

FILMONTOLOGIE DER PHONOSPHERISCHEN PRÄSENZ

AUDITIVE DIFFERENZIERUNGEN ZWISCHEN
ZEICHENPOTENZIALEN UND LEIBAFFEKTION

Lars C. Grabbe

Zusammenfassung/Abstract

Etablierte medientheoretische Analysen und Theoriemodelle zum Thema Film legen zumeist einen ausgeprägten Fokus auf die visuelle Ebene von medialen Zeichen, um dann erst eine sekundäre Wechselwirkung mit anderen Zeichensystemen zu berücksichtigen. Das Bild des Films sollte allerdings nicht als Dominante der Erklärung fungieren, da sonst vor allem die komplexe und eigenständige Dimension des Tons nicht adäquat berücksichtigt werden kann. Denn Töne, Geräusche, Musik, Melodien und Soundscapes sind keine zweitrangigen Phänomene, sondern eigenständige artifizielle Präsenzen, deren autonome Analyse viel dazu beitragen kann, der Multimodalität von Mediensystemen vollständig Rechnung zu tragen.

Hinzu gesellt sich die problematisierende Frage, ob wir es bei Tönen und Tonsequenzen überhaupt im strengen Sinn mit einem semiotisch konsistent zu beschreibenden Zeichensystem zu tun haben, oder ob es gravierende Differenzen gibt zwischen Bild, Schriftbild, lautsprachlicher Äußerung und Tonsequenzen eines charakterisierten Objekts sowie komplexen musikalischen Arrangements. Diese Differenzen können vor allem im Kontext einer Multimodalität der Wahrnehmung deutlich nachgewiesen werden, wobei Tönen dann eine primär denotative Funktion (primäre Bedeutungskonvention) abgesprochen wird. Im Gegenzug zeigt sich dann vielmehr die konnotative Funktion (Überlagerung einer primären Bedeutungskonvention) von Tönen und Tonsequenzen – im Kontext einer körperlichen Leibaffektion – als ontologische Beschreibungsgröße einer sich im Film entfaltenden phonosphärischen Präsenz.

Established media theoretical analyses and theory models in the context of film are mostly focussing the visual level of medial signs, not till then considering a secondary interdependency with other sign systems. The image of film should not serve as a dominant concept of explanation, because in this case the complex and autonomous dimension of sound could not be considered adequately. Sound, noise, music, melodies and soundscapes are no second rate phenomena, they are independent artificial presences whose analysis could contribute a lot for the complex description of the multimodality of media systems.

Additionally we have to ask a problematic question, if sounds and sound sequences correlate in a stricter sense with a semiotic sign system that is consistently describable, or if there are serious differences between images, type face, acoustic language and sound sequences of a defined object or complex musical arrangements. These differences could be clearly verified in the context of the multimodality of perception, in which sounds no longer possesses a primary function of denotation (primary meaning). In return we could reveal the primary function of connotation (overlap of primary meaning) – in the context of a bodily affection – as an ontological descriptor of a film inherent phonosphere.

Theoretische Vorüberlegungen

Dennoch gehört die Erforschung der Natur der Repräsentation an den Anfang jeder Untersuchung darüber, wie Symbole innerhalb und außerhalb der Künste funktionieren. Daß Repräsentation in einigen Künsten, etwa der Malerei, häufig und in anderen, etwa der Musik, nicht häufig vorkommt, stellt für eine Einheitsästhetik ein Problem dar [...]. (Goodman 1997: 15)

Semiotik fragt nicht zuerst nach komplexen Bedeutungsstrukturen und semantischen Potenzialen von Zeichen im Alltag oder Zeichensystemen der Kunst. Vielmehr geht es anfänglich um das Verhältnis von Ausdruck und Inhalt, Signifikant und Signifikat¹, zueinander und der Beziehung von komplexen Signifikantensystemen untereinander. Legt man diesen Maßstab an Sprache, Schrift und Bilder, so ergeben sich neben denotativen Referenzmustern (primären Bedeutungskonventionen) auch konnotative Bezüge (Bedeutungsüberlagerungen), die allesamt eine ausgeprägte Codierung aufweisen.²

1 «Bewegen wir uns von der Systemebene des Zeichens in dessen linguistische Strukturebene, so lässt sich das Zeichen als Relation von Signifikant und Signifikat bezeichnen. Der Signifikant ist das Lautbild dem eine Bedeutung/Vorstellung zugeordnet ist (Ausdrucksebene). Das Signifikat ist eben diese Bedeutung/Vorstellung (Inhaltsebene)» (Grabbe & Rupert-Kruse 2009: 24).

2 «1. Ausdruckssubstanz: z. B. die lautliche, artikulatorische, nicht funktionale Substanz, mit der sich die Phonologie befasst; 2. Ausdrucksform, die aus den paradigmatischen und syntaktischen Regeln besteht [...]; 3. Inhaltssubstanz: dies sind z. B. die emotiven, ideologischen oder einfach begrifflichen Aspekte des Signifikats, sein positiver Sinn; 4. Inhaltsform: die formale Organisation der Signifikate untereinander, aufgrund der Anwesenheit oder Abwesenheit eines semantischen Merkmals; dieser letzte Begriff ist sehr schwer zu fassen, weil es bei der menschlichen Rede unmöglich ist, die Signifikate von den Signifikanten zu trennen [...]» (Barthes 1983: 34f.).

Hierin liegt vermutlich bereits der erste Unterschied dieser Zeichensysteme zu Einzeltönen, komplexen Tonsequenzen oder melodischen Leitmotiven. Obwohl für die Relation von Signifikant und Signifikat das Prinzip der Arbitrarität gilt, d. h. die Beziehungsrelation ist nicht motiviert (vgl. Grabbe & Rupert-Kruse 2009: 24), realisieren sich Schrift, Sprache und Bild jedoch mit einem hohen Maß der inhaltlichen Präzision und Deutlichkeit. Das mag für konkrete Geräusche³ in ihrer Gesamtsequenz auch noch anwendbar sein. Extrahiert man allerdings Einzeltöne bzw. Frequenzausschnitte oder richtet das Augenmerk auf temporal ausgedehnte Tonsequenzen oder musikalische Leitmotive, dann sind diese sehr reduziert codiert und semantisch zuerst relativ schwach bis gar nicht konventionalisiert. Mit der Frage, ob es sich bei tonalen Strukturen überhaupt noch um denotierende Zeichen⁴ im Kontext einer Sinnvermittlung oder primären Bedeutungskonvention handeln kann, geht gleichzeitig eine Verschiebung der Perspektive einher. Denn statt eine semiotische Bedeutungsordnung als Argumentationsgröße nutzen zu können, muss der Blick auf den phänomenologischen Leib des Rezipienten gerichtet werden: Ein Schallereignis ist bereits vor jeglicher Konventionalisierung und Bedeutung konkret vorhanden, denn es ereignet sich im Medium der eigenen Leiblichkeit – und damit innerhalb einer ästhetischen Signatur von Empfindungen. In dieser Perspektive verschiebt sich der Fokus eindeutig von denotativen Primärbedeutungen in Richtung eines spezifisch konnotativen Potenzials, welches sich dann selbst grundsätzlich im Bezugssystem des empfindenden Körpers

3 Das Meeresrauschen denotiert die Vorhandenheit des Meeres als Ursache bzw. Quelle des spezifischen Schallereignisses.

4 «Zur Definition eines Zeichens gehört seine Fähigkeit, aufgrund eines Codes einen Zusammenhang zwischen Signifikant und Signifikat herzustellen [...]» (Eco 1977: 170).

ereignet. Diese Annahme einer korporalen oder somatischen Decodierung⁵ macht die Konnotation dann als Resultat einer leiblichen Schallerfahrung interpretierbar, die sich dem taktilen Spüren offenbart. Das Hören wird dann als Sinn begreifbar, der nichts anderes ist, als das Organspektrum des taktilen Spürens, welches dann in einem zweiten Schritt konnotative Referenzbeziehungen initiieren kann (vgl. Fiala o.J.: 6).

Bild, Schrift, Laut und Ton: Ein semiotisches Problem der Medienbeschreibung

In audiovisuellen Medienformen gehören visuelle und auditive (akustische) Strukturen zu einem multimodalen Zeichengefüge. Dass Bilder in diesem Kontext zeichenbasierte Wahrnehmungseffekte ausbilden, erscheint plausibel, denn auch sie besitzen ikonische, indexikalische oder symbolische Referenzen zwischen Signifikant und Signifikat (orientiert an der Sprache). Gleichermaßen zu akzeptieren ist die Klassifizierung von visuellen Erscheinungen, die wir im weitesten Sinne als Bilder bezeichnen können, im Kontext von Denotation und Konnotation. Das Foto einer Taube zeigt somit eine visuelle Erscheinung, die in einer Ähnlichkeitsbeziehung (ikonische Relation) zum konkret fotografisch fixierten Vogel (den wir direkt als Taube klassifizieren können) steht. Das Foto stabilisiert mittels eines chemischen Prozesses oder einer digitalen Codierung den abwesenden Vogel in einer visuellen Anwesenheit unter medialer Hilfe eines technischen Apparats. Die ähnlichkeitsbasierte und fotografisch fixierte visuelle Erscheinung der Taube denotiert folglich den realen Vogel, der somit zur kausalen Ursache des fotografischen Bildmotivs wird. Das Foto verdeutlicht demnach sehr plausibel, dass hier ikonische und indexikalische Prinzipien zur Anwendung gebracht werden. In einem weiteren Schritt sind visuelle Erscheinungen allerdings durchaus anschlussfähig für

eine erweiterte Semiotisierung durch Prozesse der Interpretation oder psychologischen Evaluation des Gesehenen. Die Taube kann also semantisiert werden, indem man das Foto nicht nur als fotografische Fixierung eines real existierenden Vogels auffasst, sondern als Darstellung der abstrakten Idee oder Metapher des Friedens. Dann werden Ikonizität und Indexikalität um eine symbolische Verweisstruktur zwischen Signifikant und Signifikat ergänzt – die Denotation mündet hier bereits in die Konnotation, da die Bedeutungseinheit Taube bereits mit individualpsychologischen Deutungsparametern der Friedensidee durchsetzt wird.

Auch der schriftbildlich fixierte Begriff der Taube lässt sich gemäß des eben angewendeten semiotischen Schemas klassifizieren. Die einzelnen Buchstaben t, a, u, b und e formen das Wort Taube, welches sich gemäß einer lautsprachlichen Vorgabe dann auch sinnhaft und funktional aussprechen lässt. Der schriftbildliche Begriff sowie das lautsprachliche Wort Taube denotieren hierbei jeweils die spezifische Vogelart, die sich beispielsweise von Möwen, Kanarienvögeln oder dem Weißkopfadler unterscheidet. Dass dieses Unterscheidungsprinzip auf einer sinnhaften mentalen Repräsentation basiert, begründet sich durch die gelernte Konvention, die Begriff und Wort sinnhaft mit der passenden Vogelart assoziiert bzw. verbindet. Diese denotative Referenz ist hierbei natürlich schon durchaus symbolisch, da die Verknüpfung von Begriff, Wort und passender Vogelart durch eine angelernte Konvention zu Stande kommt. Diese denotative Primärkonvention kann natürlich durchaus um eine konnotative Konvention erweitert werden, denn auch Begriff und Wort Taube können als Metapher und Idee des Friedens fungieren⁶, die dann weiterhin auch individuelle Emotionen, Prägungen und Affekte hervorrufen können.

Besitzen die denotative und konnotative Dimension von Zeichen in piktorialer, schriftbildlicher und sprachlicher Perspektive eine übersichtliche und präzise zu erörternde Strukturtypologie, so wird dies im auditiv-akustischen Referenzbereich zum Problemfall einer semiotischen Zuordnungsvorschrift. Vor allem bei höherstufigen musikalischen Arrangements, Leitmotiven oder

5 Für den Zusammenhang von medialen und visuellen Einflussfaktoren auf körperliche Prozesse siehe folgende Ausführung: «So wirkt z. B. die Konstitution des POV-Shots in CLOVERFIELD als Aspekt des filmischen Codes unmittelbar durch somatische Decodierung auf affektiver Ebene: Der Rezipient wird erschüttert. Die somatische Reaktion resp. Empathie des Rezipienten ist dabei als Dekodierungseffekt des vermittelten Codes anzusehen. Damit wird die somatische Decodierung des filmischen Codes zum Modus des Filmverstehens» (Grabbe & Rupert-Kruse 2008a: 308).

6 Für Konnotationen als kulturelle Elemente die den Zeichenradius produktiv erweitern, sei verwiesen auf Sonesson (1989). Auch eine «Bedeutungskomponente [...], die die Grundbedeutung überlagert und sich einer vom Kontext abstrahierenden Bedeutung entzieht» (Hepp 1999: 31).

dynamischen Soundscapes wird die Zuordnung und Unterscheidung von Denotation und Konnotation erschwert, wobei Geräusche allerdings noch einer klarer zu fassenden Systematik obliegen. Gewiss dürfte man ab einer gewissen Gradstufe des kulturellen Wissenshorizonts eine denotative Unterscheidung zwischen einem akustischen Reizangebot vornehmen können. Somit sollte eine tonale Differenzierung des Gurrens einer Taube vom aufgeregten Zwitschern des Kanarienvogels oder dem lauten und klaren Schrei eines Weißkopfadlers vorgenommen werden können. Ein simples Gurren einer Taube wird aber nur schwer für den akustischen Impuls einer Friedensmetapher erhalten bzw. diese denotieren können, jedenfalls nicht, wenn weitere Kontextinformationen oder multimodale Medienverwendungen fehlen. Ist das Foto einer Taube eine ikonische Repräsentation die nicht weiter reduzierbar oder aufteilbar ist (wir wollen das Foto nicht zerschneiden), so lassen sich Begriff und Wort Taube sehr wohl in kleinere Buchstabensegmente aufteilen oder ganz auf diese reduzieren. Zerlegt man hingegen die temporal ausgedehnte auditiv-akustische Tonsequenz des Gurrens einer Taube in ihre punktuellen Frequenzen im Millisekundenbereich, in Tempo, Rhythmuscluster oder Melodievariationen, dann wird eine eindeutige Denotation erschwert bzw. vollständig unmöglich. Gleiches gilt natürlich für komplexe Leitmotive, Kompositionen, Melodien oder bei einem spezifischen soundscore innerhalb einer Filmszene: Als temporal ausgedehnte Tonsequenz wird eine rezeptive Wirkung entfaltet, die spätestens dann nicht mehr objektiv als denotierendes Zeichen beschrieben werden kann, wenn man diese in ihre Einzeltöne (punktuellen Schallereignisse) zerlegt. Segmentiert man das Wort Taube in Buchstaben, so bleiben t, a, u, b, und e als autonome Buchstaben zurück, die entweder neu arrangiert werden können oder sich schlicht in anderen Worten wiederfinden lassen. Eindeutig bleibt an dieser Stelle aber folgender Sachverhalt. Ein Buchstabe t vereint in sich nur Eigenschaften die eine typische t-Verwendung bestimmen – ein t denotiert nur t. Ein Buchstabe t findet sich demnach nicht im Wort Haus. Hingegen lässt sich das punktuelle Frequenzspektrum von 440 Hertz, also genaugenommen Kammerton a, ebenso in einer Tonsequenz einer Verfolgungsjagd ausmachen sowie in einem gregorianischen Choral oder einem populären Musikstück.

Mit dieser Überlegung ist ein Problem der Wahrnehmung verknüpft. Buchstabe und Wortlaut des Buchstaben t sind entweder lesbar oder hörbar, beide jedoch verweisen auf ein spezifisches t-Sein als denotative Eigenschaft. Ein Kammerton a lässt sich zwar aufschreiben, ebenfalls innerhalb einer professionellen Notation, jedoch bleibt dieser schriftliche Aspekt konkret weit hinter dem akustischen Wahrnehmungsereignis des Hörens von 440 Hertz zurück. Zudem ist es relativ unwahrscheinlich, dass ein ungeübter Hörer in der Lage ist, ein punktuelles Tonereignis konkret als Kammerton a zu klassifizieren. Geräusche des Alltags (Telefonklingeln, Automotoren, Türklingel etc.), Musik im Radio, Filmmusik oder soundscore sind doch zumeist temporale Sequenzereignisse, in denen spezifische Frequenzen, Rhythmen und Melodien enthalten sind. Es geht folglich weniger um die semiotische Analyse von Einzeltönen als vielmehr um komplexe und temporal ausgedehnte Schallereignisse und deren konnotative Metabezüge, Kontexte und Semantiken.⁷ Diese temporal ausgedehnten Tonräume sind als komplexe Wahrnehmungsereignisse zu bewerten, die neben Bedeutungen und Semantiken eine nicht zu unterschätzende körperliche Perspektive aufweisen. Dumpfe Basstöne, ob im Live-Konzert oder im Kinosaal, empfindet der Rezipient als körperdurchdringendes Phänomen in der Bauchgegend, pulsierende *drone sounds* und *megadrones*⁸ im Horrorfilm verstärken visuelle Impulse von Gewalt, schrille Frequenzen im Obertonbereich

7 Guerino Mazzola (1997) betont in seinem Schema der ontologischen Topographie der Musik explizit deren Komplexität und Bezogenheit auf drei differenzierte Ebenen der Repräsentation von Realität. In dieser Orientierung ist ein musikalisches Zeichenpotenzial stets bezogen auf eine mentale, psychische sowie physikalische Realität. Musik ist folglich eine mentale Entität und ein Modus der Empfindung individualpsychologischer Emotionen. Grundlage bildet die physikalische Realität: «Peripheral and inner ear, auditory nerve, its path through multiple relays stations of the brain stem, the neo- and archicortical centers for auditory processing and memory, such as Herschl's gyrus and the hippocampal formation» (Mazzola 1997: 3).

8 «The synthesized «megadrone» is a low, continuous rumbling (like distant thunder and the threat of dangerous weather, like an earthquake, like a constantly uneasy stomach). These bass sounds reverberate in your abdominal regions. Similar long and reverberant bass sounds are used by Bach for the earthquake in the *Saint Matthew Passion*, by Berlioz for storm at sea in *The Trojans*, by circus drummers to dramatise the most spectacularly dangerous high wire acts [...]» (Tagg 1999: 21).

werden direkt im Kopf gespürt, *body sounds* (Herzschlag, Schrittgeräusche etc.) triggern empathische Prozesse und gesteigerte tonale Asymmetrie (z. B. Zwölftonmusik, Balinesisches Gamelan etc.), je nach kulturellen Hörgewohnheiten, kann für rezep-tive Irritation sorgen, aufgrund eines Bruchs mit der musikalischen Normerwartung von Takt, Frequenz, Rhythmus, Klang und Melodie. Diese Übersicht ließe sich um einiges weiterführen, an dieser Stelle reicht diese Auswahl allerdings, um auf die wesentliche Überlegung hinzuweisen: Tonerfahrungen sind komplexe Sinnesmomente, die ebenso für Alltagserfahrungen konstitutiv erscheinen, wie auch für die Wahrnehmung von Kunst. Neben Bildräume der medialen Erfahrung treten folglich stets Tonräume der Erfahrung und es konstituieren sich multimodale Erfahrungsräume des Atmosphärischen.⁹ An dieser Stelle sei nochmals auf die gleichwertige Berücksichtigung von Bild und Ton hingewiesen und ein Plädoyer formuliert, den Atmosphärenbegriff in ein Mehrebenenmodell zu überführen. In dieser Perspektive integriert die bereits als Syntheseprinzip fungierende Atmosphäre einerseits ikonosphärische Elemente des Visuellen (Film, Bild, Kamera, Schrift, Gestalt/Form, Raum etc.) sowie andererseits phonosphärische Elemente des Auditiven und Akustischen (Geräusche, Einzeltöne, Musik/Leitmotiv, Filmmusik, Sprache, Soundscapes, *auditory icons*¹⁰ oder *earcons*¹¹ etc.). Die Atmosphäre bleibt natürlich ein Gesamtphänomen, während die Teilebenen der Ikonosphäre und Phonosphäre heuristische Modellannahmen darstellen, um Detailanalysen unter gleich berechtigten Vorzeichen vornehmen zu können, ohne eine einseitige Betonung des Visuellen oder Tonalen.

Körper und Ton: Zur Multimodalität der Wahrnehmung

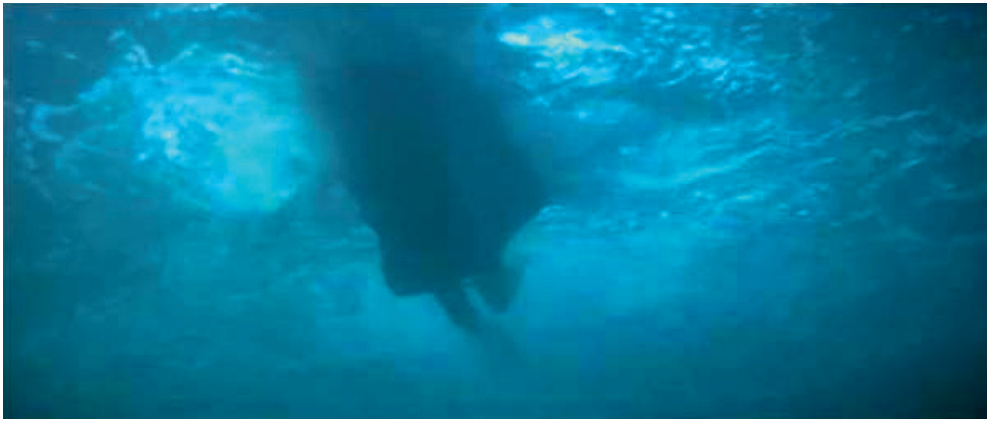
Verfügen ein Bild von einer Taube, der schriftbildlich fixierte Begriff oder der lautsprachliche Ausdruck des Wortes über ein klar definiertes Potenzial zur semiotischen Beschreibung, so wird diese Eindeutigkeit vom Tonereignis des Gurrens direkt problematisiert. Bild, Begriff und Ausdruck denotieren eindeutig das mentale Konzept «Taube», lassen zudem die Symbolisierung im Sinne einer Idee des Friedens zu und evozieren subjektive Konnotationen, welche individualpsychologische Charakteristika der Friedensidee zum Ausdruck bringen. Ein punktuelles Ton- bzw. Schallereignis, extrahiert aus einer vollständigen Tonsequenz des Gurrens, denotiert hingegen kein konkretes Konzept. Wir können feststellen, dass dieses punktuelle Schallereignis keinerlei semiotische Referenz ausweist, um den Zusammenhang zu einer real existierenden Taube (als Schallursache) herzustellen. Allerdings findet selbst im Millisekundenbereich bereits ein akustischer Effekt im Körper des Rezipienten statt, dazu wird Denotation nicht benötigt. Das Schallereignis wird somit zum leibphänomenologischen Potenzial des Affekts und reinen Spürens innerhalb der Rezeption. Es geht an dieser Stelle demnach vielmehr um eine methodische Neuorientierung, die gleichsam die Perspektive der Tonalität produktiv mit medientheoretischen Diskursen verbindet. An diesem Punkt der Argumentation wird nicht die Denotation oder verschiedenartige Referenzmodi zwischen Signifikant und Signifikat (Ähnlichkeit, kausaler Bezug, Konvention) thematisiert, sondern der phänomenologische Modus der Konnotation als Resultat eines Schallereignisses, empfunden im Medium der eigenen Leiblichkeit.

In dieser Perspektive lässt sich Konnotation phänomenologisch gewendet vielmehr als Form oder Ausprägung einer somatischen (körperlichen) Decodierung ausweisen, da unser Leib als Medium kaum Einfluss nehmen kann auf einen Ton, der sich als Schallereignis realisiert. Der Philosoph Erwin Fiala attestiert aus diesem Grund dem Ton ein immersives Moment, da er ein aktives Schallereignis ist, welches in den Gehörkanal des Rezipienten aktiv eindringt (vgl. o.J.: 5) bzw. sich als intensive Schwingung im Körper des Rezipienten ausbreitet. Ab einem bestimmten Lautstärkemoment (>65 dB) wird ein akustischer Reiz als Schmerz empfunden, ruhige Lieder wirken sanft bei durchschnitt-

9 Hans J. Wulff betont das Potenzial von Musik, dass sie bereits nach wenigen Takten spezifische Atmosphärenwerte erzeugen kann (vgl. 2012: 113).

10 Das bekannteste *auditory icon* ist das Geräusch von zerknülltem Papier während des Löschvorgangs eines Computer-Dokuments – das digital erzeugte Geräusch ist dem tatsächlichen Papiergeräusch ähnlich.

11 Hier steht die Konvention im Vordergrund, denn die Bedeutung von *earcons* muss erlernt werden, um exakt zugeordnet werden zu können, z. B. Klingeln zur Pause.



1 Das legendäre Zwei-Ton-Motiv in DER WEISSE HAI ist die tonale Repräsentation des Hais als Element der Bedrohung, noch bevor man ihn sehen kann. (Quelle: DER WEISSE HAI)

lich 72 bpm¹², dumpfe *megadrone sounds* evozieren typische Angstcharakteristika sowie negative Emotionen und plötzliche Knallereignisse generieren die affektive Schrecksekunde. Höherstufige Sequenzarrangements wie Refrains, Melodien, Riffs oder Hooklines können als autonome Toneinheiten wahrgenommen werden, die Elemente mnemonischer Repräsentation darstellen oder aktivieren (z. B. Melodien aus der Kindheit stimmen nach 30 Jahren noch fröhlich oder ein besonderes Lied erinnert an den Abschlussball, mit allen damit verbundenen Emotionen) oder das leitmotivische Thema aus DER WEISSE HAI (JAWS, Steven Spielberg, USA 1975) signalisiert im Kontext¹³ von filmischen Was-

serdarstellungen einen Antagonisten oder bringt abstrakte Bedrohungsvermutungen hervor. Diese letzten Beispiele unterstreichen dann die Annahme einer höherstufigen und weitestgehend kognitiven Konnotation, die entweder ein individualpsychologisches Moment darstellt oder aber über kulturelle Stereotypen¹⁴ intertextuell¹⁵ bei Rezipienten mental stabilisiert werden.

Ton- bzw. Schallereignisse sind zwar nicht sichtbar, es sei denn man transponiert sie in eine schriftliche Notation oder nutzt ein Instrument zur Messung von Frequenzen, sie sind allerdings körperlich konkret wahrnehmbar. Das ästhetische Moment wird dann verstehbar als ästhetische Wahrnehmungsidee, einer ganzheitlichen, körper-

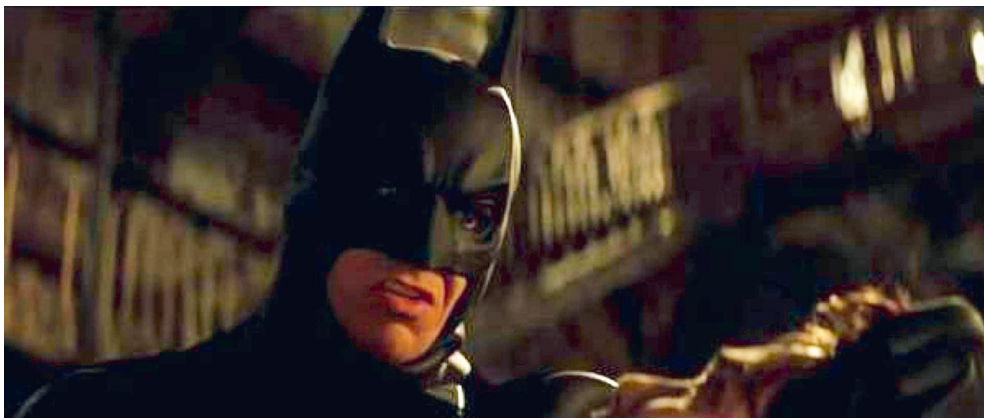
12 beats per minute.

13 Natürlich leben vor allen Dingen die spezifische Kinoerfahrung oder die häusliche Filmerfahrung von einer dualistischen Erfahrungskonstitution: Einerseits ist das Schallerleben ganzheitlich leiborientiert, und wird somit zu einer phänomenologischen Größe, andererseits wird ein auditiver Tonraum generiert, dessen Wahrnehmungsqualitäten repräsentational sowie semiotisch auf den Bildraum bezogen sind. Ton- und Schallereignisse werden demnach leibhaft sowie atmosphärisch gespürt, sie aktivieren jedoch ebenfalls einen Modus repräsentationaler Kopräsenz mit dem filmisch Dargebotenen und steigern eine immersive Partizipation. «Die wohl größte Bewusstheit des Atmosphärischen genießt der Film-Ton (hier *atmosphärischer Ton* oder kurz *Atmo* genannt). Wollen Filme einen realitätsnahen Eindruck erzielen, so müssen sie unsere spezifischen akustischen Umwelten imitieren. Die Atmo kondensiert oft stereotyp deren bestimmte und leicht dechiffrierbare Bestandteile. So setzt sich beispielsweise für den Handlungsort «Bahnhof» der Atmo-Ton aus dem Geräusch einfahrender Züge, von Schritten

auf dem Bahnsteig, Lautsprecherdurchsagen, Gepäckwagen und so weiter zusammen. Der amerikanische Terminus *sound image* deutet noch stärker darauf hin, dass es sich bei den akustischen Konstruktionen im Film um *Ton-Bilder* handelt, welche nach künstlerischen Gesetzmäßigkeiten organisiert sind» (Wulff 2012: 114).

14 Hier geht es dann um spezifische Lernkonventionen und kulturelle Wissensvariablen: «Ein dritter Typ beruht auf Gestaltphänomenen, denen er innerhalb der Medienkultur bereits so häufig begegnet ist, daß sich die damit verbundenen komplexen Motive durch vielfachen kommunikativen Gebrauch konventionalisiert haben und stereotyp geworden sind» (Wuss 1999: 11).

15 Nicht vergessen werden darf in diesem Kontext auch der Einsatz von Atmo-Ton im Film, der als akustische Konstruktion die diegetische Realität ausgestaltet und Ton-Umgebungen sowie *sound images* generiert (vgl. Wulff 2012: 114). Diese Elemente werden gelernt und können auf neue Sehgewohnheiten übertragen werden.



2 Obwohl Sprache und Bild über getrennte mediale Kanäle vermittelt werden, wird die Sprache dennoch kohärent auf die Bewegung der Lippen übertragen. (Quelle: BATMAN BEGINS)

lichen sowie sinnlichen Gewährwerdung.¹⁶ Es ist demnach eine adäquate Alltagsannahme, dass wir hinter dem Bellen eines Hundes in einem nicht einsehbaren Nachbargarten, einen real existierenden und anwesenden Hund faktisch annehmen. Das Schallereignis ist konkret an die mentale Repräsentation¹⁷ eines Hundes gekoppelt: Es gibt kein Tonereignis ohne gleichzeitige Annahme der faktischen Realität einer dazugehörigen Quelle. Das soll natürlich nicht heißen, man unterliege der fehlerhaften Annahme, dass man beim Hören von *Stairway To Heaven* gleichzeitig die faktische Anwesenheit der Band *Led Zeppelin* erwartet. Vielmehr nehmen wir ein Medium als Quelle der Musik an. Ähnlich verhält es sich bei der Wahrnehmung von Sprache innerhalb eines Films oder der Wahrnehmung von Filmmusik. Spricht Christian Bale in seiner Rolle in *BATMAN BEGINS* (Christopher Nolan, USA 2005) auf der Kinoleinwand, so sind die Quelle des Schallereignisses allerdings nicht die sich konkret bewegendenden Lippen des Schauspielers, sondern das digitale Lautsprechersystem innerhalb des Kinosaals. In der Wahrnehmung nimmt der Rezipient dieses Ereignis allerdings als kohärentes Ereignis wahr: Die Stimme gehört zu den Lippen des Sprechers (vgl. Anderson 1998: 84).

Auch die Filmmusik gehört zu einem konkreten Ereignis, welches immer auf die visuelle Fläche

der Darstellung bezogen bleibt. Das gehört zum apparativen Wissen des Rezipienten, dass Bild und Ton in einem Film zusammengehören und in dieser Zusammengehörigkeit den Film konstituieren.¹⁸ Auch bei extradiegetischer Musik werden Bild und Ton nicht separat aufgefasst, sondern als sich bedingendes Einheitsprinzip einer gegenseitigen Bezogenheit von Bild und Ton wahrgenommen.

Dieses Resonanzverhältnis (vgl. Murch 1994: XXII) von Bild und Ton lässt sich wahrnehmungstheoretisch fassen. Der Philosoph Bence Nanay argumentiert in diesem Kontext explizit für die Annahme einer «multimodality of perception» (2012: 1), die vor allem im ästhetischen Kontext empirische Begründungszusammenhänge zulässt und diese produktiv nutzt. Die Grundannahme dieser These formuliert die Möglichkeit, dass differente Modalitäten der Repräsentation (visuell, akustisch, olfaktorisch etc.) bereits im Wahrnehmungsakt selbst einheitlich aufgefasst werden. Es muss keine nachträgliche mentale Zerlegung

18 «Besitzen Objekte einen Ton, heben sie sich aus der diegetischen und narrativen Reststille ab und machen auf potentielle Bedeutung aufmerksam – tonale Struktur kann Bedeutungsstrukturen fokussieren und Initial der Hypothesenbildung sein. Werden also Objekten Tonstrukturen unterlegt (Objekt-Ton-Kopplung), die beim Rezipienten direkt decodiert und zugewiesen werden können, wie z. B. das Fauchen eines Löwen im Kontext eines Kühlschranks, so wird objektbezogene Handlungsintentionalität offeriert und als solche begriffen, ohne dass physikalische Bewegungselemente dazu kommen müssen – die tonale Struktur fungiert in diesem Fall als «Handlungsphänomen» (Grabbe & Rupert-Kruse 2008b: 296).

16 Hierin verkörpert sich eine spezifisch ästhetische Signatur des subjektiven Denkens (vgl. Welsch 2003: 46f.).

17 Zum Verhältnis von internen und externen Repräsentationen sei verwiesen auf Malafouris (2007).

oder Zusammenführung von repräsentationalen Inhalten erfolgen. Es ist demnach möglich, dass wir einen akustischen Sprechakt im Kinosaal auf den aktiven und visuell dargestellten Sprechakt einer Figur übertragen, oder eine asynchrone und fehlerhafte Synchronisation doch als synchron und fehlerfrei auffassen. Deutlicher wird dieses Phänomen anhand des empirisch erforschten Bauchredner-Effekts. Hier wird die Puppe als sprachaktive Instanz wahrgenommen, obwohl der akustische Sprechakt direkt vom Bauchredner stammt. Der akustische Impuls passt nicht zur Puppe, wird aber im Wahrnehmungsakt auf diese übertragen und in dieser Synchronizität kohärent¹⁹ aufgefasst. In dieser Perspektive ist der visuelle Sinn dominanter als der auditive Sinn. Diese kreuzmodale Illusion ist im Wahrnehmungsakt stabil realisiert und lässt somit eine zweite Grundannahme zu. Repräsentationen werden nicht grundsätzlich als mentale Konstrukte verstanden, da sie sich bereits im aktiven Wahrnehmungsakt realisieren können und dort bereits kreuzmodale Repräsentationen in actu ermöglichen. Diese Repräsentationsstruktur kann auch dominant durch den auditiven Stimulus realisiert werden, wie z. B. bei der Filmmusik²⁰, oder durch experimentell nachgewiesene Faktoren:

But there are more surprising examples of multimodal perception: if there is a flash in your visual scene and you hear two beeps while the flash lasts, you experience it as two flashes. This is one of the not so many examples where vision does not trump audition: the two beeps in our auditory sense modality influence the processing of the one flash in our visual sense modality and, as a result, our visual experience is as of two flashes. (Nanay 2012: 2)

Die Strukturen unserer Wahrnehmung ermöglichen demnach wahrnehmungsbasierte Repräsentatio-

nen²¹, bei denen sich je unterschiedliche Grade und Dominanzinduktionen von Visualität und Audition wechselwirksam ereignen. Neben die Multimodalität von medialen Strukturen und Mediensystemen tritt die Multimodalität der Wahrnehmung als anthropologischer und biologischer Faktor.

Ton/Bild-Relation und phonosphärische Elementaria

Verstehen wir Film zunächst in seiner technischen Bestimmung als ein audiovisuelles Repräsentationssystem welches multimodale Ton/Bilder hervorbringt (vgl. Grabbe & Rupert-Kruse 2013b: 18), dann ist dieser Grundannahme eine intermediale Wesensbestimmung der verschiedenen Mediensysteme – vornehmlich von Bild und Ton – inhärent. Denn Bild und Ton sind in der technisch apparativen Vermittlung zwar getrennte Modi der kommunikativen Artikulation, diese Trennung wird allerdings durch eine mediale Transformation (vgl. Merleau-Ponty 2010: 79) oder mentale Fusion (vgl. Murch 1994: XIX) in ein interaktionistisches Relationsgefüge (vgl. Grabbe & Rupert-Kruse 2013b: 19) überführt. Dieser Interaktionismus zeigt sich dabei nicht als vereinfachtes Nebeneinandertreten zweier differenter Modi, sondern als ganzheitlich mediale Modifikation:

[Der] Klang fügt dem Bild nicht nur etwas hinzu, der Klang verwandelt das Bildgeschehen, das seinerseits das modifiziert, was akustisch vernehmbar ist. Gesprochene Sprache, Musik und andere Geräusche bilden eine integrale Dimension des filmischen Bewegungs-bildes. (Keppler 2010: 447)

Die interaktionistische Bestimmung von Bild und Ton erzeugt während der audiovisuellen Repräsentation einen kausalen Gesamteindruck und damit

19 Interessante Ergebnisse zur Wahrnehmung widersprüchlicher Impulse bei gleichzeitigem kognitivem Ausgleich liefert der McGurk-Effekt. Zahlreiche Beispiele für die Versuchsanordnung lassen sich auf Youtube finden.

20 «Die Wirksamkeit isolierter empathischer cues auf der Ebene des Tons wurde u. a. durch Experimente von Annabel Cohen bewiesen. Cohen unterlegte sich bewegende geometrische Figuren mit unterschiedlichen Musiken und ließ sich dann von den Probanden ihres Experiments die Szenen beschreiben. Die Ergebnisse zeigten, dass durch unterschiedliche Musiken nicht nur die Figuren anders wahrgenommen wurden, sondern auch – je nach Bedeutung der Musik – die verschiedenen Handlungen unterschiedlich interpretiert wurden» (Grabbe & Rupert-Kruse 2008b: 296). Zusätzlich sei verwiesen auf Cohen und Marshall (vgl. 1988: 95f.).

21 Ein vergleichbar evidentes Theoriemodell zur spezifischen Wahrnehmungsnahe von Zeichenstrukturen legt Klaus Sachs-Hombach vor: «Die Interpretation des Zeichens wird zurückgeführt auf Wahrnehmungskompetenzen, die sich direkt auf den Zeichenträger und dessen präsentationalen Gehalt beziehen. Die unterschiedlichen Prozesse der Gegenstandswahrnehmung können, wie Sachs-Hombach argumentiert, demnach einerseits basale Bedeutungsstrukturen manifestieren (z. B. das Klingelzeichen als Index für das Ende der Schulstunde), andererseits sind, aufgrund speziellerer Wahrnehmungsmodi, ebenfalls spezifischere Bedeutungskomplexionen realisierbar (z. B. die *Seeing-In-Theory* nach Richard Wollheim, die das Erkennen von Tiergestalten in Wolkenformationen thematisiert)» (Grabbe 2011: 121; dazu auch Sachs-Hombach 2001).

ein neues synkretistisches Objekt innerhalb der Wahrnehmung des Rezipienten – weder nur Bild oder nur Ton, sondern ein Resonanzmoment²² aus beiden Faktoren.

Diese interdependente Fusionsordnung von Bild und Ton lässt sich in einem trichotomen Ordnungsschema zum Ausdruck bringen, welches innerhalb von Filmtheorie und Produktionsästhetik zur Anwendung kommt (Schneider 1997; Raffaseder 2010). Gemäß eines horizontalen Ähnlichkeitsspektrums (links, zentral, rechts) wird die Relation aus Bild und Ton als Paraphrase, Polarisation oder Dissonanz/Kontrapunkt (vgl. Schneider 1997: 24) gekennzeichnet. Die Paraphrase wird auf der linken Seite für die maximale Ähnlichkeit von Bild und Ton eingeordnet, während auf der rechten Seite die Dissonanz als maximales Differenzmoment verortet wird. Zentral zwischen diesen Ausprägungen liegt die Polarisation:

In medientheoretischer Perspektive erscheint es als äußerst plausibel und erforderlich, auf der Basis eines Ähnlichkeit-Differenz-Schemas, verschiedene Ebenen der Bild-Ton-Interaktion zu erarbeiten. Die jeweiligen Facetten dieses Schemas sollten jedoch nicht ausschließlich einen primär piktorialen Fokus aufweisen, sondern den Ton als eigenständige Sinn- und Handlungsdimension integrieren. In dieser semiotischen Perspektive wird das Verhältnis von Bild und Ton grundlegend als ein Resonanzphänomen beschreibbar, bei dem die tonale Dimension die Wahrnehmung des Visuellen – im Sinne einer akustischen Deutungsschablone – vorweg nehmen kann. Diese höherstufige und semiotische Perspektive scheint sich allerdings wesentlich stärker auf eine rezeptive bzw. funktionale Ebene zu beziehen, auf der sich die höherstufigen Modi der Bild-Ton-Interaktion artikulieren. Notwendige Grundlage für höherstufige Rezeptionen bildet eine technische Ebene, als Bedingung der Möglichkeit zur Präsentation variabler Sinnesinhalte.

(Grabbe & Rupert-Kruse 2013b: 34)

Paraphrase, Polarisation und Dissonanz sind prozessorientierte Elemente, die ihrerseits Konvergenzstadien von Bild und Ton zum Ausdruck bringen: Ähnlichkeit, Zentralität und Differenz. Paraphrasen basieren prinzipiell auf einer Ähnlichkeitsstruktur,

womit eine eher «identitätslogische Übereinstimmung von Bildinhalt und Ton gemeint ist (z. B. Klingelndes Telefon innerhalb der Diegeese; Meeresrauschen und Einstellung auf das Meer)» (Grabbe & Rupert-Kruse 2013b: 35). Polarisation hingegen drückt eine Verstärkung eines Bildinhalts aus bzw. generiert diese, indem der Ton zusätzliche Bedeutungen hinzufügt:

[Der] Bildinhalt wird akustisch überformt. Bild und Ton sind bei der Polarisation zentrisch verklammert, um ein filmisches Artefakt um eine symbolische Bedeutung bzw. Konvention zu erweitern und beispielsweise Emotionen auf Seiten des Rezipienten zu evozieren (z. B. Klaviersonate innerhalb einer Liebeszene).

(Grabbe & Rupert-Kruse 2013b: 35)

Die Dissonanz folgt schließlich einem Gegensatzprinzip der Störung und Irritation. Sie stiftet eine explizite Differenz von Bild und Ton (z. B. Dissonante Tonmodulationen in Horror-Filmen), bei der eine Resonanz von Bild und Ton von interpretativen Leistungen des Rezipienten abhängt (vgl. Pauli 2011: 12).

In ihrer grundsätzlichen Dimension der Wechselwirkung von Bild und Ton, lässt sich das Ähnlichkeit-Differenz-Schema in ein temporales Zeitmodell überführen, welches die Typologie der sequentiellen Zeitlichkeit²³ integriert. Konkret formt sich demnach bereits im bewussten Jetztthorizont der Rezeption eine Resonanz von Bild und Ton, die das Aktualmoment bestimmt: «Hier kann sich eine akustische Dimensionierung des Visuellen, ein tonales Interpretament visueller Informationen, vollziehen» (Grabbe & Rupert-Kruse 2013b: 35). Eine notwendige Erweiterung dieser Annahme findet sich in dem Kontext von Retention²⁴ und Rekognition, d.h. bereits im Vorfeld gemachte Tonerfahrungen²⁵ im Umgang mit Schallereignissen bilden eine Art Tongedächtnis aus:

23 Ohne Zeit, kein Ton. Diese Aussage ist ein notwendiger Begründungszusammenhang, um Temporalität und Rezeption zusammenzubringen: «Dieser funktionale Horizont wird zusätzlich zum Zeithorizont, in der Annahme einer temporalen Tonalität des filmischen Repräsentationssystems» (Grabbe & Rupert-Kruse 2013b: 35).

24 Retention ist «für die Konstitution der Dauer von Daten entscheidend» (Lohmar 2008: 9), als Funktion einer erinnernden Vergegenständlichung und Aktualisierungsmoment. Die Protention hingegen manifestiert die «Erwartung des Kommenden» (ebd.: 95).

25 «Wir haben es daher bei der Filmerfahrung mit einem hochgradig von Seiten des filmischen Stimulusmaterials kon-

22 Im Vorwort von Michel Chions *Audio-Vision: Sound on Screen* findet das Konzept «conceptual resonance» (Murch1994: XXII) Verwendung: «The sound makes us see the image differently, and then this new image makes us hear the sound differently [...]» (ebd.).



3 Im Kinosaal kommt der Eigentümer des Sternenzerstörers aus KRIEG DER STERNE von hinten rechts noch bevor man ihn sieht, dann schiebt der Ton das Bild langsam von oben in den sichtbaren Bereich (Quelle: STAR WARS, George Lucas, USA 1977)

Auf dieser Prozessstufe ist Rezeption stets bezogen auf ein inter- und intratextuelles Tongedächtnis. Inter-textuelle Tonalität fungiert als akustisches Stereotypenphänomen, durch das zurückliegende akustische Erfahrungen auf den Filminhalt (Dramaturgie, Requisite, Emotionalität etc.) bezogen werden können (z. B. der spezifische Klang amerikanischer Polizeisirenen ermöglicht die raumzeitliche Lokalisierung des Bildinhalts). Auch können intratextuelle Tonmotive durch Repetition eine leitmotivische Funktion übernehmen, durch die Aufmerksamkeit fokussiert, Spannung evoziert oder Emotionen gesteigert werden können. Bei spezifischer Rezeptionswirkung können intratextuelle Motive dann selbst zu intertextuellen Leitmotiven werden, die mit Situationen, Genres oder Heldentypen gekoppelt sind (z. B. Leitmotiv in INDIANA JONES; der imperiale Marsch in STAR WARS). (Grabbe & Rupert-Kruse 2013b: 35f.)

Natürlich bildet die rezeptive Erwartungshaltung ein zentrales Schlüsselmoment für das Erleben von Spannung oder empathische Prozesse mit filmischen Figuren. Dieses Element der Zukünftigkeit lässt sich ebenfalls in die temporale Dimension integrieren, wobei Tonstrukturen dann entweder visuelle Informationen inszenatorisch einleiten bzw. in das filmische Bild der Anwesenheit führen, oder aber vollständig vorbereiten.²⁶

trollierten Akts der Rezeption zu tun, der sich grundlegend über das zeitlich-narrative Moment des Films, zwischen den Polen Retention, Perception und Potention, strukturiert» (Grabbe & Rupert-Kruse 2013a: 27).

26 «Ebenso wie sich nicht sichtbare Bilder im hors-champ, im Umfeld des Bildes – oder besser: im Außerhalb des Bildrahmens – befinden, die durch die Kamerabewegung jederzeit

Dies lässt sich als akustische Hypothese kennzeichnen: «Mit dieser tonalen Hypothese gelingt dann ein akustischer Fokus auf eine visuell und zeitlich absente oder noch nicht realisierte Dimension (z. B. Das Leitmotiv in Jaws kündigt den Hai an, noch bevor man ihn sieht)» (Grabbe & Rupert-Kruse 2013b: 36).

Die Kategorisierung der phonosphärischen Präsenz verläuft innerhalb des erläuterten Ähnlichkeits-Differenz-Schemas in explizit temporaler Orientierung. Dabei korreliert diese Perspektive mit der film- und kognitionspsychologischen Annahme von intersensorischen Beeinflussungsstrukturen, die im aktiven Rezeptionsprozess kognitiv zusammengeführt werden. Stichwort für diese Theorie ist das Konstrukt der Synästhesie, die ebenfalls eine somatisch-leibliche Dimension der Wahrnehmungsstrukturierung integriert:

Erforscht wurden Synästhesiephänomene vornehmlich unter drei Aspekten. Erstens können Schwellenveränderungen für eine Sinnesmodalität durch Reize einer anderen herbeigeführt werden; laute akustische Reize erhöhen etwa die Sehschwelle. Zweitens hat man sich mit wechselseitiger Beeinflussung heteromodaler Wahrnehmungsqualitäten befaßt; starke und hohe Töne hellen z. B. Farben auf, tiefe Töne verdunkeln sie. Drittens lassen sich experimentell erzeugte Kopplun-

in die visuelle Kadrierung des Films gerückt werden können, können sich auch nicht sichtbare Bilder im hors-son befinden, die, vom Ton hervorgebracht, auf der Leinwand sichtbar werden können. Das hors-son ist somit im Kontext eines generativen Prinzips als Äußerlichkeit des Akustischen zu sehen» (Martin 2010: 147).

gen von heteromodalen Qualitäten beobachten, etwa im Rahmen bedingter Reflexe; so konnten z. B. Temperaturempfindungen durch in vorangegangenen Experimenten erzeugte akustische Reize beeinflusst werden.

(Wuss 1999: 296)

Peter Wuss bestätigt mit dieser Zusammenfassung der empirischen Forschungsergebnisse indirekt die bereits weiter oben formulierte Annahme einer Multimodalität der Wahrnehmung (Nanay 2012; Sachs-Hombach 2001) sowie der wahrnehmungsbasierten Repräsentation. Film scheint ein besonders geeignetes Medium zu sein, in welchem das Synästhesie-Konzept verstärkt als Form einer somatischen Decodierungsleistung zum Tragen kommt, da Film prinzipiell auf eine ausgeprägte Sinnvermutung auf Seiten des Rezipienten angewiesen ist. Der Rezipient ist im filmischen Kontext besonders sensibilisiert, um semantische Horizonte auszuloten und Handlungsparameter zu erfassen. Der Rezipient befindet sich in einem evaluierenden Orientierungsprozess, «bei dem alle Sinnesorgane aktiviert werden, Informationen aus dem Geschehen zu extrahieren. So scheint der Synästhesie-Effekt der Filmrezeption vor allem ein semantisch gesteuerter zu sein» (Wuss 1999: 297).

Fazit

Semantische und konnotative Effekte als Kategorie des Leibes zu begreifen erscheint vor einem auditiv-akustischen Hintergrund als plausibel. Dabei knüpft dieser Diskurs zwar an phänomenologischen Überlegungen an die u. a. von Vivian Sobchack bereits in den 1990er Jahren postuliert wurden, allerdings exemplifiziert die tonale Dimension eine bewusste und notwendige Akzentverschiebung. Der Ton und die komplexe Tonsequenz sind körperbezogene Phänomene, die Rezeptionsprozesse mit dem Medium Film in temporaler Perspektive synchronisieren. Es sind nicht nur Kamera und das visuelle Display die man als Argument für eine körperliche Aktivierung des Rezipienten anführen kann. Vielmehr dringt der Film selbst immersiv als Schallereignis in den Rezipienten ein (Fiala o.J.) und realisiert ein vollständiges, filmisches Interface²⁷ (Bewegtbild-Interface): Als somatische Kom-

ponente der leiblichen Affektion und leiblichen Konnotation.

Hier liegt ein evidenter Ansatzpunkt begründet, da sich biologische mit medientheoretischen Diskursen produktiv verbinden lassen, wenn der Bereich des rein Visuellen um die tonale und temporale Komponente erweitert wird. Zudem lässt sich eine theoretische Unschärfe ausgleichen, die allein damit zusammenfällt, die multimodale Komplexität des Films zu Gunsten einer monomodalen Perspektive zu vernachlässigen:

Positiv festhalten lässt sich, dass eine substanzielle philosophische Auseinandersetzung mit bewegten Bildern [...] den Bereich des rein Optischen weit überschreiten muss. Dabei geht es weniger darum, dass zu dem Gesehenen parallele taktile Qualitäten – wohl möglich imaginativ – «hinzugefühl» werden können. Entscheidend ist vielmehr, erstens die in der Wahrnehmung tatsächlich miteinander und mit der Motorik interagierenden Sinne zu berücksichtigen und zweitens zu verfolgen, wie das Mosaik der vom bewegten Bild dargebotenen Modalitäten (auch in zeitlicher Hinsicht) zusammengesetzt ist. Da es sich bei diesen Bildern um bewegte handelt und sie daher auf vielfältige Weise und besonders körperlich wirken können, ist schließlich und drittens nicht zu vergessen, dass «motion pictures» auch «e-motion pictures» (Carroll 2008: 147) sind.

(Liebsch 2013: 120f.)

In der Perspektive der Zusammenführung von Bild, Ton und Zeit lässt sich die «bodily activated subjectivity» (Jeong 2013: 79) im Sinne Vivian Sobchacks adäquat im Kontext von Multimodalität der Wahrnehmung und Multimodalität des Films (re-) konfigurieren. Der Film ist damit nicht mehr nur Prothese, sondern wird zu einem Initialmoment aktiver Rezeptionsprozesse – «a liberational potential of our body» (Jeong 2013: 81).

interface) (vgl. Jeong 2013: 6). Weiterhin wird der Interface-Begriff einer Neukonzeption unterworfen, wobei die zunächst technische Dimension um eine explizit körperlich-biologische Dimension ergänzt wird. Die Sinne des Menschen werden selbst zu Grundbedingungen einer Schnittstelle zwischen Subjekt und Objekt. Jeong bezeichnet sie als Quasi-Interfaces. Folglich fungiert «[...] the cinematic interface as a comprehensive combination of medium-specific interfaces and quasi-interfaces» (Jeong 2013: 8).

27 Seung-hoon Jeong entwickelt eine komplexe Theorie des cinematic interface als Schnittstellen-Phänomen zwischen Rezipient und Film. Er entwirft eine trichotome Verhältnisbestimmung zwischen Kino (kulturelles Interface), Körper des Subjekts (embodied interface) und Leinwand (perceptual

Literatur

- Anderson, Joseph D. (1998): *The Reality of Illusion. An Ecological Approach to Cognitive Film Theory*. Carbondale & Edwardsville: Southern Illinois University Press.
- Barthes, Roland (1983): *Elemente der Semiotologie*. Frankfurt a.M.: Suhrkamp.
- Carroll, Noël (2008): *The Philosophy of Motion Pictures*. Malden, MA & Oxford: Blackwell.
- Cohen, Annabel. J. & Marshall, Sandra K. (1988): Effects of Musical Soundtracks on Attitude toward Animated Geometric Figures. In: *Music Perception*, Vol. 6, No. 1. S. 95–112.
- Eco, Umberto (1977): *Zeichen. Einführung in einen Begriff und seine Geschichte*. Frankfurt a.M.: Suhrkamp.
- Fiala, Erwin (o.J.): *Versuch einer Semiotik des Phänomens »Musik«*, Online: http://www.gewi.kfunigraz.ac.at/grelle.musik/musicdigCulture/files/archiv/Fiala_Erwin/ErwinFiala_Versuch%20einer%20Semiotik%20des%20Ph%C4nomens.doc [26.06.2014].
- Goodman, Nelson (1997): *Sprachen der Kunst. Entwurf einer Symboltheorie*. Frankfurt a.M.: Suhrkamp.
- Grabbe, Lars C. (2011): *Georg Simmels Objektwelt. Verstehensmodelle zwischen Geschichtsphilosophie und Ästhetik*. Stuttgart: ibidem.
- Grabbe, Lars C. & Rupert-Kruse, Patrick (2008a): Gewalt als Erlebnis. Die somatische Decodierung als Strategie des Filmlebens in CLOVERFIELD. In: *Kodikas/Code. Ars Semeiotica. An International Journal of Semiotics*, Volume 31, No. 3–4. Tübingen: Gunter Narr. S. 299–312.
- (2008b): Empathische cues und die Möglichkeit einer Affekt-Sprache des Films. In: *Kodikas/Code. Ars Semeiotica. An International Journal of Semiotics*, Volume 31, No. 3–4. Tübingen: Gunter Narr. S. 289–298.
- (2009): Roland Barthes: Zeichen, Kommunikation und Mythos. In: *Schlüsselwerke der Cultural Studies*. Hg. von Andreas Hepp/Friedrich Krotz/Tanja Thomas. Wiesbaden: VS/GWV. S. 21–31.
- (2013a): Filmische Perspektiven holonisch-mnemischer Repräsentation. Versuch einer allgemeinen Bildtheorie des Films. In: *IMAGE – Zeitschrift für interdisziplinäre Bildwissenschaft*, Ausgabe 17. Hg. von Rebecca Borschtschow/Lars C. Grabbe/Patrick Rupert-Kruse. Köln: Halem. S. 23–36.
- (2013b): Äquilibration und Synkretismus. Überlegungen zu einer interaktionistischen Theorie der Filmbildrezeption. In: *Multimodale Bilder. Zur synkretistischen Struktur des Filmischen*. Hg. von Lars C. Grabbe/Patrick Rupert-Kruse/Norbert M. Schmitz. Darmstadt: Büchner. S. 16–41.
- Hepp, Andreas (1999): *Cultural Studies und Medienanalyse. Eine Einführung*. Opladen/Wiesbaden: Westdeutscher Verlag.
- Jeong, Seung-Hoon (2013): *Cinematic Interfaces. Film Theory After New Media*, New York & London: Routledge.
- Keppeler, Angela (2010): Die wechselseitige Modifikation von Bildern und Texten in Fernsehen und Film. In: *Sprache intermedial. Stimme und Schrift, Bild und Ton*. Hg. von Arnulf Deppermann und Angelika Linke. Berlin & New York: De Gruyter. S. 447–467.
- Liebsch, Dimitri (2013): Wahrnehmung, Motorik, Affekt. Zum Problem des Körpers in der phänomenologischen und analytischen Filmphilosophie. In: *IMAGE – Zeitschrift für interdisziplinäre Bildwissenschaft*, Ausgabe 17. Hg. von Rebecca Borschtschow/Lars C. Grabbe/Patrick Rupert-Kruse. Köln: Halem. S. 103–123.
- Malafouris, Lambros (2007): Before and Beyond Representation: Towards an Enactive Conception of the Palaeolithic Image. In: *Image and Imagination: a Global History of Figurative Representation* Hg. von Colin Renfrew und Iain Morley. Cambridge: McDonald Institute for Archaeological Research. S. 289–302.
- Martin, Silke (2010): *Die Sichtbarkeit des Tons im Film. Akustische Modernisierungen des Films seit den 1920er Jahren*. Marburg: Schüren.
- Mazzola, Guerino (1997): *Semiotic Aspects of Musicology: Semiotics of Music*, <http://recherche.ircam.fr/equipes/repmus/mamux/documents/musiksemiotik.pdf> [26.06.2014].
- Merleau-Ponty, Maurice (2010): Das Kino und die neue Psychologie. In: *Philosophie des Films*. Hg. von Dimitri Liebsch. Paderborn: mentis. S. 70–79.
- Murch, Walter (1994): Foreword. In: Michel Chion: *Audio-Vision – Sound on Screen*. Hg. von Claudia Gorbman. New York: Columbia University Press. S. VII–XXIV.
- Nanay, Bence (2012): *The Multimodal Experience of Art*, Online: <http://webh01.ua.ac.be/bence.nanay/The%20multimodal%20experience%20of%20art.doc> [26.06.2014].
- Pauli, Jonathan (2011): *Sound-Design in der Vorproduktion. Wie das narrative Potential des Sound-Designs im Film voll ausgeschöpft werden kann*. Masterarbeit, Hochschule der Medien Stuttgart. Online: www.hdm-stuttgart.de/~curdt/Pauli.pdf [26.06.2014].
- Raffaseder, Hannes (2010): *Audiodesign*, 2. Auflage. München: Hanser..
- Sachs-Hombach, Klaus (2001): *Bildbegriff und Bildwissenschaft*. In: *kunst – gestaltung – design*, Heft 8. Hg. von Dietfried Gerhardus und Sigurd Rompza. Saarbrücken: St. Johann. S. 3–26.

- Schneider, Norbert J. (1997): *Komponieren für Film und Fernsehen. Ein Handbuch*. Mainz: Schott.
- Sonesson, Göran (1989): *Pictorial concepts: Inquiries into the semiotic heritage and its relevance for the analysis of the visual world*. Lund: University Press.
- Tagg, Philip (1999): *Introductory notes to the Semiotics of Music*. Online: <http://www.tagg.org/xpdfs/semi-otug.pdf> [26.06.2014].
- Welsch, Wolfgang (2003): *Ästhetisches Denken*, 6., erweiterte Auflage. Stuttgart: Reclam.
- Wulff, Hans J. (2012): Prolegomena zu einer Theorie des Atmosphärischen im Film. In: *Filmische Atmosphären* (Zürcher Filmstudien 30). Hg. von Philipp Brunner/Jörg Schweinitz/Margit Tröhler. Marburg: Schüren. S. 109–123.
- Wuss, Peter (1999): *Filmanalyse und Psychologie. Strukturen des Films im Wahrnehmungsprozeß*. Berlin: Edition Sigma.