht wechseln können. Allerdings gilt die Schulterperspektive dieses speziellen Titels als eher unausgereift, wodurch die Ergebnisse der Studie zumindest fraglich sind. (Vgl. Burgar 2019)

hörs.<sup>9</sup> Selbst Gegner, die von der Kamera noch nicht eingefangen werden, verraten sich oft bereits durch Schlurferäusche oder das charakteristische Stöhnen. So ist auch ohne Blickkontakt klar, dass sich in dem Raum eine Bedrohung aufhält, und die Spielenden können sich bereithalten. Der Gehörsinn ist dem Gesichtssinn in einer Hinsicht, der Ortung, überlegen und ergänzt das Sehen daher optimal. Gerrig und Zimbardo beschreiben Hören und Sehen als komplementäre Sinne, die sich gegenseitig ergänzen und erst in Kombination ihr gesamtes Potential entfalten.

Dem Hören und dem Sehen kommen komplementäre Funktionen in unserer Erfahrung der Welt zu. Wir hören oftmals einen Reiz, bevor wir ihn sehen, insbesondere wenn er sich hinter unserem Rücken oder auf der anderen Seite von undurchsichtigen Objekten wie beispielsweise einer Wand befindet. Obwohl unsere Sehfähigkeit besser als unsere Hörfähigkeit dazu geeignet ist, Objekte zu identifizieren, wenn sie sich erst einmal in unserem Sehfeld befinden, so sehen wir die Objekte oftmals nur deshalb, weil wir unsere Ohren dazu benutzt haben, um den Augen die richtige Richtung zu weisen.

(Gerrig und Zimbardo 2008, 129)

Die Komplementarität dieser beiden Sinne entsteht aus der Tatsache, dass das Gehör mit erstaunlicher Präzision Schallquellen im Raum lokalisieren kann und dabei ohne räumliche Grenzen agiert. Das Ohr kann mühelos Geräusche von den Seiten oder von hinten wahrnehmen und ihre Richtung identifizieren, also auch Ereignisse jenseits des Sichtbereichs. Verfügen Spielende über eine Möglichkeit, den in den meisten modernen Spielen angebotenen Surround-Sound wiederzugeben, lassen sich virtuelle Schallquellen aus allen Richtungen erzeugen.<sup>10</sup> Auf diese Weise können die Spielenden beispielsweise die Schritte von Feinden hinter sich hören und darauf reagieren. Eine entsprechend wichtige Position hat die klangliche Gestaltung von Computerspielen seit einiger Zeit. (Vgl. Ruf et al. 2019)

Crathorne attestiert dem Computerspiel drei akustische Kategorien: Musik, Klangobjekte/Effekte und Umgebungsgeräusche. (2010, 15) Damit verzichtet Crathorne auf eine wichtige vierte Kategorie, die Metasounds, die Jørgensen als »extra-diegetic sound« bezeichnet. (2006, 1) Damit hat der Ton in Computerspielen vier wesentliche Dimensionen: Sounds, Dialoge, Musik und Metasounds.<sup>11</sup> Sounds sind

9 Vgl. Lombard und Ditton 2006: »In general, our visual and aural senses dominate our perception [...]« (o.S.)

10 Aktuelle Weiterentwicklungen der Surround-Technik ermöglichen auch Schallquellen an der Raumdecke, beispielsweise bei Dolby Atmos.

11 Studien zur klanglichen Präsentation von Computerspielen sind vergleichsweise selten und qualitativ durchwachsen. So fassen beispielsweise Zhang und Fu sämtliche Formen von Tonereignissen in Computerspielen unter den Begriff »background music«, (2015, 1) wodurch feinere Unterscheidungen von Klangeffekten nicht mehr möglich sind.

die grundlegendste akustische Kategorie. Bestimmte Ereignisse erzeugen Sounds. Diese Erfahrung machen Menschen schon in frühester Kindheit und sie bleibt konstant, solange die Sinnesorgane ihre Arbeit tun. In Computerspielen lassen sich drei Arten von Sounds unterscheiden, von denen zwei aus dem Film bekannt sind: Atmos, Geräusche, Funktionsgeräusche. Eine Atmo (kurz für »Atmosphäre«) bezeichnet das zu erwartende Grundgeräusch eines Ortes, beispielsweise gedämpfter Straßenlärm in einer Stadtwohnung oder Vogelgezwitscher im Wald. (Vgl. Friedrich 2008, 46) Atmos sind ein Teil der Welt, sie transportieren im Idealfall eigentlich keine Information außer der, dass alles so ist, wie es sein sollte; damit sind sie erzählerische Sounds, die zur Glaub- und Lebhaftigkeit der Spielwelt beitragen. In der Tat kann man wohl behaupten, dass eine Atmo, die auffällig ins Bewusstsein tritt, ihre eigentliche Funktion nicht erfüllt. Das nämlich unterscheidet sie von den Geräuschen: wenn Snake die Schritte der Wachen hört, wenn sich Türen schließen oder wenn andere Geräusche die Tätigkeiten und Geschehnisse in der Spielwelt verraten. Atmos markieren Gleichheit, Geräusche dagegen markieren Verschiedenheit. Ein Geräusch zeigt eine Bewegung an, eine Veränderung in der Spielwelt, von der veränderten Position einer Wache (Schritte) über den veränderten Zustand eines Gegenstands (Tür schließt).<sup>12</sup> Geräusche können entsprechend erzählerische oder spielerische Funktion einnehmen, oder auch beides zugleich. (Vgl. Jørgensen 2006, 4)

Im Film sind damit alle akustischen Ereignisse abgedeckt, aber im Computerspiel ergibt es Sinn, noch eine weitere Kategorie zu unterscheiden: das Funktionsgeräusch. Wenn eine Wache sich bewegt, ist dieses Geräusch für andere Wachen erwartet; Schrittgeräusche von Snake dagegen können die Wachen alarmieren. Es gibt also einen entscheidenden Unterschied zwischen Geräuschen, die von anderen Figuren oder der Welt selbst erzeugt werden, und solchen, die von den Spielenden qua Spielfigur erzeugt werden. Gerade in METAL GEAR SOLID und HITMAN spielen diese Funktionsgeräusche eine zentrale Rolle bei der Lösung vieler Rätsel, weil sie die Aufmerksamkeit der anderen Figuren binden können. So kann ein geschickter Münzwurf in HITMAN die Blickrichtung aller Figuren in einem bestimmten Umkreis auf einen Punkt lenken, was einen unbeobachteten Korridor für 47 schafft, während Snake in MGS durch gezielte Klopfgeräusche Gegner in schwer einzusehende Winkel der Levelarchitektur locken und anschließend unbemerkt ausschalten kann. Das Funktionsgeräusch ist also ein spielergelenktes Geräusch und entsprechend zumindest tendenziell der spielerischen Sphäre zuzuordnen.

Das zweite klassische Audio-Element ist die Sprache. Unter diesen Begriff fällt im weiteren Sinne jede Form von gesprochenem Wort, also Dialoge, Monologe, Er-

<sup>12</sup> Faulstich (2002, 141) beschreibt Geräusche als »handlungsfunktional und häufig auch ideo-logierelevant«, unterscheidet aber nicht weiter zwischen Geräuschen und Atmos, vielleicht, weil diese Kategorien ineinander übergehen können.

zählerstimmen und mehr. Nicht jedes Computerspiel nutzt gesprochene Dialoge; besonders in vielen älteren Spielen oder in Spielen, die sich ästhetisch an diesen älteren Titeln orientieren, erfolgen Dialoge vorwiegend schriftlich. Daher bleibt die inhaltliche Komponente der Sprache einem anderen Sinn vorbehalten: dem Sprachsinn. Im Zusammenhang mit dem Gehör interessiert vor allem die klangliche Umsetzung der Dialoge.<sup>13</sup> Die akustische Komponente berücksichtigt dabei nicht den Inhalt des Gesagten, sondern den Ton. Das Timbre der Stimmen, ihre Sprachmelodie und Stimmung sowie parasprachliche Ausdrücke wie Ächzen, Schlucken oder schweres Atmen werden über die akustische Ebene kommuniziert und können schriftlich bestenfalls näherungsweise ausgedrückt werden. (Vgl. Mahl und Schulze 1982) Außerdem kann der Dialog natürlich technisch verfremdet werden, sodass bei Bedarf eine »Entmenschlichung« der Sprechenden akustisch umgesetzt werden kann. Eine Figur, bei der die Signifikanz des Sprachklangs besonders deutlich wird, ist Psycho Mantis im METAL-GEAR-Universum. In METAL GEAR SOLID V: THE PHANTOM PAIN bleibt die Figur ohne Stimme; bei ihrem ersten Auftritt in METAL GEAR SOLID (Konami 1998) dagegen monologisiert Psycho Mantis ausgiebig. Die Figur trägt eine spezielle Maske, und entsprechend klingt ihre Stimme gedämpft, gefiltert. Erzählerische Bedeutung erlangt dieser Stimmklang dadurch, dass Psycho Mantis mit enormen psychischen Fähigkeiten ausgestattet ist. Dass seine Stimme trotzdem durch die Maske erklingt, weist darauf hin, dass er entweder nicht zur gedanklichen Kommunikation mit Snake imstande ist oder sich bewusst dagegen entscheidet. Beide Alternativen können Auswirkungen auf das Gesamtverständnis der Figur haben. Ein anderer prominenter Stimmeffekt in MGS ist die Verzerrung der Stimme von Kontaktten, die per Funk mit Snake sprechen und die Stimmen damit klar von solchen unterscheiden, die in der unmittelbaren räumlichen Nähe von Snake erklingen.

In HITMAN bleibt der Funkkontakt zwischen Agent 47 und seiner Missionsplanerin Diana unverzerrt und ist nicht nur ungewöhnlich klar und frei von den Verfremdungen, die wir aus unserem telekommunikativen Alltag kennen, sondern anscheinend selbst der virtuellen Wirklichkeit enthoben. Es sind Situationen möglich, in denen 47 und Diana eine Unterhaltung führen, in denen also 47 aktiv spricht, während er sich gleichzeitig an eine ahnungslose Person heranschleicht. Selbst im Kontext des überlebensgroßen Agent 47 und seiner genetisch gesteigerten Leistungsfähigkeit wirkt eine Unterhaltung mit normaler Sprechstimme, die von Personen in Greifreichweite nicht gehört wird, unglaublich. Das Spiel bietet keine Erklärung für solche Situationen an.

---

13 Vertiefend zum Stimmklang im Computerspiel vgl. Stingel-Voigt 2019.

Ebenfalls in diesem Sinne unglaublich ist der Einsatz von Musik in MGS und HITMAN.<sup>14</sup> Allgemein lassen sich zwei Möglichkeiten zur Verwendung von Musik unterscheiden: Originalmusik und Lizenzmusik. (Vgl. Crathorne 2010, 18)<sup>15</sup> MGS nutzt Originalmusik, also Kompositionen, die speziell für das Spiel erstellt und aufgenommen wurden,<sup>16</sup> sowie an vereinzelten Stellen auch Lizenzmusik, die bereits bestehende, zumeist populärmusikalische Werke für die Verwendung im Spiel lizenziert.<sup>17</sup> Zu Beginn des Spiels ertönt beispielsweise das Stück »The Man Who Sold The World« in seiner Darbietung von Midge Ure aus dem Jahr 1982. Die Tatsache, dass es sich dabei um eine Coverversion eines im Original von David Bowie stammenden Songs handelt, gewinnt besondere Bedeutung, wenn sich am Ende des Spiels herausstellt, dass der Protagonist ebenfalls nicht das Original, sondern eine Kopie des eigentlichen Snake ist.

Entscheidend am Einsatz der Originalmusik in MGS ist, dass sie nur für die Spielenden existiert; weder Snake noch irgendeine andere Figur der virtuellen Welt können die musikalische Untermalung wahrnehmen; zumindest gibt es keine Ansätze, die Präsenz der Musik innerhalb der Diegese zu rechtfertigen, auch wenn sie nicht explizit geleugnet wird. Dabei ist der Spielverlauf nur selten nicht von Musik begleitet, wenngleich sich der Klangteppich zumeist hintergründig und zurückhaltend gibt und nur in spielerisch signifikanten Situationen anschwillt. Beim Erreichen wichtiger Missionspunkte wird die Musik präsenter und unterstreicht damit die Bedeutsamkeit des Geschehens. Besonders dynamisch zeigt sich ein solcher Musikeinsatz wiederum in HITMAN: Nach erfolgreichem Abschluss aller Missionsziele markiert Musik die Aufforderung zum Aufsuchen eines Ausgangs und steigert sich in Pegel, Arrangement und Vortragsdynamik, je näher die Spielenden die Spielfigur zu einem Ausgang bringen. Die Musik wird zum »Metonym des Spielfortschritts«. (Whalen 2004b, 14)<sup>18</sup> So erhält die Musik eine steuernde Funktion gemäß der Kategorisierung von Zofia Lissa (1965; zitiert bei Faulstich 2002, 144f; vgl. auch Krohn 2008).<sup>19</sup> Sie wird zum Hilfsmittel für die Spielenden, die

14 Eine vollständige musiktheoretische Auseinandersetzung mit der Musik eines Computerspiels kann im Rahmen dieser Arbeit nicht geleistet werden. Empfehlenswert ist die Arbeit von Brame (2009), während Tekrø (2018) die Musiktheorie außen vor lässt und Crathorne (2010) zwar musiktheoretisch fundiert vorgeht, dafür aber Aspekte und Erkenntnisse der Game Studies vernachlässigt.

15 Crathornes zusätzliche Unterscheidung zwischen orchesteraler und populärstilistischer Originalmusik ist hier nicht übernommen.

16 Das gilt teilweise auch für Musik, die theoretisch Lizenzmusik sein könnte, die also innerhalb der Diegese gehört wird.

17 Zur besonderen Wirkung von Lizenzmusik im Film vgl. Well 1976, 283f.

18 »Also, at the syntagmatic level, music often serves as a metonym for progress in the game.«

19 Auf eine genaue Auseinandersetzung mit der Kategorisierung von Lissa oder anderer ähnlich gelagerter Untersuchungen (Schneider identifiziert gleich 20 verschiedene Funktionen von Musik!) wird hier aus verschiedenen Gründen verzichtet. Zum einen beziehen sich diese

nicht immer sofort offensichtlichen Ausgangsmöglichkeiten schneller zu finden, und übernimmt damit eine klar spielerische Funktion.

Den meisten Spielenden ist eine solche musikalische Untermalung außerhalb der Erzählebene bekannt: Im Film ist die Nutzung dramatisierender Musik längst etabliert und tatsächlich so allgegenwärtig, dass eher ihr Fehlen (wie beispielsweise in ALIEN von 1979) auffällt. Während es also natürlich wahr ist, dass diese Form von Musik in der realen Erfahrungswelt keine Entsprechung findet, so handelt es sich doch um ein etabliertes Mittel akustischer Gestaltung im Film wie auch im Computerspiel.

Es gibt auch intradiegetische Musik in MGS, die in Anlehnung an Faulstichs »Musik im Film« als »Musik im Spiel« im Gegensatz zur »Spielmusik« bezeichnet werden kann. (Vgl. Faulstich 2002, 142) Diese Bezeichnung beschreibt wiederum Musik, die nun doch von den Figuren innerhalb der virtuellen Welt wahrgenommen wird – bei MGS sogar ganz explizit und spielmechanisch relevant. Gemeint ist Musik, die in der Spielwelt stattfindet, beispielsweise durch live spielende Musiker oder durch Radios. So kann Snake die Grenze zwischen Musik im Spiel und Funktionsgeräuschen verwischen, indem er ein Radio ein- oder ausschaltet, um eine Ablenkung für Wachen zu schaffen. Aufgrund der Tatsache, dass die Musik hier eine spielrelevante Rolle einnimmt, kann man argumentieren, dass es sich in diesem Fall nicht um Musik im Spiel handelt, sondern um ein Funktionsgeräusch, das eben zufällig musikalisch ist. Den Tanz auf dem Grat zwischen intra- und extradiegetischem Musikeinsatz wagt die bereits erwähnte Eröffnungssequenz von MGS. Snake greift nach einer Musikkassette, die mit dem Songtitel beschriftet ist, steckt die Kassette in ein Abspielgerät und drückt auf Play, woraufhin zunächst Kassettenknistern ertönt. Wenn die Musik und mit ihr die cineastische Eröffnungssequenz beginnt, ist von Kassetten-Artefakten keine Spur mehr; die Audioqualität ist makellos. Erst, wenn Snake im Krankenhaus zu Bewusstsein kommt, verändert

---

Untersuchungen auf den Film und müssten für die Anwendbarkeit auf Computerspiele ohnehin genau überdacht und gegebenenfalls adaptiert werden. Ein solches Vorhaben würde den Rahmen dieser Arbeit sprengen. Zum anderen ist fraglich, ob es einer so genauen Kategorisierung bedarf, wenn sich die Funktion eines Musikeinsatzes in den meisten Fällen auch intuitiv erfassen lässt, wenngleich dann sicherlich nicht mit derselben Tiefe wie in einer präzisen musiktheoretischen Analyse. Auch Faulstich gibt zu Bedenken, die Systeme von Lissa und Schneider würden nur vorgestellt, »um dazu anzuregen, *selber* film-, figuren-, szenen-, setting-, atmosphärenspezifische Funktionen der Musik zu ermitteln und sie produktbezogen jeweils zu begründen und in die Analyse und Interpretation einzubeziehen.« (Faulstich 2002, 147; vgl. auch Lissa 1965 und Schneider 1990, 90f) Für das Computerspiel gibt es Modelle zur Funktion der Musik von Whalen (2004a), d’Escriván und Collins (2011, 239), Moormann (2015, 138) und Jünger (2016, 20ff). Einige Studien postulieren als Hauptfunktion der Musik die Steigerung der Immersion und kommen zu der wenig überraschenden Erkenntnis, dass Musik die Immersion erhöhen kann. (Rogers et al. 2019; Sanders und Cairns 2010; Zhang und Fu 2015)

sich der Klang der Musik wieder und wird auch räumlich lokalisierbar: Die Musik dringt aus einem kleinen Radio im Zimmer. Diese Art der Darstellung ist wiederum nicht neu, sondern vom Kino geprägt, wo die Umdeutung extradiegetischer Musik zu von den Figuren der Erzählung konsumierter Radiomusik zu einem etablierten Stilmittel geworden ist.

Schließlich gibt es noch die Kategorie der Metasounds.<sup>20</sup> Damit sind Geräusche oder Klänge gemeint, die unmittelbar mit Handlungen im Spiel zusammenhängen, aber nicht innerhalb der Spielwelt ertönen, sondern nur für die Spielenden zu hören sind. Wenn Snake beispielsweise Munition auftreibt, ertönt ein deutliches Ladegeräusch, ohne dass Snake tatsächlich seine Waffe geladen hätte. Bewegen die Spielenden die Figur über eine Ressource, deren Kapazitätsgrenze in Snakes Inventar erreicht ist, wird dieser Umstand ebenfalls durch ein charakteristisches Geräusch kenntlich gemacht. Diese Signale haben keine tatsächliche Quelle innerhalb der Spielwelt, sondern dienen der funktionalen Kommunikation mit den Spielenden. Metasounds sind naturgemäß extradiegetische Elemente und damit der spielerischen Ebene zugehörig.

Auch der Tastsinn spielt für Computerspiele eine wichtige Rolle, allerdings auf andere Weise als Hören und Sehen. Während unsere beiden wichtigsten Sinne vorwiegend angesprochen werden im Sinne der Nachahmung von Ereignissen, mit deren Beschaffenheit wir aus realweltlichen Sinneserfahrungen vertraut sind, kommt dem taktilen Erleben eine Sonderrolle zu. Es mag Fälle geben, in denen der Tastsinn ebenfalls angesprochen wird; der Regelfall ist das aber sicherlich nicht. Vielmehr wird die Haptik in einer übertragenden, einer vermittelnden Rolle für das Computerspiel eingesetzt. Die Finger ertasten nicht die virtuellen Objekte, sondern das Steuergerät, über das Einfluss auf die virtuelle Welt ausgeübt werden kann. Der Tastsinn ist also ganz entscheidend für das Computerspiel, weil er instrumentell zum Ertasten des Werkzeugs eingesetzt wird, das die Kommunikation mit der virtuellen Welt ermöglicht. Er wirkt dabei zusammen mit dem kinästhetischen Sinn, der die benötigte Stellung der Finger für bestimmte Handlungen erfasst. Gregersen und Grodal sprechen vom »mapping« von Funktionen (2009, 69), die den Tastsinn weit über seine eigentliche Kapazität hinaus befähigen, auf eine Umgebung einzuwirken. Denn qua Controller ist er nicht mehr »nur« in der Lage, die Form und Oberflächenbeschaffenheit von Gegenständen zu erfühlen. Mapping ist in dieser Betrachtung durch Lernen und Automatisieren mehr als die Übertragung von Funktionen; komplexe Vorgänge wie das Einsteigen in ein Fahrzeug werden über eine einzige Taste auslös- und eben somit auch erfährtbar. Es werden also nicht nur abstrakte Steuerkonzepte auf Tasten verteilt, sondern durch die fühlbare Ergonomie eines Gamepads taktil erfahrbar gemacht. Schleicht sich Snake von

<sup>20</sup> Jørgensen (2006) spricht von »earcons« in Unterscheidung zu den intradiegetischen »auditory icons«.

hinten an einen Gegner heran, kann ein einziger Tastendruck einen Würgegriff auslösen, der anschließend durch ein weiteres Kommando erweitert werden kann und es den Spielenden erlaubt, die so eingefangenen Gegner zu verhören, bewusstlos zu machen oder zu töten. Diese komplexen Vorgänge werden jeweils von einer einzigen Taste verkörpert.

Die meisten Steuerbefehle ähneln sich von Spiel zu Spiel und zumindest zentrale Elemente werden vergleichsweise schnell standardisiert, sofern sie sich als ergonomisch sinnvoll erweisen. Verwirrung entsteht eigentlich nur dann, wenn neue Steuerelemente eingeführt werden und die Entwickler sich noch nicht darüber einig sind, wie diese Elemente am besten genutzt werden können. MGS nutzt beispielsweise die beiden Analog-Sticks, die fast alle aktuellen Spielcontroller zur Verfügung stellen, in etablierter Weise: Der linke Stick bewegt die Spielfigur durch die virtuelle Welt, der rechte Stick wendet den Blick. Diese Form der Steuerung wird so heute in vielen Spielen verwendet, für die sie anwendbar ist. Als die doppelten Analog-Sticks aber erstmals eingeführt wurden, hatte sich nachvollziehbarerweise noch kein Standard etabliert und die Entwickler experimentierten mit verschiedenen Möglichkeiten, sie einzusetzen. So gab es beispielsweise Rennspiele, die Gas und Bremse nicht auf die althergebrachten Oberseitentasten und ebenfalls nicht auf die heute etablierten Schultertasten legten, sondern auf den rechten Stick. (Vgl. *MIDNIGHT CLUB: LOS ANGELES*, Rockstar 2008)

Computerspiele und Technologien, die den Tastsinn konkret ansprechen, hat es allerdings auch immer wieder gegeben. Zum einen gab es Versuche mit optionalen Peripheriegeräten wie der KOR-FX Vibrationsweste, die eigentlich für die Rüttelmotoren im Gamepad gedachte Informationen ausliest und die Weste entsprechend vibrieren lässt. Auf diese Weise soll ein körperlicher Eindruck von Körpertreffern oder auch nahen Explosionen vermittelt werden. (Vgl. KOR-FX Herstellerwebsite) Bislang konnten sich solche Lösungen nur als Spielerei, nicht aber als echte Ergänzung zur Spielerfahrung durchsetzen. Etwas anders verhält es sich mit »HD Rumble«, einer Technologie, die Nintendo der Switch-Spielkonsole eingebaut hat. (Vgl. Nintendo Herstellerwebsite) Diese nutzt eine andere Art der Vibrationserzeugung als bisherige Controller und ermöglicht damit ein feineres, differenzierteres haptisches Feedback für die Spielenden. So lässt sich in manchen Spielen beispielsweise die Bewegung von virtuellen Murmeln auch erfühlen. Ob diese Technologie sich aber weiter verbreitet oder es bei einem einmaligen Versuch bleibt, lässt sich momentan noch nicht abschätzen. In jedem Fall zeigt diese Entwicklung aber, dass die Erschließung weiterer Sinneskanäle für das Medium Computerspiel noch in vollem Gange ist.

Nach der Orientierung auf dem Controller und der Erfassung von Erschütterungsfunktionen hat der Tastsinn in neueren Spielen noch eine dritte Funktion. Aktuelle Gamepads bieten üblicherweise drucksensitive Schultertasten, also Bedienelemente, die nicht nur eine binäre Unterscheidung zwischen Drücken und

Nichtdrücken erkennen, sondern auch den Grad des Drucks, seine Intensität erfassen. So kann in einem Rennspiel die Beschleunigungstaste graduell betätigt werden, um mehr oder weniger Gas zu geben. In HITMAN beispielsweise lässt ein halber Druck auf die Schusstaste Agent 47 den Atem anhalten, sodass für einige Sekunden die Zielpräzision steigt, bis mit dem vollständigen Durchdrücken der Schuss abgefeuert wird.

Schließlich gibt es noch eine Reihe von Primärsinnen, die bislang noch nicht oder zumindest kaum medial ausgewertet werden. Sicherlich gab es in die eine oder andere Richtung bereits Versuche, die schnell wieder aufgegeben wurden – Geruch und Geschmack beispielsweise eignen sich nicht für eine mediale Nutzung (vgl. Ryan 2001, 56), obwohl ihr Einsatz, wenn der notwendige Aufwand betrieben wird, positive Auswirkungen auf das Spielerleben hat. (Vgl. Lombard und Ditton 2006) Den Grund für die Nichttauglichkeit dieser Sinne sehen die Forscher in der schwer greifbaren Art der entsprechenden Reize (»lingering nature«; Ghinea und Ademoye 2010, 657), die daher röhrt, dass sie nicht physikalisch, sondern chemisch arbeiten. (Murray et al. 2013, o.S.) Zwei weitere bislang unerschlossene Sinne werden voraussichtlich auch weiterhin außen vor bleiben, weil die über sie wahrgenommenen Ereignisse sehr eng mit dem Wohlbefinden des Menschen verknüpft sind: Temperatur und Schmerz. Es wäre durchaus denkbar, über entsprechende Anzüge beispielsweise die Hitze eines Schauplatzes empfindbar zu machen. So gar die Instrumentalisierung von Schmerz ist technisch nicht unmöglich. Beides erscheint aber als Medienerfahrung wenig erstrebenswert.

Ein wenig anders verhält es sich mit den beiden übrigen Primärsinnen Gleichgewicht und Kinästhesie. Neben der beschriebenen Beteiligung des kinästhetischen Sinns am taktilen Erleben des Steuergeräts sowie mitunter recht erfolgreichen Implementierungen von Bewegungssteuerung in Konsolen (bspw. Nintendo Wii) hat vor einigen Jahren eine Technologie die Marktreife erlangt, die sich in der Science Fiction seit Jahrzehnten behaupten konnte: virtuelle Realität (VR) mit tragbarer Display-Technologie, sogenannten *Head-Mounted Displays*. Mit Oculus, HTC Vive und Playstation VR haben sich gleich mehrere Systeme im Massenmarkt positionieren können. Dabei spielt natürlich der optische Aspekt eine wichtige Rolle, da durch die doppelten Bildschirme und deren Nähe zu den Augen ein stereoskopisches, also dreidimensionales Umgebungs bild erzeugt werden kann. 3D-Optik ist allerdings auch durch andere technische Mittel umsetzbar. Die wirklich revolutionäre Besonderheit der VR-Systeme ist ihre Interaktion mit der Spielwelt. Sensoren erkennen die Position der Brille (und damit: des Kopfes) im Raum und reagieren auf Bewegungen. Ein Drehen des Kopfes zu einer Seite verändert also auch den dargestellten Bildausschnitt entsprechend, wodurch die Illusion entsteht, sich wirklich in der virtuellen Welt zu bewegen. (Vgl. Koenis 2019, 11) Durch die Kopplung der virtuellen Bewegung mit der tatsächlichen Position der Spielenden in der Realität und damit der Parallelisierung der jeweils wirkenden Kräfte werden der Gleichge-

wichtssinn und der kinästhetische Sinn in das Spielgeschehen involviert.<sup>21</sup> Isbister schreibt über bewegungsbasierte Steuermöglichkeiten: »game designers can now use players' bodies themselves as a medium for shaping emotions« (2017, 107). In Spielen, in denen man nicht nur die Spielfigur, sondern auch den Blick vergleichsweise frei bewegen kann, verliert man mit traditionellen Controllern schnell die Orientierung über die eigene Position im Raum. In einem System, in dem die echte Schwerkraft als ständige Kalibrierung dient, bleibt die Orientierung ohne Weiteres erhalten. Eine besonders anschauliche Beschreibung des Effekts stammt von Computerspiel-Journalist Rob Pearson, der seine Erfahrung mit der VR-Version von THE ELDER SCROLLS V: SKYRIM (Bethesda 2011; kurz SKYRIM) beschreibt:

Anyone who has played VR knows that it gives you a sense of scale of things that you just don't get when you're playing games traditionally, so you really can appreciate the size of the caverns you're in, the size of a mountain in Skyrim and the size of a giant frostbite spider.

(PlayStation Access 2017: 17:14)

Pearson beschreibt sein Erstaunen über die tatsächliche Größe einer riesigen Frostbissspinne (ein Gegner in SKYRIM), die ihm erst beim Spielen der VR-Version des Spiels wirklich bewusst wurde, obwohl er den Titel zuvor bereits unzählige Male gespielt hatte. Der entscheidende Punkt ist, dass sich solche Größenverhältnisse leicht verzerren lassen, wenn der Winkel der Kamera frei bestimbar ist und kein echter Referenzpunkt geboten wird. Ein Punkt X erscheint in der Mitte des Bildschirms; ob dabei die Kamera aber parallel zum Boden verläuft oder bereits in einem steilen Winkel nach oben gerichtet ist, lässt sich ohne Referenz nur schwer einschätzen. Mit der VR-Technologie ermöglichen der Gleichgewichtssinn und das Gefühl für die eigene Position im Raum eine präzisere Einschätzung des tatsächlichen Blickwinkels und damit eine »realistischere« Beurteilung von Größenverhältnissen. VR-Technologie ist also mehr als nur die Mimesis des Blicks, sie umfasst darüber hinaus Repräsentationen von Gleichgewicht und Kinästhesie.

Leider befindet sich die Entwicklung von Spielen für VR-Systeme noch ganz am Anfang. Zwar hat die Technologie sich mit einer gewissen Breitenwirkung durchgesetzt, der vergleichsweise hohe Anschaffungspreis stellt aber immer noch eine große Hürde für die »Massenvirtualisierung« dar. Der entsprechend noch recht kleine Markt rechtfertigt für Entwickler kaum die hohen Kosten von dedizierter VR-Spielentwicklung. Die beeindruckendsten Titel sind daher lediglich Portierungen, also Adaptierungen von bereits für Nicht-VR-Systeme erschienenen Spielen

21 Laurel (2016, o.S.) weist allerdings darauf hin, dass momentan die Bewegung der Spielfigur (controllergesteuert) noch nicht unabhängig von der Blickrichtung (brillengesteuert) möglich ist.

für VR. Entsprechend wird es wohl auch noch dauern, bis die Entwickler kreativ mit den Möglichkeiten der Technologie arbeiten und subversiv mit ihr umgehen.

Von den neun Primärsinnen lassen sich also fünf als Vermittlungsinstanzen im Computerspiel nutzen: Sehen und Hören als dominante Sinnesmodalitäten nutzen weitestgehend auch die klassischen Techniken und verbinden das Computerspiel am engsten mit seinen mediengenealogischen Vorgängern Fernsehen oder Musik. Dem Tastsinn, in klassisch rezeptiven Medien höchstens als peripherer Zusatzreiz beispielsweise durch die physische Beschaffenheit des Papiers eines Romans relevant, kommt im Computerspiel durch die Notwendigkeit einer taktilen Verbindung per Steuerapparat eine besondere Bedeutung zu. Abgesehen von kamera- oder sensorengestützten Bewegungssteuermodi orientiert der Tastsinn die Finger auf dem Controller, vom Gamepad über Tastatur und Maus bis hin zum spezifischen Modul. Er arbeitet damit eng mit dem kinästhetischen Sinn zusammen, der die notwendigen feinmotorischen Abläufe erfasst und in längerem Spielverlauf als Muskelgedächtnis automatisiert. Die Kinästhesie ist weiterhin für VR-Anwendungen relevant, für die schließlich auch der Gleichgewichtssinn unterstützend eingebunden werden kann (und willkürlich oder unwillkürlich wird). Geruch, Geschmack, Wärme und Schmerz werden entweder aufgrund ihrer technisch nur schwer umsetzbaren Funktionsweise oder aufgrund der für ein Produkt der Unterhaltungsindustrie wenig vorteilhaften Assoziationen zumindest bisher höchstens in experimentellen Zusammenhängen eingesetzt.

## 2.2 Sinnlichkeit jenseits des Sensorischen: Sekundärsinne

Die kognitiven Sinne sind in ihrer Funktion komplexer als die organischen Sinne, weil sie eben keine eigenen physiologischen Orte haben (auch wenn Steiner das behauptet; vgl. 2017, 11f), sondern Informationen der Primärsinne zweitverwerten oder auch Impulse aus mehreren Sinneskanälen zusammenführen. Dass solche Mechanismen möglich und auch immer noch der Sinneswahrnehmung zuzuordnen sind, deuten Gerrig und Zimbardo in der Erläuterung des kinästhetischen Sinns an:

Oftmals integriert das Gehirn Informationen des kinästhetischen Sinnes mit Informationen der Berührungssinne. Beispielsweise kann unser Gehirn nicht die volle Bedeutung der Signale von den einzelnen Fingern erfassen, ohne die genaue Lage der Finger zueinander zu kennen. Angenommen, Sie ergreifen ein Objekt mit geschlossenen Augen. Ihr Berührungssinn erlaubt Ihnen die Annahme, dass es sich um einen Stein handelt, aber erst Ihr kinästhetischer Sinn wird Ihnen Wissen darüber verschaffen, wie groß der Stein ist.

(Gerrig und Zimbardo 2008, 138)

Es ist also durchaus nicht ungewöhnlich, dass Reize von Sinnesrezeptoren abgegriffen und neben ihrer organisch-sinnlichen Verarbeitung auch zur Verfeinerung oder Erweiterung anderer Sinnesreize verwendet werden. Empfindung und Wahrnehmung sind, wie bereits erwähnt, nicht identisch; das eine ist die rein physische Aufnahme und Weiterleitung von Reizen, das andere ist ihre Interpretation und Kontextualisierung. Man könnte auch sagen, die betroffenen Rezeptoren dienen nicht beispielsweise als Rezeptoren des Tastsinns, sondern als Rezeptoren für mehrere Sinne zugleich. Was Rudolf Steiner beschreibt, bevor er in seiner Interpretation der Phänomene in wissenschaftlich schwer haltbare Esoterik übergleitet, sind Perzepte. Die medizinische Psychologie hat den Begriff »Sinn« für den Empfindungsapparat reserviert; wenn man aber seine Anwendung auf die Perzeptionsebene erlaubt, werden die von Steiner beschriebenen Empfindungen eben zu Sinneswahrnehmungen. Ein Eindruck, den Steiner auf seiner esoterischen Grundlage vorstellt und entsprechend benennt, ist der »Lebenssinn«:

Wann gehen wir am meisten in uns selber hinein? Nun, wenn wir im allgemeinen Lebensgefühl dasjenige wahrnehmen, was wir im wachen Zustande immer eben als unser Bewusstsein haben, dass wir sind, dass wir uns innerlich erfühlen, dass wir sind. Das wird vermittelt durch den Lebenssinn.

(Steiner 2017, 14)

Diese abstrakte Annäherung an ein vages Empfinden ließe sich etwas konkreter als ein »Innensinn« beschreiben, also die Wahrnehmung des eigenen inneren Zustands. In Anlehnung an den kinästhetischen Sinn, der letztlich die Position der Extremitäten im Raum protokolliert, bietet der Innensinn eine Bestandsaufnahme des Körperinneren. Da diese Modalität eher hypothetischen Charakter hat, lässt sie sich kaum als tatsächlichen Sinneskanal betrachten, außer in der eher vagen Form der Affiziertheit. In dieser Arbeit wird die Funktion des Innensinns daher so aufgefasst, dass das Ziel eines Titels die zumindest zeitweise Unterdrückung von Informationen aus dem Innensinn ist. Dann nämlich, wenn sich die Rezipierenden nicht mehr darüber im Klaren sind, was gerade in ihrem eigenen körperlichen Inneren geschieht, verlieren sie sich selbst in der Realität eines Titels. Die Empfindung des Innensinns, oder genauer die Nichtempfindung des Innensinns, ist ein Indikator für die Phänomene von Involvierungen, die in Kapitel 6 untersucht werden. Damit ist der Innensinn kein direkt ansprechbarer Kanal, sondern vielmehr eine Summe aus verschiedenen unspezifischen Empfindungen, die gemeinsam einen bestimmten Zustand hervorrufen können.

In engem Zusammenhang mit dem Innensinn lässt sich aus der Steinerschen Esoterik ein weiterer Sinn extrahieren. Steiner spricht davon, dass beim Verständnis anderer Menschen über den Sprachsinn hinaus das Begreifen der Gedanken und des Ichs der anderen Person notwendig sind: »Und es gehört, um die Gedan-

ken eines anderen Wesens wahrzunehmen, wiederum einfach ein anderes Sinnesorgan dazu, als es der bloße Wortesinn ist.« (Steiner 2017, 13) Und weiter:

Und auch das Ich ist noch nicht wahrgenommen, wenn man bloß die Gedanken wahrnimmt. Gerade aus demselben Grunde, warum wir abgesondert den Hörsinn vom Sehsinn statuieren, müssen wir, wenn wir auf die feineren Gliederungen der menschlichen Organisation eingehen, auch einen besonderen Ichsinn, einen Sinn für die Ich-Wahrnehmung statuieren.

(Steiner 2017, 13f)

Was Steiner unter »Gedankensinn« und »Ichsinn« fasst, wird hier kombiniert und als Beschreibung einer Erfahrung vorgeschlagen, die zentral für das Individuum ist: die Unterscheidung zwischen der eigenen Identität und der einer anderen Person. Es geht damit um die Fragen von Individualität, Identifikation und Empathie. Dieser »Identitätssinn« steht dem Innensinn durchaus nah; der Fokus beim Innensinn liegt allerdings stärker auf physiologischen Phänomenen, während der Identitätssinn stärker auf einem Konzept von Persönlichkeit beruht. Der Identitätssinn stellt die Frage: »Wer bin ich?« Gerade diese Frage wird im Zusammenhang mit der Präsenz qua Spielfigur im Computerspiel immer wieder neu verhandelt. Die Diskussion der entsprechenden Konzepte erfolgt in Kapitel 4.

Deutlich konkreter lässt sich fassen, was Steiner den »Wortesinn« nennt:

Es ist eine vollständige Verkennung, dass man glaubt, mit dem Gehörsinn sei auch schon dasjenige erschöpft, was der Wortesinn in sich enthält: [...] Es ist auch in Bezug auf die organische Gliederung ein Unterschied vorhanden zwischen dem bloßen Hören des Tones und dem Wortewahrnehmen. Das Hören des Tones ist vermittelt durch das Ohr, das Wortewahrnehmen ist durch andere Organe vermittelt, welche ebenso physischer Natur sind wie diejenigen, die den Gehörsinn vermitteln.

(Steiner 2017, 11f)

Wenngleich es bezeichnend für die mangelnde Wissenschaftlichkeit Steiners ist, dass er immer wieder wie selbstverständlich von »Organen« spricht, die bestimmte Sinne verkörpern, ohne aber diese Organe spezifizieren zu können, lässt die Neurolinguistik durchaus eine Art von sprachlichem Sinn vermuten. Es gilt heute als weitestgehend unstrittig, dass die Codierung und Decodierung von Sprache in Abschnitten des Gehirns abläuft, die nicht für auditive oder visuelle Signalverarbeitung zuständig sind, sondern eben für die spezifische Verarbeitung von Sprache, unabhängig von der Übertragungsmodalität – gesprochene und geschriebene Sprache wird im gleichen Teil des Gehirn verarbeitet. (Vgl. Friederici 1994; Price et al. 2003; Nitsch 2009; Müller 2013; Kemmerer 2015) Während die Rezeptoren, über die das Sprachsignal in die Verarbeitungskette gelangt, also den Gesichts- oder Gehörsinnen zuzuordnen sind (je nach Sprachmodus), wird anschließend nicht

etwa gesprochene Sprache im Gehörbereich des Gehirns und gelesene Sprache im Gesichtsteil des Gehirns verarbeitet; vielmehr werden die Daten aus beiden sensorischen Kanälen extrahiert und den jeweiligen Hirnbereichen für Grammatik und Lexikon zugeführt. (Vgl. Indefrey 2004, Friederici 2011 sowie Dietrich und Gerwien 2017 *passim*)

Inwiefern sowohl die akustische als auch die optische Sprachwahrnehmung tatsächlich dieselben Mechanismen nutzen, ist bislang nicht endgültig geklärt. Dietrich und Gerwien beschreiben verschiedene Ansätze, die für oder gegen eine gemeinsame Verarbeitung akustischer und optischer Sprache sprechen (2017, 219ff). Yuki Asano identifiziert deutliche Hinweise darauf, dass ein gemeinsames Sprachzentrum jegliche Sprachreize von außen verarbeitet:

Die Modelle weisen darauf hin, dass die Dekodierungsprozesse des Lesens und Hörens oder die Enkodierungsprozesse des Schreibens und Sprechens ähnlich verlaufen, mit anderen Worten, dass die Verarbeitung des akustisch oder visuell dargebotenen Inputs auf dem gleichen Mechanismus beruht.

(Asano 2011, 14)

Folgerichtig beschreibt Asano die Vorgänge akustischer oder visueller Natur auch in Einheit: »Die Dekodierung, bzw. Inputverarbeitung ist der Prozess, der mit dem Wahrnehmen von Lauten oder Schriftzeichen beginnt und mit ihrem Verstehen endet.« (Ebd.) Wenn der Sprachsinn also als eigener Sinneskanal angenommen werden kann, ergibt sich daraus eine Verschränkung von Informationen über mehrere Kanäle hinweg. Als Sekundärsinn nutzt der Sprachsinn den sensorischen Input anderer Kanäle, nämlich von Sehen und Hören, für die perzeptuelle Aufarbeitung.<sup>22</sup> Für die Analyse erscheint es für die meisten Fälle vorteilhaft, die akustischen und visuellen Informationen von den semantischen zu trennen. Allerdings sollte diese Trennung nicht zu streng vollzogen werden. Es ergibt beispielsweise wenig Sinn, Tonfall und eventuelle Spezialeffekte auf den Stimmen der Sprecher getrennt von den textlichen Inhalten zu behandeln. Insofern ist es zur Strukturierung der Analyse eventuell sogar von Vorteil, den Aspekt des Dialogs, hier zunächst unter dem akustischen Gesichtspunkt diskutiert, gänzlich in den Bereich des Sprachsins zu verschieben.

Die Darstellung von Sprache im direkten Zusammenhang mit dem Computerspiel bietet aber immer noch ein weites Feld, das für die Analyse erschlossen werden kann. In MGS können verschiedene Spracheinsätze eindeutig den Sphären zugeordnet werden. So ist der Funkkontakt mit Revolver Ocelot spielerisch zu begreifen; Ocelot gibt Tipps zur Spielsteuerung und zu Spielmechaniken sowie

22 Eine zusätzliche Kombination mit dem Tastsinn bei blinden Menschen und Braille-Schrift ist denkbar, für diese Arbeit aber nicht relevant, weil im Computerspiel zumindest bisher nicht realisiert.

zur weiteren Organisation von Missionen. Gespräche von Wachen untereinander haben tendenziell eher eine erzählerische Wirkung, besonders dann, wenn sich die Wachen in einer unbekannten Sprache unterhalten. So sprechen die Sowjet-Truppen in der ersten Mission von MGS miteinander logischerweise Russisch. Wer zum Großteil der Zielgruppe des Spiels gehört, der des Russischen nicht mächtig ist, kann den Unterhaltungen keinerlei Informationen entnehmen; damit dienen sie einzig der Etablierung einer Atmosphäre der Spielwelt, der Charakterisierung der Gegner als »fremd« und der Unterstützung eines realistischen Settings.

In HITMAN unterhalten sich die Figuren in den verschiedenen Levels ebenfalls miteinander und mit Agent 47. Diese Gespräche finden ausnahmslos auf Englisch statt, unabhängig von der Herkunft der sprechenden Figuren oder dem Handlungsort. Der überwiegende Teil dieser Gespräche ist dabei spielrelevant, es werden entscheidende Informationen über mögliche Lösungen der Rätsel preisgegeben oder es gibt Hinweise, welche anderen Herangehensweisen noch möglich wären. Die Dialoge sind fast immer missionszentriert, ein trivialer Austausch zwischen den virtuellen Personen findet nicht statt. Insofern sprechen die Figuren eigentlich immer mit den Spielenden, nie miteinander. Das ist natürlich wenig glaubhaft, denn die Erfahrung aus dem Lebensalltag lehrt, dass Menschen sich eigentlich ununterbrochen über Dinge austauschen, die für Unbeteiligte keine direkte Relevanz haben, auch wenn sie die Unterhaltung freiwillig oder unfreiwillig mithören. So bröckelt die Fassade einer in sich geschlossenen Spielwelt in HITMAN, wenngleich diese Fassade ohnehin nicht mit letzter Konsequenz errichtet wird. Hier spielt sicherlich wieder das Spielprinzip eine Rolle: HITMAN ist ein Rätselspiel, keine Simulation. Es geht nicht um eine glaubhafte und lebendig wirkende virtuelle Welt, zumindest nicht in erster Linie; es geht um die Lösung von komplexen Rätseln, die lediglich in den Farben der Simulation bemalt sind.

Ein schriftlicher Spracheinsatz findet in HITMAN abgesehen von Menüs sowie spezifischen Eingabeaufforderungen für besondere Aktionen im Spiel kaum statt. Generell wird in modernen Spielen vergleichsweise wenig Text eingesetzt. Vor den technischen Möglichkeiten der Audio-Implementierung war Text das einzige sprachliche Werkzeug für Spieleanthropologen und damit geradezu allgegenwärtig, jeder Dialog wurde in Form von Textkästen abgehandelt. In Dialogsystemen findet sich dann auch heute noch Text in vielen Spielen, nämlich immer dann, wenn eine Auswahl verschiedener Antwortmöglichkeiten besteht. Von THE ELDER SCROLLS V: SKYRIM über die MASS-EFFECT- (BioWare 2007) und DEUS-Ex-Serien (Eidos/Square Enix 2000) bis hin zu THE WITCHER III: THE WILD HUNT (CD Projekt RED 2015) und jüngst CYBERPUNK 2077 (CD Projekt RED 2020) können die Spielenden in Dialogen aus mehreren Antwortoptionen aussuchen. Der Text verkörpert dabei eine vorsprachliche Ebene, er zeigt quasi den Gedanken an, bevor eine Antwort gegeben wird. Erst mit Bestätigung der Auswahl wird die Antwort

realisiert – in SKYRIM implizit, in den anderen genannten Beispielen durch zusätzliche gesprochene Sequenzen der Hauptfigur.

Wiederum schwieriger in gängigen Kategorien der Wahrnehmung zu begreifen ist das Phänomen der Zeit, das aber trotzdem als Sinn eingeordnet werden kann. (Vgl. Stadler 2008, 7f) Die Zeitwahrnehmung des Menschen ist nicht stabil, sondern in vielfältiger Weise manipulierbar (vgl. Eagleman 2008), wohl auch deshalb, weil sie auf Material aus vielen verschiedenen Quellen zurückgreift: »Psychological time relies on sensory, perceptual, attentional, and memory processes at different time scales. [...] Psychological time is influenced by many different factors, such as attentional processes and contextual changes.« (Block und Gruber 2014, 132) Das Konzept der Zeitwahrnehmung beschäftigt Psychologen und Philosophen seit langer Zeit und die systematische Forschung spätestens seit den 1960er Jahren. (vgl. Allan 1979 und Warden 2016, 24f)<sup>23</sup> Eine Schwierigkeit bei der Untersuchung des menschlichen Zeitempfindens liegt in seinem nichtlinearen Verhalten, wie Hinrichs ausführt:

Kurzweilige Zeiträume, die mit belebenden Ereignissen gefüllt sind, erscheinen in der Erinnerung lang, während das gegenwärtig Langweilige in der späteren Rückschau sich verkürzt darstellt oder dem Vergessen schneller anheimfällt.

In der Lust scheint die Zeit schneller zu vergehen, bei Schmerz und Hunger erscheint der Zeitablauf verlangsamt.

[...]

Störungen des Zeitgefühls sind auch beim Gesunden nichts Seltenes, insbesondere dann, wenn die Vigilanz gegenüber der Umwelt verändert ist, also etwa bei Schläfrigkeit, Intoxikationen oder unter den Bedingungen sensorischer Deprivation. Daß Zustände noch tieferer Regression, Schlaf und Psychose, »jedes Zeitmaß fehlt«, hatte Freud bereits 1900 angenommen.

(Hinrichs 1984, 347f)<sup>24</sup>

Was Hinrichs als »Vigilanz gegenüber der Umwelt« bezeichnet, spielt auch in der aktuellen Forschung zur Zeitpsychologie unter dem Begriff der »Aufmerksamkeit« eine wichtige Rolle. (Vgl. Angrilli et al. 1997; Nordin et al. 2013, 1091) Dennoch legt Eagleman nahe, dass Zeitempfinden keine rein kognitive Erfahrung ist, sondern auch eine unwillkürliche Komponente besitzt. (2008, 131) In jedem Fall sind Zeit und ihr Vergehen basale menschliche Erfahrungen, die sich auch in der Erzähltheorie bereits umfangreich niedergeschlagen haben. Die Begriffe von Erzählzeit

23 Sanders und Cairns verfolgen den Beginn dieser Untersuchungen bis zum Ende des 19. Jahrhunderts zurück (vgl. 2010, 161), allerdings beziehen sie sich mehr auf die Untersuchung des Phänomens an sich und weniger auf empirische Experimentalforschung.

24 Hinrichs zitiert Freuds TRAUMDEUTUNG. Vgl. Freud 2010, 108.

und erzählter Zeit sowie zeitbezogene Konzepte wie Ordnung, Dauer und Frequenz (vgl. Martinez und Scheffel 2005, 30ff) sind narratologisch etabliert, deshalb aber natürlich nicht automatisch für das Computerspiel nutzbar. So argumentiert Juul für die Verwendung der Begriffe »Spielzeit« und »Ereigniszeit« (»game time« und »event time«; Juul 2006). Die erzähltheoretischen Begriffe von Erzählzeit und erzählter Zeit (Juul verwendet die Begriffe »story time« und »discourse time«) seien auf das Computerspiel nicht anwendbar, weil kein zeitlicher Abstand zwischen Handlung und Erzählakt angenommen werden könne. (Ebd.; vgl. auch Korbel 2009, 9) Entsprechend unterscheidet Juul zwischen der Zeit, die in der virtuellen Welt vergeht (Ereigniszeit) und jener, die die Spielenden beim Spielen aufwenden (Spielzeit). Weiterhin problematisiert er Konzepte wie die Prolepse, da die zukünftige Konfiguration der virtuellen Welt ja von den Handlungen der Spielenden abhängen müsse und entsprechend nicht vorweggenommen werden könne. Die Zeitverhältnisse der spielerischen Sphäre und jene der erzählerischen Sphäre unterscheiden sich also grundsätzlich.

Es ist vielleicht gerade dieser komplexe Zusammenhang, der Entwickler immer wieder zur Integration zeitflüider Mechanismen in ihre Titel inspiriert. So gilt die oft als »bullet time« bezeichnete Verlangsamung der Zeit (entsprechend der Zeitlupe beim Film) inzwischen als etabliertes gestalterisches Mittel. (Vgl. Korbel 2009, 49) Während die narratologische Entsprechung des zeitdehnenden Erzählens vor allem der Betonung einer Situation und der Vermittlung besonders vieler Details dient, ermöglicht die *bullet time* im Computerspiel ebenso spektakuläre wie ansonsten schlicht unmögliche Manöver. Denn während sich die Zeit im virtuellen Raum dehnt, bleibt der Zeitverlauf für die Spielenden logischerweise konstant; sie können sich also relativ zur virtuellen Welt sehr viel schneller bewegen. In MAX PAYNE (Remedy 2001), das Spiel, das die *bullet time* popularisiert hat, können die Spielenden durch die Verlangsamung der Ereigniszeit deutlich präziser zielen und beispielsweise mehrere Gegner ausschalten, bevor diese reagieren können. In MGS findet sich eine automatische Version von *bullet time* im sogenannten »Reflex Mode«, der ausgelöst wird, wenn Snake von einem feindlichen Soldaten entdeckt wird. Die Zeitdehnung gibt den Spielenden ausreichend Gelegenheit, auf den Entdecker zu reagieren und einen Alarm zu verhindern. Die Verlangsamung ist dabei insofern eine Nachahmung der menschlichen Erfahrungswelt, als das Zeitempfinden wie erwähnt aufmerksamkeitsabhängig ist. Berühmt sind die Erzählungen, dass sich ein traumatischer Moment zeitlich besonders ausgedehnt hätte; dieser Effekt wird heute vor allem der erhöhten Aufmerksamkeit und der nachträglichen Zuschreibung einer längeren Dauer zugesprochen. (Vgl. Terr 1983; Stetson et al. 2007; Yamada und Kawabe 2011; García-Pérez 2014) Weiterhin kann eine solche Spielmechanik eine Annäherung an eine möglicherweise überlegene Wahrnehmung der Spielfigur darstellen, wie sie bereits bei der visuellen Perspektive erläutert wurde; die Bezeichnung als »Reflex Mode« deutet diese Interpretation ebenfalls an.

Neben der Verlangsamung der Ereigniszeit hat sich auch ihre Umkehrung inzwischen als gängige Mechanik etabliert. Was in PRINCE OF PERSIA: THE SANDS OF TIME (Ubisoft 2003) noch eine zentrale, spielbestimmende Eigenschaft des Titels war, ist heute besonders bei Rennspielen beliebt. (Vgl. GRID 2, Codemasters 2013; FORZA: HORIZON 4, Playground/Turn 10 2018) Besondere Signifikanz hat die »Rückspul-Funktion« in BRAID (Number None/Hothead 2008), bei dem das Level-Design vollständig um die Zeitmanipulation herum aufgebaut ist. Der Kunstgriff hier ist das narrative Ende des Spiels, bei dem die Mechanik erzählerisch umgekehrt wird. So finden die Spielenden heraus, dass Spielfigur Tim nicht eine Prinzessin vor einem grausamen Ritter rettet, sondern selbst der Verfolger ist und die Prinzessin vor ihm flieht. (Vgl. Stamenković und Jaćević 2015, 194f)

Zeit ist auch ein wichtiger Faktor des Bewusstseins schlechthin, wie Sarrazin et al. (2008) feststellen, und spielt eine Rolle bei der Attribuierung von Handlungen zu einem selbst.<sup>25</sup> Der Frage, ob menschliches Handeln wirklich auf einem freien Willen basiert oder ob die Entscheidung zu einer Handlung erst retrospektiv hinzugefügt wird, ist dabei noch Gegenstand der Forschung.<sup>26</sup> Einig ist man sich aber darüber, dass es ein Bewusstsein für das eigene Handeln gibt, eine Zuweisung von Geschehnissen in der umgebenden Welt an bewusste Initiativen des eigenen Körpers. »One form of bodily self consciousness is a sense of agency, of oneself as the agent of action.« (Marcel 2003, 50) David et al. definieren diese Empfindung: »The sense of agency refers to the ability to recognize oneself as the author of one's own actions and distinguish these from actions caused or controlled by other sources.« (2011, 1; vgl. auch Beck et al. 2017, 54) Anhand dieses Sinns<sup>27</sup> erkennen Menschen ihre Wirkmacht in der Welt und bewerten ihre Interaktion mit den sie umgebenden Gegenständen. Vor diesem Hintergrund leuchtet es ein, dass gerade diese Sinneswahrnehmung für das Medium Computerspiel von ganz entscheidender Bedeutung ist.

Im Gegensatz zu klassischen, rein rezeptiven Medien funktioniert das Computerspiel ausschließlich unter der Voraussetzung der Partizipation, also der Herauslösung der Rezipierenden aus einer passiven Rolle und der Einnahme einer aktiven Mitwirkung am Geschehen. Das Computerspiel ist ein interaktives Medium; es verlangt die Interaktion der Rezipierenden, verlangt ihre Aktivität und damit eine

25 »Time is a fundamental dimension of consciousness. Many studies of the ›sense of agency‹ have investigated whether we attribute actions to ourselves based on a conscious experience of intention occurring prior to action, or based on a reconstruction after the action itself has occurred.« (Sarrazin et al. 2008, 602)

26 Vgl. Haggard et al. 2002.

27 In der Bezeichnung als »sense of agency«, was mit der hier verwendeten Bezeichnung von »Aktivitätssinn« korrespondiert, zeigt sich Einigkeit in der Forschung. (Vgl. Marcel 2003, David et al. 2011)

Wahrnehmung eigenen Handelns, eine Korrelation zwischen einer Handlungsabsicht und einem dieser Absicht entsprechenden Geschehen in der Welt, das wiederum von den Sinnen erfasst und rückgemeldet wird, um die Empfindung von *agency* zu erzielen.<sup>28</sup> Eine Handlung nach dem Aktivitätssinn besteht also aus drei Komponenten: der Handlungsabsicht, dem Geschehen und einer sensorischen Rückkopplung, aus deren Gesamtheit sich der Zusammenhang zwischen Absicht und Geschehen rekonstruieren lässt. Der Ablauf muss dabei drei Regeln folgen: Die Absicht muss der Handlung vorhergehen (Priorität), die Absicht muss zur Handlung passen (Konsistenz) und die Absicht muss der einzige Grund für die Handlung sein (Exklusivität). (Vgl. Wegner und Wheatley 1999, 483ff)

Der Aktivitätssinn findet sich von den Sphären wiederum sehr unterschiedlich angesprochen. Erzählerische *agency* ist ein komplexes Thema, für das es einige Modelle gibt, die in Kapitel 3 vorgestellt werden. Die spielerische Aktivitätsempfindung entsteht aus dem Zusammenwirken von Steuerung und sensorischem Feedback, also aus der Realisierung von Steuereingaben in der virtuellen Welt. Menschen sind gut darin, solche abstrakten Verbindungen zu internalisieren und als direkte Verbindungen zu betrachten – man vergleiche nur das Drehen eines Lenkrads und die resultierende Richtungsänderung eines Fahrzeugs. Entsprechend lassen sich spielübergreifend verbreitete Verbindungen wie das Hochdrücken eines Sticks und die Vorwärtsbewegung der Figur oder das Heben des Blicks schnell automatisieren, zumal eine gewisse formale Ähnlichkeit zwischen der Steuerbewegung und der resultierenden Figurenreaktion besteht. Andere Funktionen wie die Navigation durch Menüs mit den Tasten des D-Pads oder der Wechsel zwischen aufrechtem Gang, Hocke und Kriechen mit einer Taste erfordern zunächst eine bewusstere Steuerleistung, laufen mit zunehmender Spielzeit aber schließlich auch wie von selbst ab.

Aufgrund des hohen Abstraktionsgrads der Steuerung eines Computerspiels über ein Gamepad oder eine Tastatur werden Steuerschemata immer dann besonders interessant, wenn sie sich in ihrer Form mimetisch einer echten Bewegung angleichen. In GOD OF WAR 3 (SCE Santa Monica 2010) gibt es eine Szene, die speziell auf den Aufbau des Playstation-Controllers zugeschnitten wurde.<sup>29</sup> Das Sony Gamepad verfügt über zwei analoge Sticks, die zusätzlich zu ihrer Bewegung eine Druckfunktion haben, also eingedrückt werden können, um zusätzliche Funktionen zu erreichen. Im für seine explizite Brutalität bekannten GOD OF WAR 3 kämpft Spielfigur Kratos unter anderem gegen den Meeresgott Poseidon. Im Laufe des Kampfes wird die Eingabeaufforderung »L3 + R3« eingeblendet, was für das gleichzeitige Eindrücken der beiden Stick-Tasten steht. Die Bewegung der Spielenden,

28 Vgl. Haggard 2003, 111; Sarrazin et al. 2008, 603; Weiss et al. 2014, 82.

29 GOD OF WAR ist eine Spielserie von Sony, also von derselben Firma, die auch die Playstation herstellt. Die Spiele dieser Reihe sind daher nicht auf anderen Systemen verfügbar.

um der Eingabeaufforderung folgen und den Kampf gewinnen zu können, besteht also aus einem gleichzeitigen und symmetrischen Eindrücken der beiden Sticks mit den Daumen beider Hände. Es handelt sich um ein ungewöhnliches Kommando, dessen Sinn sich aber sogleich erschließt: Nach erfolgter Eingabe drückt Kratos seinem Widersacher beide Daumen in die Augen und entscheidet den Kampf so für sich.<sup>30</sup>

Es gibt generell Versuche, Steuerkonzepte für Control-Pads stärker mimetisch an den Bewegungen zu orientieren, die sie auslösen, oder ihnen zumindest eine möglichst intuitive Struktur zu geben. Die typische Taste zum Abfeuern einer Waffe in Shootern hat sich beispielsweise vollständig verlagert. War in den 1990ern noch eine der vier Oberflächentasten (zum Beispiel A, B, X und Y beim Super Nintendo; Kreuz, Kreis, Dreieck und Viereck bei der Sony Playstation) die typische Aktionstaste für einen Schuss, liegt diese Funktion heute üblicherweise auf einer Schultertaste, wodurch die für einen Schuss erforderliche Bewegung sich morphologisch einem Finger am Abzug annähert. Kampfkunst-Spiele wie TEKKEN (Namco 1995) oder MORTAL KOMBAT (Midway 1992) ordnen verschiedenen Tasten eine Extremität zu, sodass eine nachvollziehbare Systematik entsteht, die sich die Spielenden als Automatismus aneignen können. In ASSASSIN'S CREED (Ubisoft 2007) wird ein ähnlich physiologisches Modell genutzt: Von den rautenförmig angeordneten Tasten steuert die untere die Beine, die beiden Tasten links und rechts sind den jeweiligen Armen zugeordnet und die obere Taste kontrolliert den Kopf.

Die perzeptuelle Sinnesebene erschließt also fünf weitere Ebenen der Wahrnehmungsorganisation für die Strukturierung des Computerspiels: den Innensinn als möglichen Marker für das Phänomen der Immersion; den Sprachsinn als bedeutenden Kanal zur Kommunikation komplexer Konzepte; den Identitätssinn als elementares Erlebnis virtueller Präsenz; den Zeitsinn zur chronologischen Organisation des Erlebens; und schließlich den Aktivitätssinn als Rückkopplung der eigenen Handlungsfähigkeit. Im Zusammenspiel mit den zuvor bereits erläuterten physiologischen Sinnen ergibt sich ein Katalog aus zehn potentiell wirksamen Kanälen, die vom Computerspiel bedient und entsprechend in der Analyse genutzt werden können. Jeder Sinneskanal kann unterschiedlich wirken, je nachdem, ob er vorwiegend in der spielerischen oder der erzählerischen Sphäre aktiv ist.

---

30 Eine Interpretation dieser Szene kann nicht umhin, festzustellen, dass die mimetische Involverierung der Spielenden in diesen grausamen Akt zugleich durch die Perspektive aufgebrochen wird, denn die Kamera blickt in diesem Moment nicht durch Kratos', sondern durch Poseidons Augen. Die Spielenden sind damit in einem Moment zugleich Täter und Opfer, sie erleben einen Akt außergewöhnlich brutaler Gewalt und nehmen zugleich automatisch Abstand zur allein von ihrem Zorn regierten Figur Kratos ein.

