

Auditives Storytelling im Spielfilm

Die Funktion der Tonspur im Kontext der Narration

Diplomarbeit

Ausgeführt zum Zweck der Erlangung des akademischen Grades

Dipl.-Ing. für technisch-wissenschaftliche Berufe

am Masterstudiengang Digitale Medientechnologien an der
Fachhochschule St. Pölten, **Masterklasse experimentelle Medien**

von:

Christoph Eymannsberger, BSc

DM151509

Betreuer/in und Erstbegutachter/in: Dipl.-Ing Thomas Wagensommerer, MA BA

Zweitbegutachter/in: Mag. Markus Wintersberger

[St. Pölten, 15.05.2018]

Ehrenwörtliche Erklärung

Ich versichere, dass

- ich diese Arbeit selbständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfe bedient habe.

- ich dieses Thema bisher weder im Inland noch im Ausland einem Begutachter/einer Begutachterin zur Beurteilung oder in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

Diese Arbeit stimmt mit der vom Begutachter bzw. der Begutachterin beurteilten Arbeit überein.

.....

Ort, Datum

.....

Unterschrift

Kurzfassung

Aufbauend auf der Arbeit im Audiodepartment des mittellangen Spielfilmes „Der Dunkle Reigen“ (Bücheler, 2017) sowie auf Erkenntnissen aus Wahrnehmungspsychologie und Psychoakustik wird in dieser Arbeit die oft unterschätzte Bedeutung der akustischen Kommunikation in Bezug auf die Narration im Spielfilm untersucht. Thematisch bewegt sich die Arbeit in den Gebieten Audioproduktion, Sounddesign und Storytelling. In der Theorie erarbeitete Gestaltungsmittel und -Konzepte werden anhand einer Analyse der verschiedenen Elemente der Tonspur (Sprache, Geräusche und Musik) in den Filmen „Mulholland Drive“ (David Lynch, 2001), „2001: A Space Odyssey“ (Stanley Kubrick, 1968), „Melancholia“ (Lars von Trier, 2011) sowie „Der Dunkle Reigen“ (Dominik Bücheler, 2017), hinsichtlich deren dramaturgischen und narrativen Funktion untersucht. Die Analyse zeigt, dass die Tonspur als eigenständiger Kommunikationskanal im Spielfilm einen wesentlichen Beitrag hinsichtlich der Dramaturgie und Narration leisten kann, vorausgesetzt, es werden die komplexen Bedeutungszusammenhänge, die durch die Verknüpfung der visuellen und auditiven Sinnesreize sowie den durch die Narration kommunizierten Informationen und Rahmenbedingungen im Prozess der Wahrnehmung bei den FilmzuseherInnen berücksichtigt.

Abstract

Based on findings retrieved from research in the fields of perceptive psychology and psychoacoustics, the objective of this thesis is to research the significance of acoustic communication concerning the narrative and dramatic composition in fiction film. Insights are provided into the interaction of audio production, sound design and storytelling. Key concepts regarding the narrative use of sound design in storytelling from other works in this field are gathered and validated by the analysis of the individual elements of the soundtrack (speech, music and effects) in David Lynch's „Mulholland Drive“ (2001), Stanley Kubrick's „2001: A Space Odyssey“ (1968), and Lars von Trier's „Melancholia“ (2011), as well as „Der Dunkle Reigen“ (2017), a short movie by Dominik Bücheler. The analysis confirms that the soundtrack is an independent communication channel which can contribute to the narration of the story, alongside with the image and the dramatic composition.

Inhaltsverzeichnis

Ehrenwörtliche Erklärung	II
Kurzfassung	III
Abstract	IV
Inhaltsverzeichnis	V
1 Einleitung	7
1.1 Thematik und Ziele der Arbeit	7
1.1 Forschungsfragen und Thesen	9
1.2 Aufbau und Methodik	11
1.3 Terminologie und Definitionen	12
2 Grundlagen der akustischen Kommunikation	14
2.1 Systematisierung von akustischen Ereignissen	14
2.1.1 Einteilung nach den Anforderungen des Spielfilms	14
2.1.2 Einteilung nach physikalischen Eigenschaften	20
2.2 Die auditive Wahrnehmung	24
2.2.1 Grundlagen der auditiven Wahrnehmung	24
2.2.2 Ebenen der auditiven Wahrnehmung	28
2.2.3 Wahrnehmbare Eigenschaften akustischer Ereignisse	30
2.2.4 Wahrnehmung von Sprache	46
3 Wirkung akustischer Ereignisse	49
3.1 Bedeutung erster Ordnung	50
3.2 Bedeutung höherer Ordnung	54
3.3 Formen der semantischen Verknüpfung	57
4 Die Tonspur als narratives Gestaltungsmittel	65
4.1 Grundlagen	65
4.2 Bild- Ton Beziehungen	69
4.2.1 Grundlagen	69
4.2.2 Bedeutung der Bild- Ton Beziehungen für die Narration	74
4.3 Hörperspektive	77
4.4 Konzepte und Techniken	80
4.4.1 Subjektivierung	80
	V

4.4.2	Die Beziehung zur Diegese	84
4.4.3	Konzepte der akustischen Gestaltung	87
5	Empirische Analyse	91
5.1	Methodik, Definition der Parameter	91
5.2	„Melancholia“ (Lars von Trier, 2011)	92
5.2.1	Narration und Dramaturgie	93
5.2.2	Tongestaltung	94
5.3	„Mulholland Drive“ (David Lynch, 2001)	99
5.3.1	Narration und Dramaturgie	99
5.3.2	Tongestaltung	103
5.4	„2001: A Space Odyssey“ (Stanley Kubrick, 1968)	106
5.4.1	Narration und Dramaturgie	107
5.4.2	Tongestaltung	109
5.5	„Der Dunkle Reigen“ (Dominik Bücheler, 2017)	113
5.5.1	Narration und Dramaturgie	113
5.5.2	Tongestaltung	114
6	Fazit	117
	Literaturverzeichnis	121
	Filmverzeichnis	125
	Abbildungsverzeichnis	126

1 Einleitung

1.1 Thematik und Ziele der Arbeit

„Das Sehen ist [...] ein motorisch aktiver Prozess, das Hören ist motorisch passiv“. (Görne 2017, S. 33–34) Die visuelle Wahrnehmung ist eine bewusste, muskulär gesteuerte Tätigkeit, während das Ohr über keine vergleichbare muskuläre Steuerung und der damit einhergehenden Filterung des wahrnehmbaren Bereichs verfügt. Die auditive Wahrnehmung arbeitet vielmehr unterbewusst. Der Gehörsinn hat im alltäglichen Leben oft nur die Rolle eines Hilfsorgans, das Informationen reflexartig und unbewusst zur Unterstützung des Sehsinnes verarbeitet. Völlig bewusst wahrgenommen werden akustische Informationen meist nur bei unzureichender Stimulierung des Sehsinnes (Sonnenschein 2001, S. 72). Im Gegensatz zum Auge kann das Ohr jedoch nicht verschlossen werden, es nimmt Schall grundsätzlich aus allen Richtungen wahr und wertet rund um die Uhr die akustische Umwelt aus, auch im Schlaf“ (Görne 2017, S. 33–34). Dadurch wird die Tongestaltung von den FilmzuseherInnen oft erst dann bewusst wahrgenommen, wenn ein offensichtliches Problem vorliegt, welches es erschwert, die vermittelten Informationen aufzunehmen (beispielsweise bei Unverständlichkeit der Sprache oder beim Auftreten von technischen Gebrechen).

Die Passivität des Hörens erklärt einerseits die Problematik, dass der Tonspur in der Produktion oft wesentlich weniger Aufmerksamkeit gewidmet wird als der Bildebene, sowie dass die akustische Kommunikation und deren Möglichkeiten in der Filmanalyse und -theorie oft nicht oder nur unzureichend behandelt werden. (Flückiger 2007, S. 100) Andererseits erklärt sie die unterschwellige Natur der akustischen Kommunikation, worauf die Möglichkeiten der Tongestaltung, dem Filmpublikum unbemerkt Informationen zu vermitteln und unterbewusst Emotionen zu beeinflussen, basieren. (Raffaseder 2010, S. 19)

Sound, much more than the image, can become an insidious means of affective and semantic manipulation. On one hand, sound works on us directly, physiologically (breathing noises in a film can directly affect our own respiration). On the other, sound has an influence on perception: through the phenomenon of added value, it interprets the meaning of the image, and makes us see in the image what we would not otherwise see, or would see differently. (Chion und Gorbman 1994, S. 20)

Wie bereits Kritiker zur Zeit der Einführung des Tonfilms festgestellt haben, sollte es sich bei der Gestaltung der auditiven Kommunikation im Spielfilm nicht um die bloße Umformulierung der auf der Bildebene gezeigten Inhalte handeln (wie es auch heute manchmal der Fall ist), sondern die Tonebene kann ein eigenständiges Mittel in der Erzählstruktur darstellen, nicht nur zur Ergänzung der visuellen Eindrücke. (Flückiger 2007, S. 133–134) Ein echter Mehrwert kann jedoch nur entstehen, wenn die Wirkung und die Bedeutung der verschiedenen Elemente der Tonspur auf das Filmpublikum berücksichtigt werden, zum einen für sich alleine, zum anderen im Gesamtkontext (der im Spielfilm durch die Narration vorgegeben ist), sowie auch in Kombination mit den durch das Bild vermittelten Informationen.

„Zu oft werden Tonspuren im medialen Kontext weitgehend nach subjektiven ästhetischen Vorlieben konzipiert, produziert und bewertet, während es an objektiven Qualitätskriterien nach wie vor mangelt.“ (Raffaseder 2010, S. 236) Wie Hannes Raffaseder kritisiert auch Jörg Lensing in diesem Zusammenhang das zunehmende Bedürfnis nach Individualismus vieler junger Filmemacher, die im Bestreben besonders originell zu sein, kreative Entscheidungen rein aus dem Bauch heraus treffen, ohne sich der Wirkung der Tongestaltung auf das Filmpublikum bewusst zu sein. Diesem Ansatz zufolge seien Probleme, die bei den RezipientInnen bei der Interpretation der Inhalte, aufgrund eines unbedachten Umgangs mit den Gestaltungselementen, auftreten, selbst verschuldet. Ohne von „seichter Massenunterhaltung, geschweige denn einer Stagnation der gestalterischen Entwicklung“ zu sprechen, plädiert Lensing für eine bewusste Auseinandersetzung mit filmischen und außer-filmischen Hör-Konventionen und einem darauf aufbauenden, bedachteren Umgang mit solchen. „Es bleibt zu hoffen, dass der [...] Film, alleine schon aus ökonomischem Druck diese Irrungen des

Individualismus nie so weit treiben kann, dass er aufhört mit einem Publikum zu kommunizieren.“ (Lensing 2006, S. 198–200)

Im Zuge der Themenfindung und Vorbereitung der Arbeit wurden die O-Ton Aufnahmen während der Produktion, sowie der Dialog-Schnitt in der Postproduktion des mittellangen Spielfilms „Der Dunkle Reigen“ (Bücheler 2017) durchgeführt. Die Erfahrungen aus diesem Projekt lieferten der Grundstock für die Erarbeitung der Problematik. Aufbauend darauf ist das Ziel dieser Arbeit, die Möglichkeiten der akustischen Kommunikation im Bereich fiktionaler Narration im Spielfilm zu analysieren, herauszuarbeiten und deren Anwendung aufzuzeigen. Sie kann zum einen als Hilfestellung und Nachschlagewerk hinsichtlich der Tongestaltung dienen. Andererseits soll die Arbeit ein Bewusstsein für das Thema schaffen und falls noch nicht vorhanden, Filmproduzent/Innen dazu anregen, Schlüsselpersonen des Sounddepartments bereits sehr bald in die Planung miteinzubeziehen, um schon in der Vorproduktion die Möglichkeiten der Tongestaltung zu erörtern und in der Konzeption und der anschließenden Produktionsphase einzuplanen, um so das häufig nicht genutzte Potential der akustischen Kommunikation ausloten zu können. Ebenso sollen Unterschiede und Gemeinsamkeiten in der akustischen Gestaltung in Bezug auf den narrativen Kontext und der Dramaturgie von Spielfilmen erörtert werden, sowie die Unterschiede zwischen einem österreichischen Independent Film zu größeren, internationalen Produktionen hinsichtlich dieser Thematik darstellen.

1.1 Forschungsfragen und Thesen

Die wesentliche Fragestellung, die im Zuge dieser Arbeit beantwortet werden soll, ist jene nach der Funktion der Tongestaltung in Bezug auf die Dramaturgie, sowie deren Rolle im Gesamtkontext der Handlung. Um die Grenzen des Themengebietes zu definieren, bezieht sich die Arbeit auf den Bereich der fiktionalen Narration im Medium Film. Viele der Grundvoraussetzungen und Konzepte lassen sich aber auch auf andere Bereiche der akustischen Gestaltung, insbesondere anderer audio-visueller Medien,

anwenden. Dafür ist eine Darstellung der Grundvoraussetzungen der akustischen Kommunikation im Spielfilm notwendig, die durch die Beantwortung folgender Fragen gegeben ist:

Welche Wirkungen und Bedeutungen können durch die akustische Kommunikation ausgelöst bzw. vermittelt werden?

Wie werden diese Bedeutungen kommuniziert?

Darauf aufbauend orientiert sich die Erarbeitung der Funktion der Tongestaltung im Spielfilm in Bezug auf Dramaturgie und Narration an folgenden Fragestellungen:

Wie kann durch die akustische Gestaltung ein Mehrwert in Bezug auf Narration und Dramaturgie im Spielfilm entstehen?

Welche zugrundeliegenden Konzepte und Theorien kommen zur Anwendung?

Welche Unterschiede zeigen sich in der Analyse der zu untersuchenden Filmen, und welche Gemeinsamkeiten?

Sind die in der Theorie dargestellten Konzepte ausreichend und deren Anwendung sinnvoll, um einen Mehrwert in Bezug auf Narration und Dramaturgie im Spielfilm zu schaffen?

Worin unterscheidet sich ein österreichischer Independent Film von größeren, internationalen Produktionen in Bezug auf die narrative Funktion der Tonspur?

Zum jetzigen Zeitpunkt kann man folgende Thesen als Ergebnis der im Zuge der Arbeit durchgeführten Recherchen und Analysen formulieren:

Der Einsatz von Sprache, Geräuschen und Musik als narratives Element wird bei Spielfilmen berücksichtigt. Die Tonspur ist nicht nur bloßes Beiwerk zum Bild, sondern ein eigenständiges Instrument zur dramaturgischen Gestaltung im Spielfilm.

Die in der Theorie erarbeiteten Konzepte sind in der Tongestaltung hilfreich und schaffen darüber hinaus eine Grundlage für die nachfolgende Analyse in Bezug auf die Thematik.

Die analysierten Filme unterscheiden sich konzeptuell, sowie in der Ausführung bzw. im Stil der Produktion voneinander. Es lassen sich aber gemeinsame Ansätze, Methoden und Techniken herausarbeiten, mit denen in allen Filmen gearbeitet wurde.

1.2 Aufbau und Methodik

Zur Darstellung der theoretischen Grundlagen und des aktuellen Wissensstandes, sowie zur Analyse der Filme ist eine umfassende Literaturrecherche notwendig. Ziel der Recherche ist es, den derzeitigen Stand der Forschung zu erfassen und die Analyse darauf aufzubauen. Ausgehend von den in der Literatur beschriebenen Konzepten werden Erkenntnisse hinsichtlich der Forschungsfragen erarbeitet und diese Erkenntnisse anhand der Analyse von vier Produktionen überprüft.

In Kapitel 2 werden die notwendigen Grundlagen der akustischen Kommunikation sowie der auditiven Wahrnehmung in Bezug auf die akustische Gestaltung von Spielfilmen erarbeitet und in Kapitel 3 die möglichen Wirkungen sowie die Bedeutungen, die die Elemente der Tonspur für das Filmpublikum erlangen können dargestellt. In Kapitel 4 werden die hinsichtlich der Dramaturgie und Narration wesentlichen, theoretischen Ansätze und Gestaltungskonzepte erarbeitet, welche als Grundlage für eine in diesem Zusammenhang zielführende Tongestaltung dienen und auf denen die anschließende Analyse aufbaut.

Durch die analytische Aufarbeitung und den Vergleich der akustischen Gestaltung unterschiedlicher Produktionen in Bezug auf deren Handlung und Dramaturgie werden die in der Theorie erarbeiteten Konzepte und die gewonnenen Erkenntnisse überprüft, wodurch eine objektive Darstellung der Möglichkeiten akustischer Gestaltung von Spielfilmen in Bezug auf die Forschungsfrage gegeben ist. Die Filme „Mulholland Drive“ (David Lynch, 2001), „2001: A Space Odyssey“ (Stanley Kubrick, 1968) und „Melancholia“ (Lars von Trier, 2011) sollen dabei einen Überblick über das Medium „Spielfilm“ bieten, um die Forschungsfragen genre-übergreifend und von der Zeit der Entstehung unabhängig beantworten zu können.

Analysiert werden dabei die in Kapitel 2.1.1 beschriebenen Elemente der Tonspur (Sprache, Geräusche und Musik). Deren Bedeutung wird zuerst alleine und anschließend in Bezug auf die Bildebene erörtert, sowie die Bedeutung im narrativen Gesamtkontext festgelegt. Durch die Aufteilung in verschiedene Elemente der Tonspur

lassen sich die meist in Klangkulissen verwobenen Elemente bewusster wahrnehmen und analysieren. So kann deren Funktion in dramaturgisch wichtigen Abschnitten der Filme, im Kontext der anderen Tonelemente sowie in Bezug zum Bild bestimmt werden und die unterschiedlichen Ansätze der Tongestaltung hinsichtlich deren narrativer Funktion herausgearbeitet werden.

1.3 Terminologie und Definitionen

Im Folgenden sollen einige Begriffe erklärt werden, die für das grundlegende Verständnis dieser Arbeit wichtig sind:

Die Verwendung des Begriffs Film bezieht sich in dieser Arbeit nicht auf das physikalische, zur analogen Aufzeichnung und Wiedergabe verwendete Trägermedium, sondern auf die, im Englischen sehr treffend als *Motion Picture* bezeichnete und zum Massenmedium gewachsene Kunstform. Die in dieser Arbeit behandelte Gattung ist der Spielfilm, wobei sich der Begriff immer auf diese Gattung bezieht, sofern nicht im Text explizit auf eine andere Bedeutung hingewiesen wird. Beim Spielfilm werden „Ereignisse mit Schauspielern nach dramaturgischen und ästhetischen Gesichtspunkten in Szene gesetzt“. (Keutzer et al. 2014, S. 279)

„Auf der Ebene der Filmgestaltung meint Dramaturgie die konkrete Bauform des Werks, seine Erzählweisen, Gliederung und Komposition.“ (Keutzer et al. 2014, S. 195) Die Ereignisse werden narrativ, also in erzählender Form, vermittelt und sind oft fiktional. Im Englischen bietet sich hier auch der Begriff *fictional film* an, der aber auch Animationsfilme umfasst, die in dieser Arbeit nicht behandelt werden. Jedoch kann es auch Spielfilme mit nicht- fiktionaler Narration geben, wie beispielsweise Autobiographien. David Bordwell bezeichnet die Narration im Spielfilm als einen Prozess, in dem durch die konkrete Anordnung der Ereignisse, sowie deren Darstellung durch die kognitiven Leistungen der FilmzuseherInnen eine Geschichte konstruiert wird. (Bordwell 1985, S. 53) „Eine Analyse der Narration steht in der Tradition der literaturwissenschaftlichen Erzähltheorie und fragt in grundlegender Weise danach,

was eine Erzählung oder Geschichte konstituiert, wie Erzählungen aufgebaut sind und ob hinter der Vielfalt des Erzählten wiederkehrende Strukturen zu erkennen sind.“ (Keutzer et al. 2014, S. 194)

Der Begriff Sounddesign wird in der Arbeit gleichbedeutend mit der akustischen Gestaltung bzw. Tongestaltung im Spielfilm verwendet.

Ein weiterer wichtiger Begriff ist das *Schallsignal*, das einerseits in einem informationstheoretischen Kontext eine akustische Information darstellt, wohingegen in einem technischen Kontext die Darstellung der Amplitude von Schall im Verhältnis zur Zeit oder zur Frequenzzusammensetzung gemeint ist. Als *akustisches Ereignis* oder *Schallereignis* werde ich im Folgenden den physikalischen Prozess bezeichnen, dessen Eigenschaften technisch messbar sind und der allen wahrnehmbaren Geräuschen oder Klängen zugrunde liegt. Der Begriff *Hörereignis* bezieht sich auf die subjektive Wahrnehmung von akustischen Ereignissen und beschreibt die durch ein Schallereignis ausgelöste Hörempfindung, die im Prozess der auditiven Wahrnehmung gebildet wird und der, wie in Kapitel 3 beschrieben von den RezipientInnen eine Bedeutung zugewiesen wird.

2 Grundlagen der akustischen Kommunikation

2.1 Systematisierung von akustischen Ereignissen

2.1.1 Einteilung nach den Anforderungen des Spielfilms

Nach Hannes Raffaseder werden die in audiovisuellen Medien verwendeten akustischen Informationen grundsätzlich in die Kategorien Sprache, Musik und Geräusche unterteilt.

Bei der Sprache unterscheidet man weiter zwischen Dialog (die Sprache entstammt den in der Szene agierenden Personen und ist Teil der filmischen Realität) und Off-Text bzw. Kommentar (die Sprache ist nicht Teil der filmischen Realität). (Raffaseder 2010, 241) In der Filmproduktions- Praxis wird der Kommentar auch als Off- Voice oder als Voiceover bezeichnet. Jörg Lensing verwendet für die Unterteilung der Sprache die Begriffe On-Voice als Bezeichnung für Dialoge und Monologe (also der Sprache der Charaktere im Bild) und den Begriff Off-Voice für innere Monologe und Kommentare. (Lensing 2006, S. 48)

Die Tonspur im Spielfilm ist wie die auditive Wahrnehmung im Allgemeinen (mehr dazu in Kapitel 2.3.3.) auf die menschliche Stimme und die Sprache fokussiert. Sind in der akustischen Umwelt durch das Sprachorgan produzierte Laute, Silben und Worte vorhanden, werden diese vor anderen akustischen Ereignissen wahrgenommen und rücken diese in das Zentrum der Aufmerksamkeit. (Chion und Gorbman 1994, S. 6) Im Kontext der Narration gibt es die so genannte theatrical speech (=schauspielerische Sprache), die die Kommunikation der Charaktere untereinander beschreibt. Dagegen wird unter textual speech meist eine Stimme aus dem Off bzw. ein Kommentar verstanden, die zu einer Erzählerfigur gehört oder Gedanken oder Erinnerungen von Charakteren wiedergibt. Die emanation Speech beschreibt ganz oder teilweise

unverständliche Sprache von Charakteren, deren Bedeutung in keinem Bezug zur Handlung steht. Die Fokussierung auf den semantischen Inhalt der Sprache wird daher vermieden. Diese Form wird zur Gestaltung der akustischen Umgebung verwendet, um zum Beispiel eine Restaurant- Atmosphäre zu schaffen, ohne von der eigentlichen Handlung abzulenken. (Sonnenschein 2001, S. 156–157)

Die in der Gestaltung der Tonspur verwendeten Geräusche sind Atmos und Effekte. „Unter einer Atmo werden jene Hintergrundgeräusche verstanden, die für die akustische Wahrnehmung einer Umgebung charakteristisch sind.“ (Raffaseder 2010, S. 242) Die im Alltag meist nicht bewusst wahrgenommenen Geräuschkulissen bestimmter Orte sind für die räumliche und zeitliche Beschreibung einer Szene sehr wichtig, da diese ohne die charakteristischen Hintergrundgeräusche unnatürlich wirkt. Durch die Atmo werden Informationen über die Uhrzeit, die Jahreszeit und auch über die geschichtliche Epoche vermittelt, oder beispielsweise darüber, ob die Szene in einem Raum oder im Freien, in einer ländlichen Gegend oder im urbanen Bereich stattfindet. Auch für die vermittelte Stimmung ist unter anderem auch die Atmo entscheidend, da bestimmte akustische Umgebungen verschiedene Assoziationen des Filmpublikums hervorrufen können. Aus diesem Grund ist die Verwendung von Atmos von den Originalschauplätzen, abhängig von der gewünschten Wirkung, nicht immer zielführend, sofern es sich nicht um eine dokumentarische, quasi-realistische Darstellung der akustischen Umgebung handelt. (Lensing 2006, S. 104–108) In der Postproduktion werden zusätzlich zur Originalaufnahme der Atmo vom Filmset noch weitere Atmo- Aufnahmen aus Sound-Bibliotheken hinzugefügt, um eine mehrdimensionale, klanglich dichtere und im narrativen Kontext wertvolle akustische Umgebung zu schaffen. (Lensing 2006, S. 104–105)

Geräusche, die bewusst eingesetzt werden, um die Aufmerksamkeit auf bestimmte Ereignisse zu lenken und so einen noch engeren Bezug zur Handlung aufweisen, werden als Soundeffekte bezeichnet. Hard-Effects sind Geräusche „die ganz synchron zu einem im Bild sichtbaren Ereignis gesetzt werden“ und so die visuellen Eindrücke verstärken bzw. verdeutlichen. Bei Soft-Effects ist die Synchronität vom visuellen und akustischen Ereignis nicht zwingend notwendig. Sie stellen eine subtilere Form der Unterstützung der Geschehnisse auf der Bildebene dar, oder auch eine von dieser

unabhängige, rein akustische Kommunikation. Soft-Effects werden häufig auch mit den übrigen Hintergrundgeräuschen in die Atmo eingebettet, mit dem Unterschied, dass diese im Vergleich zu den übrigen Hintergrundgeräuschen einen größeren narrativen Wert in Bezug auf die Handlung aufweisen. Eine besondere Vorgehensweise in der Gestaltung der Soundeffekte ist der „selektive Geräuscheinsatz“, worunter man die gezielte Hervorhebung einiger weniger Geräusche versteht, während die akustische Umwelt ausgeblendet bzw. in den Hintergrund gerückt wird. Der selektive Geräuscheinsatz ermöglicht es, die Aufmerksamkeit des Filmpublikums noch stärker auf einige wenige, bestimmte Ereignisse zu lenken. Sogenannte Keysounds werden meist als Soft-Effects oder als Bestandteil der Atmo eingesetzt, um bestimmte, meist durch Konvention etablierte Assoziationen mit Orten, Zeiten oder Szenarien zu kommunizieren. So ist beispielsweise in fast allen Filmen in den Nachtszenen ein bellender Hund zu hören. (Raffaseder 2010, S. 243–245) Im Zusammenhang mit Effekten wird oft der Begriff *Larger-than-life* verwendet, womit die bewusste Überzeichnung von Schallereignissen gemeint ist. (Raffaseder 2010, S. 245) Dabei wird auf eine realistische Darstellung des akustischen Ereignisses zugunsten einer durch Verfremdung und Layern (=das Überlagern verschiedener Geräusche, die mitunter in keiner Verbindung mit der eigentlichen Schallquelle stehen) erzeugten, eindrucksvoller wirkenden Darstellung verzichtet. (Sonnenschein 2001, 55,60)

Neben Sprache und Geräuschen zählt auch Musik zu den bedeutenden Bestandteilen in der akustischen Gestaltung von Spielfilmen.

Die klassische Filmmusik hat ihren Ursprung in der Begleitung von Filmvorführungen durch Pianisten, Organisten oder auch eines ganzen Orchesters in der Stummfilmära Anfang des 20. Jahrhunderts vor der Einführung des Tonfilms. (Lensing 2006, S. 152) Diese Form der Musik ist nicht als eigenständiges Musikstück gedacht (obwohl die Filmmusik oft zusätzlich als Soundtrack vermarktet wird), sondern im Gesamtkontext eines Filmes zu sehen. Das bedeutet, dass sie komponiert wird, um eine bestimmte Funktion in Kombination mit der Bildebene und den anderen Elementen der Tonspur zu erfüllen oder um eine bestimmte Wirkung im narrativen Kontext zu erzielen. Eine besondere Form der Filmmusik ist Source-Music, welche dadurch gekennzeichnet ist, dass „die Quelle [...] im Bild sichtbar oder zumindest verifizierbar ist.“ Beispiele dafür sind

Radios, Straßen- oder Livemusiker, die Teil der Szene sind. Dabei ist es wichtig, dass die klanglichen Eigenschaften der Source-Music der im Bild zu sehenden Schallquelle entsprechen und dass die Eigenschaften des Raumes, in der sich diese befindet, berücksichtigt werden, da hier beträchtliche Unterschiede zu hören sind. (Raffaseder 2010, S. 246)

Die Funktion der Musik kann dramaturgisch, also zur Unterstützung des Spannungsverlaufs sein, wobei dabei oft filmmusikalische Stereotypen bedient werden, wie zum Beispiel schnelle, treibende Rhythmen in Actionfilmen, die eine Verfolgungsjagd untermalen. (Raffaseder 2010, S. 247) Subtiler eingesetzt funktioniert die Musik oft als *emotional signifier* (=emotionale Referenz, emotionaler Vorbote), wodurch emotionale Prozesse der Charaktere oder wichtige Ereignisse verdeutlicht werden. (Sonnenschein 2001, S. 155) Die Beeinflussung von Stimmung und Emotionen in eine bestimmte Richtung bezeichnet Hannes Raffaseder als die persuasive Funktion der Filmmusik. (Raffaseder 2010, S. 247–248) Eine ebenfalls sehr gebräuchliche Einsatzmöglichkeit ist die narrative bzw. epische Funktion, wobei es sich um die Etablierung von Ort und Zeit der Handlung durch musikalische Symbole und Klischees (also durch Musikstile oder -stücke, sowie Instrumente, die für ein bestimmtes Land oder eine bestimmte Zeitepoche üblich sind) handelt (Lensing 2006, S. 48). So wird beispielsweise ein Dudelsack für gewöhnlich mit Schottland assoziiert (Raffaseder 2010, S. 247–248) oder manche Popsongs mit deren Entstehungsjahr oder einem politischen oder zeitgeschichtlichen Ereignis. Nicht zuletzt hat Filmmusik auch eine strukturelle Funktion. Beispielsweise können durch deren Einsatz Schnitte verdeckt oder betont werden, sowie mehrere Szenen wie eine Einheit erscheinen lassen oder voneinander trennen. Zusätzlich können Wiederholungen oder Variationen von bereits etablierten musikalischen Motiven und Themen auch zur Unterstützung der Handlung eingesetzt werden um Charaktere zu beschreiben oder um Verbindungen zwischen unterschiedlichen, nicht zusammenhängenden Szenen des Films herzustellen (Sonnenschein 2001, S. 155). Dabei werden solche sich wiederholende Strukturen als Leitmotive bezeichnet.

Die Einteilung der Tonspur in Sprache, Geräusche und Musik beruht unter anderem auf den technischen und logistischen Anforderungen der Filmproduktion, da bei größeren

Projekten meist unterschiedliche Personen bzw. Teams für die Bearbeitung der eben besprochenen Kategorien zuständig sind und die unterschiedlichen Elemente erst in der Endmischung zu der finalen Tonspur zusammengefügt werden. (Raffaseder 2010, S. 17)

Hinsichtlich der narrativen Struktur ist in der Filmproduktion noch eine andere Art der Einteilung üblich, die sich auf die Unterscheidung der Informationen nach deren Verortung oder Nicht- Verortung in der filmischen Realität bezieht. Für diese Unterscheidung wird in der Erzähltheorie der Begriff Diegese verwendet, den beispielsweise Susan Hayward folgendermaßen definiert:

Diegesis refers to narration, the content of the narrative, the fictional world as described inside the story. In film it refers to all that is really going on on-screen, that is, to fictional reality (Hayward 1996, S. 67)

Genauer bezeichnet die Diegese die Konstruktion einer zeitlich und räumlich sinnvollen Vorstellung der erzählten Welt, wobei diese Vorstellung individuell vom Autor, von den Schauspielern und schließlich von den FilmzuseherInnen gebildet wird. Nicht zu verwechseln ist der Begriff Diegese aber mit der aus dem Altgriechischen stammendem Wort *diêgêsis*, was übersetzt Erzählung bzw. Erörterung bedeutet. Allerdings besteht insofern ein Zusammenhang, da die Diegese durch eine Diegesis vermittelt wird. (Mag. Dr. Anton Fuxjäger 2007, S. 18–19) Das Adjektiv diegetisch bezieht sich daher auf Ereignisse und Handlungen, die innerhalb der erzählten Welt stattfinden. Typische Beispiele für nicht-diegetische Elemente auf visueller Ebene stellen Inserts, Titel oder andere eingeblendete Textelemente dar.

Nach Thomas Gröne und David Sonnenschein können auch die akustischen Informationen einer Tonspur diegetisch oder nicht diegetisch sein. (Sonnenschein 2001, S. 152) Für eine einfachere Sichtweise lässt sich die Vorstellung heranziehen, dass alle akustischen Ereignisse, die auch von den Charakteren innerhalb der filmischen Welt bzw. einer Szene gehört werden können, als diegetisch bezeichnet werden. Dazu zählen Dialoge oder beispielsweise Musik aus einem Radio in der Szene, in der die Charaktere verortet sind.

Nicht diegetische Elemente der Tonspur sind entweder extradiegetisch (also nur für das Filmpublikum hörbare Elemente wie Filmmusik oder abstrakte Soundeffekte) oder metadiegetisch. Darunter sind Hörereignisse zu verstehen, die nur in der Wahrnehmung eines Charakters im Film existieren, wie zum Beispiel innere Monologe des Protagonisten. Die diegetischen Klänge und Geräusche können für eine dokumentarische, quasi-realistische Gestaltung der Tonspur bereits ausreichend sein, oder „die Grundlage eines aufwändigeren Sounddesigns mit meta- und extradiegetischen Elementen“ bilden. (Görne 2017, S. 20)

Weiter unterscheidet David Sonnenschein bei diegetischen Sounds zwischen *on-screen* und *off-screen*, also danach, ob die Schallquelle im Bild zu sehen ist. Ein entscheidender Unterschied besteht in der Notwendigkeit der Synchronität von Bild und Tonebene bei *on-screen* Sounds. Akustische Ereignisse, deren Schallquelle nicht im Bild sichtbar ist, aber die von den Charakteren der filmischen Realität gehört werden, fallen daher auch in die diegetische Kategorie. Diese können wiederum aktiv oder passiv sein, wobei aktive *off-screen* Sounds Aufmerksamkeit erregen, bzw. das Filmpublikum neugierig machen, wer oder was die Schallquelle ist und woher die Klänge oder Geräusche kommen. Passive *off-screen* Sounds eignen sich dagegen dafür, die akustische Umgebung der Szene zu gestalten (zum Beispiel mit Atmos) oder auch um flüssige Übergänge bei Schnitten der Bild- oder der Tonebene zu realisieren. Als eigene Kategorie führt Sonnenschein sogenannte *off-track* Sounds ein, worunter akustische Ereignisse fallen, die auf der Tonspur nicht vorhanden sind, jedoch durch die Handlung und die Diegese die Vorstellung entsteht, als wären diese vorhanden. Als Beispiel hierfür kann ein Telefonat dienen, bei dem zwar der Charakter seinen Gesprächspartner hört, das Filmpublikum aber nur anhand der Reaktionen auf die Aussagen des Gegenübers schließen kann. Nicht-diegetische Sounds funktionieren im Allgemeinen meist als interpretatives Element und werden eingesetzt um emotionale Andeutungen zu machen. (Sonnenschein 2001, S. 152–154)

2.1.2 Einteilung nach physikalischen Eigenschaften

Eine notwendige Voraussetzung, um akustische Informationen sinngemäß kommunizieren und bearbeiten zu können, ist die Eigenschaften von akustischen Ereignissen zu identifizieren sowie deren Parameter zu beschreiben.

Nach David Sonnenschein sind die für die akustische Gestaltung relevanten Eigenschaften der **Rhythmus** (beschreibt die Abfolge bzw. die zeitliche Anordnung von akustischen Ereignissen im makroskopischen Zeitbereich, also z. B. regelmäßig oder unregelmäßig), die **Intensität** der Schwingung (gemessen in Dezibel [dB]) und die **Tonhöhe** (definiert durch die Frequenz [Hz]). Weiters beschreibt die **Klangfarbe**, auch Timbre genannt, die Gesamtheit von Klängen (periodischen Schwingungen) und Geräuschen (unregelmäßigen Schwingungen) sowie folglich auch die Zusammensetzung des Spektrums der Information und ermöglicht zum Beispiel die Unterscheidung von verschiedenen Musikinstrumenten durch die unterschiedlichen vom Instrument verursachten Obertöne. Die **Form** des Schallereignisses auf einer makroskopischen Zeitskala (auch Envelope genannt) beschreibt, ob es sich um eine impulsartige Schwingung oder um ein kontinuierliches, gleichbleibendes Signal handelt. (Sonnenschein 2001, S. 65–69) Besonders die Einschwingphase (Attack-Time) von Klängen oder Geräuschen mit einem semantischen Mehrwert ist wichtig, um diese einordnen bzw. erkennen zu können. Wird zum Beispiel die Aufnahme eines Musikinstrumentes rückwärts abgespielt, fehlen die für den charakteristischen Klang eines Instrumentes verantwortlichen Anschlagsgeräusche am Anfang der Töne. (Sonnenschein 2001, S. 93) Die **Geschwindigkeit** (schnell oder langsam) definiert bei impulsartigen Signalen den Abstand dieser, und die **Organisation** bestimmt, ob die Information einem System folgt (zum Beispiel den Regel der Grammatik einer Sprache) oder ob diese rein zufällig angeordnet ist. (Sonnenschein 2001, S. 65–69)

Hannes Raffaseder beschreibt die Tonhöhe, die Lautstärke und die Klangfarbe als die drei wesentlichen Eigenschaften von Schallereignissen. Weiters unterscheidet man zwischen Klängen (Tonhöhe, Lautstärke und Klangfarbe sind bestimmbar) und Geräuschen (nur die Lautstärke ist eindeutig identifizierbar). (Raffaseder 2010, S. 35)

Die als Töne und Geräusche wahrgenommenen Luftdruckschwankungen sind Wellen bzw. Schwingungen und demnach eine zeitabhängige Veränderung einer physikalischen Größe. Im Falle der auditiven Wahrnehmung handelt es sich um eine Veränderung des Luftdrucks. Finden diese Luftdruckschwankungen zwischen 20 und 20000 mal in der Sekunde statt, befinden sie sich in dem für den Menschen hörbaren Bereich und es handelt sich somit um Schallwellen. Die Anzahl der Schwingungen pro Sekunde wird als die Frequenz (gemessen in Hz) bezeichnet. Der Weg, den die Schallwelle in dieser Zeit zurücklegt, wird als die Wellenlänge bezeichnet. Die Amplitude bezeichnet die maximale Auslenkung, also das Luftdruckmaxima und somit den Wert der größten Auslenkung der Schwingung. (Hellbrück und Ellermeier 2004, S. 54)

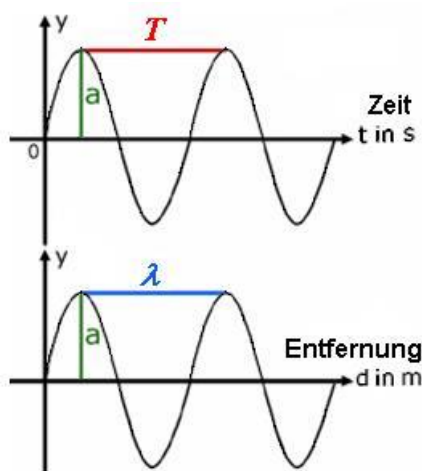


Abbildung 1: (Sengpiel 2017a)
y ... Schalldruck (p)
a ... Amplitude
T ... Periodendauer der Schwingung
λ ... Wellenlänge

Es besteht ein direkter Zusammenhang zwischen der Wellenlänge λ und der Frequenz f ($\lambda = c/f$), wobei c die Ausbreitungsgeschwindigkeit darstellt. Diese hängt von der Dichte und der Elastizität des Mediums ab, in dem sich die Schallwellen ausbreiten, und beträgt in der Luft bei einer Temperatur von 20°C etwa 343m/s . Bei einer Temperaturänderung weicht diese um einige wenige Prozent ab. (Friesecke 2007, S. 21–22)

Treten die Luftdruckschwingungen periodisch auf, also bestehen sie aus regelmäßig wiederkehrenden Mustern, spricht man von Klängen, da sich aus der Periodendauer T die Grundfrequenz und damit die Tonhöhe des akustischen Ereignisses ableiten lässt. (Raffaseder 2010, S. 58) Wichtig ist dabei, dass sich natürlich vorkommende Klänge nicht

aus einer einzigen, sondern einer Vielzahl von Schwingungen mit unterschiedlicher Periodendauer zusammensetzen. (Friesecke 2007, S. 119) Die Periodendauer entspricht dabei der tiefsten im Schallsignal verkommenden Teilfrequenz (Friesecke 2007, S. 154), wobei der Zusammenhang $f = 1/T$ gilt. (Hellbrück und Ellermeier 2004, S. 57) Alle anderen Frequenzen sind ganzzahlige Vielfache dieser Grundfrequenz und werden als Obertöne bezeichnet. (Raffaseder 2010, S. 59) Die Ausbreitung von Schall ist auch in anderen Medien als Luft möglich, im Falle der auditiven Wahrnehmung handelt es sich allerdings in der Regel um die Übertragung von Schallwellen in der Luft. (Hellbrück und Ellermeier 2004, S. 54)

Ein wesentlicher Parameter in der Entstehung, Ausbreitung und Wahrnehmung von Schall, sowie in der Beschreibung physikalischer Eigenschaften von Schallereignissen, ist die Zeit. Da Schall ein flüchtiges Medium darstellt (er beruht grundsätzlich auf der Veränderung von physikalischen Größen in Abhängigkeit von der Zeit), ist ein Anhalten oder der Stillstand nicht möglich. (Raffaseder 2010, 21,24, 40) Deutlich wird dies am Beispiel einer pausierten Videoaufnahme, bei der man zwar das Standbild sehen, jedoch die Tonspur keine Momentaufnahme wiedergeben kann. Bei statischen Klängen oder Geräuschen handelt es sich um sich wiederholende Schwingungen, die sehr knapp beisammen liegen und so als ein kontinuierliches Ereignis wahrgenommen werden.

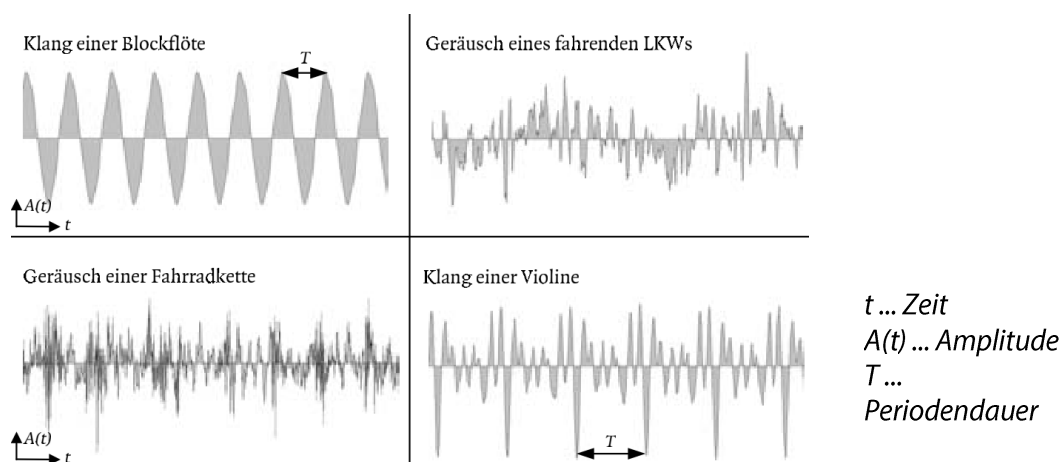


Abbildung 2: Darstellung des Signalverlaufs verschiedener Klänge und Geräusche (bzw. deren Amplitude) in Abhängigkeit von der Zeit auf mikroskopischer Ebene. (Raffaseder 2010, S. 58)

Die Betrachtung von Schallsignalen im Zeitbereich kann auf drei Arten erfolgen. Der mikroskopische Zeitbereich, also die Betrachtung des Signales im Bereich von 0,05 bis 50 Millisekunden stellt die einzelnen Druckmaxima und - Minima, bzw. deren Amplitudenwerte, dar. Zusätzlich lassen sich Eigenschaften wie die Periodendauer von Klängen (mehr dazu in Kapitel 2.2.4.2. und 2.3.3.4.) ablesen. Der „Zeitbereich der Übergangsklänge und Transienten“ fällt in den Bereich von 50-150 Millisekunden und eignet sich insbesondere, um zum Beispiel die für die Erkennung der Klangfarbe wesentliche Einschwingphase (=Transient) eines akustischen Ereignisses zu betrachten.

Im „makroskopischen Zeitbereich“ (die Darstellung des Signalverlaufs im Bereich 150 Millisekunden) lassen sich Eigenschaften wie der Rhythmus und das Tempo, sowie bei noch längerer Betrachtung auch der Aufbau sowie Veränderungen in der Dynamik von einem oder der Abfolge von vielen akustischen Ereignissen (wie zum Beispiel bei einem Musikstück) darstellen. (Raffaseder 2010, S. 64) Betrachtet man ein Schallsignal nicht auf mikroskopischer Ebene, kann man den zeitlichen Verlauf von Schallsignalen grundsätzlich in vier Phasen unterteilen (siehe Abbildung 3). Die Attack-Time beschreibt die Dauer der Einschwingphase, nach der ein akustisches Ereignis die maximale Lautstärke erreicht.

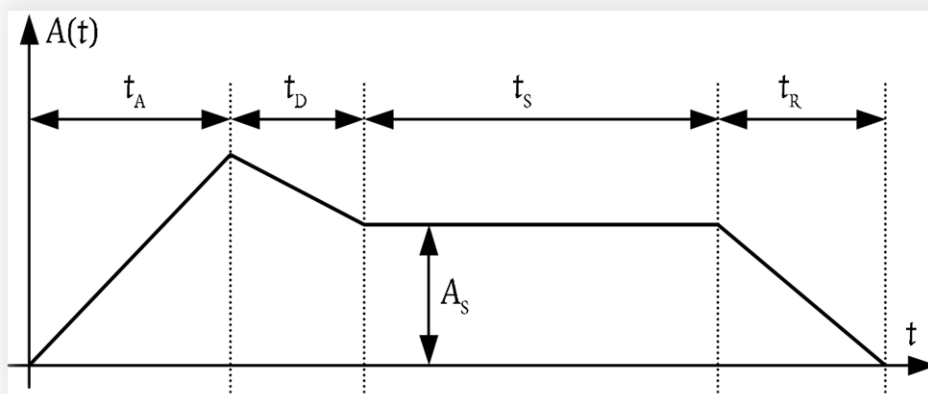


Abbildung 3: Darstellung eines Schallsignals als ADSR- Hüllkurve (Raffaseder 2010, S. 65)

$A(t)$... Amplitude

t ... Zeit

$t(A)$... Attack- Time

$t(D)$... Decay- Time

$t(S)$... Sustain- Time

$t(R)$... Release- Time

Die Attack-Time beschreibt die Dauer der Einschwingphase, nach der ein akustisches Ereignis die maximale Lautstärke erreicht. Ihr folgt die als Decay-Time beschriebene Dämpfungsphase, nach der sich die Pegel von Schallereignissen mit einer länger andauernden Anregung auf die Aushalte-Phase (Sustain-Time) einpendeln, wobei während dieser die Lautstärke annähernd konstant bleibt. Wird die Anregung der Schallquelle beendet, folgt die Auskling-Phase (=Release-time), die von den Schwingungseigenschaften der Schallquelle sowie von der Nachhallzeit des Raumes bestimmt wird. Dieses Modell bezeichnet man als die Hüllkurve eines akustischen Ereignisses, wobei oft die Abkürzung ADSR dafür verwendet wird. Zur besseren Vorstellung des Hüllkurven- Modells lassen sich die Klänge von Instrumenten heranziehen. Bei Schlag- oder Zupfinstrumenten bestehen die erzeugten Klänge nur aus einer kurzen Einschwingphase und der Dämpfungsphase, nach welcher der Klang wieder verstummt, wobei die Decay- Time von Schlaginstrumenten für gewöhnlich kürzer ist als bei Zupfinstrumenten. Bei Blas- oder Streichinstrumenten ist die Dauer der Anregung nicht nur ein kurzer Impuls, sondern länger anhaltend. Nach einer kurzen Einschwing- und Dämpfungsphase folgt die von der Dauer der Anregung (also zum Beispiel der Zeit, bis der Bogen einer Violine einmal über die Seite gestrichen wurde) bestimmte Aushalte- Phase und nach Beendigung der Anregung die Auskling- Phase. (Raffaseder 2010, S. 64)

2.2 Die auditive Wahrnehmung

2.2.1 Grundlagen der auditiven Wahrnehmung

Um die Tonspur in einem audiovisuellen Medium bewusst gestalten zu können, ist es notwendig, die Grundlagen der auditiven Wahrnehmung zu kennen, um die Wirkung von akustischen Ereignissen auf die RezipientInnen voraussehen zu können.

Thomas Görne unterteilt den Prozess der auditiven Wahrnehmung in zwei Stufen. Die physiologische Stufe der Wahrnehmung umfasst die durch den Aufbau des Ohres bestimmte Wandlung vom physikalischen Ereignis in einen Sinnesreiz, sowie die unterschiedlichen, durch physikalische Parameter ausgelösten Empfindungen wie die Lautstärke oder die Tonhöhe. Die kognitive Stufe der Wahrnehmung ist die sehr viel schwieriger zu erklärende Filterung und Strukturierung der Sinnesreize. Dabei werden die Schallereignisse Objekten zugeordnet, die mithilfe der wahrgenommenen Informationen kategorisiert werden und eine Bedeutung zugewiesen bekommen. (Görne 2014, S. 30)

Die als Schall bezeichneten Luftdruckschwankungen versetzen im Außenohr eine Membran (das Trommelfell) in Schwingung. Diese Schwingungen werden durch die Gehörknöchelchen im Mittelohr (Hammer, Amboß und Steigbügel) verstärkt und wiederum über das ovale Fenster auf die Gehörschnecke bzw. auf die sich darin befindende Flüssigkeit übertragen. Dadurch wird (je nach Frequenz) Druck auf verschiedene Haarzellen, die entlang der Schnecke angesiedelt sind, ausgeübt. Die mit den Haarzellen verbundenen Nervenzellen werden angeregt und elektrische Impulse an das Gehirn geleitet, wo eine Hörempfindung entsteht. (Hellbrück und Ellermeier 2004, S. 91–96). Die elektrischen Impulse werden zuerst durch verschiedene Verknüpfungen in den unteren Hirnschichten, unter anderem auch im Limbischen System, wo der Hormonhaushalt (zum Beispiel die Ausschüttung von Endorphinen oder Stresshormonen wie Adrenalin) geregelt, sowie primäre Körperfunktionen gesteuert werden (Raffaseder 2010, S. 108). Das Limbische System spielt auch eine wesentliche Rolle für die Empfindung von Emotionen (wiederum in Zusammenhang mit den primären Körperfunktionen, zum Beispiel Schweißproduktion bei Angstzuständen). (Stephan und Walter 2003, S. 53) Erst danach findet eine bewusste Verarbeitung der akustischen Sinnesreize in den für das Hörempfinden verantwortlichen Teilen der Großhirnrinde statt (Raffaseder 2010, S. 108). Somit werden Hörempfindungen zunächst weniger analytisch im Unterbewusstsein verarbeitet (Lensing 2006, S. 15), was ein essentieller Aspekt der auditiven Wahrnehmung in Bezug auf die akustische Gestaltung von audiovisuellen Medien ist, da so durch Musik oder den Einsatz von Soundeffekten gezielt Emotionen und Gefühlszustände der RezipientInnen beeinflusst werden können.

Die Wahrnehmung der akustischen Umwelt setzt sich aus einer Vielzahl solcher Sinnesreize, ausgelöst durch einzelne Schallereignisse, zusammen (Raffaseder 2010, S. 24). Dabei gibt es Parallelen zur Gestaltpsychologie, da auch die auditive Wahrnehmung dazu neigt, akustische Informationen nach bestimmten Mustern zu verarbeiten. (Raffaseder 2010, S. 24) Laut Michael Chion erfolgt die Wahrnehmung der Umwelt in kurzen Intervallen von zwei bis drei Sekunden, wobei die Hörempfindungen innerhalb eines solchen zeitlichen Abschnitts als eine zusammengehörige *Gestalt* wahrgenommen werden. Die vom Gehör aufgenommenen Empfindungen werden innerhalb dieser Zeit bereits verarbeitet und auf deren Informationsgehalt analysiert, wobei es sich dabei um einen kontinuierlichen, sich immer wiederholenden Prozess handelt. (Chion und Gorbman 1994, S. 9–10) Unvollständige Informationen werden dabei als störend empfunden oder werden unbewusst vervollständigt. Beispielsweise werden bei teilweise unverständlichen Sätzen die fehlenden Wörter unbewusst ergänzt und dadurch deren semantische Bedeutung verstanden, ebenso wie fehlende Töne in Melodien ergänzt werden. Gleichzeitig auftretende akustische Ereignisse werden als eine Klangquelle wahrgenommen (verstärkt durch den Maskierungseffekt, bei dem ein lauterer Signal durch Frequenzüberlagerungen ein leiseres verdeckt). Zeitlich kurz aufeinanderfolgende Ereignisse werden zu einer Information kombiniert, wenn diese eine logische Einheit ergeben. Besteht beispielsweise das Öffnen einer Tür aus einem Klicken und einem darauffolgenden Knarren, werden diese beiden Geräusche demnach nicht als einzelne, sondern als eine zusammengehörige akustische Information über eine sich öffnende Tür wahrgenommen. Aufeinanderfolgende Töne ergeben in einem musikalischen Kontext dadurch eine Melodie. (Sonnenschein 2001, S. 80–83)

Diese Einheiten bzw. Muster aus mehreren einzelnen Schallereignissen werden als akustische Streams bezeichnet. Es können bis zu fünf solcher Streams gleichzeitig verarbeitet werden. Welchen davon Aufmerksamkeit geschenkt wird, kann der Hörer in jedem Moment individuell bestimmen. (Raffaseder 2010, S. 24) Die Anzahl der Streams, denen dabei gleichzeitig Aufmerksamkeit geschenkt werden kann, ist bestimmt durch die Ähnlichkeit der Eigenschaften der akustischen Ereignisse. Sind diese sehr verschieden, ist es durchaus möglich, mehreren Ereignissen gleichzeitig

Aufmerksamkeit zu schenken. Werden jedoch beispielsweise mehrere menschliche Stimmen gleichzeitig wahrgenommen, sind die Eigenschaften der akustischen Informationen zu ähnlich, wodurch es schwierig ist, mehr als einem Gespräch sinnerfassend zu folgen. Die anderen Gespräche rücken in den Hintergrund, oder die Aufmerksamkeit springt zwischen den einzelnen Gesprächen hin- und her. (Sonnenschein 2001, S. 93) Bei akustischen Ereignissen, denen nicht bewusst Aufmerksamkeit geschenkt wird, können elementare Klangeigenschaften wahrgenommen werden (beispielsweise, ob es sich bei einem zweiten Schallereignis um menschliche Sprache oder um einen Sinuston handelt). Der semantische Inhalt des zweiten akustischen Ereignisses dagegen wird nicht wahrgenommen (zum Beispiel wenn außerhalb des Aufmerksamkeitsfokus bei einem Gespräch die Sprache gewechselt wird). Dieser Aspekt der Wahrnehmung wird 1953 von Collin Cherry durch den so genannten Cocktailparty-Effekt dargelegt (Görne 2017, S. 46), der „den Umstand beschreibt, dass Menschen auch in relativ lauter Umgebung in der Lage sind, die Wahrnehmung auf ein bestimmtes akustisches Ereignis zu richten“. Der Cocktailparty-Effekt zeigt unter anderem auch, dass akustische Informationen, die nicht bewusst wahrgenommen, dennoch zu einem gewissen Grad verarbeitet werden, da die akustische Umgebung ständig abgehört und ausgewertet wird und jederzeit einem anderen akustischen Ereignis Aufmerksamkeit geschenkt werden kann. (Raffaseder 2010, S. 19)

Beim Prozess der auditiven Wahrnehmung werden somit Luftdruckschwingungen, die durch ein Schallereignis mit physikalisch beschreibbaren Eigenschaften entstehen, durch die physiologischen Vorgänge im Ohr in Sinnesreize umgewandelt. Durch die Verarbeitung und Interpretation dieser in der kognitiven Stufe der auditiven Wahrnehmung entsteht eine subjektive Empfindung, die in der Psychoakustik als Hörereignis bezeichnet wird. (Görne 2017, S. 35) Hannes Raffaseder verwendet in Bezug auf die menschliche Wahrnehmung den Begriff Lautereignis. (Raffaseder 2010, S. 33)

2.2.2 Ebenen der auditiven Wahrnehmung

David Sonnenschein unterteilt die auditive Wahrnehmung grundsätzlich in *hearing* und *listening*, also in das passive Hören und das aktive Zuhören. (Sonnenschein 2001, S. 77) Beim passiven Hören wird die akustische Umwelt in kurzen Zeitfenstern analysiert, die Sinnesreize werden kategorisiert, mit Erfahrungswerten abgeglichen und daraus Handlungen abgeleitet. Beim aktiven Zuhören werden tiefere Zusammenhänge zwischen den Hörereignissen und eigenen Erfahrungen und Erwartungen hergestellt und anschließend interpretiert. (Raffaseder 2010, S. 30–31)

Darüber hinaus unterteilt Raffaseder die auditive Wahrnehmung in vier verschiedene Ebenen. Das „Hören zur Überwachung der Umwelt“ entspricht einer gänzlich unbewussten Wahrnehmung der akustischen Umwelt, wobei akustische Sinnesreize, ohne Bewertung von Ursachen bzw. ohne diese in einen zeitlichen oder räumlichen Kontext zu setzen, verarbeitet werden und (teilweise reflexartige) Reaktionen hervorrufen.

Beim „Hören zur Orientierung in der Umwelt“ werden die Reize in einen zeitlichen und räumlichen Zusammenhang gesetzt und falls möglich, mit der Ursache des akustischen Ereignisses verknüpft. Handlungen, die dadurch ausgelöst werden, sind nicht-reflexartige Reaktionen, die aufgrund der verarbeiteten akustischen Information ausgeführt werden. Auf der dritten Ebene, dem „Hören zur Kommunikation mit der Umwelt“, werden Hörereignisse in einem größeren räumlichen und zeitlichen Kontext interpretiert und aufgrund von Erfahrungen und Konventionen wird ihnen eine entsprechende Bedeutung beigemessen. Die hergestellten Zusammenhänge dienen als Grundlage für die Interaktion mit der Umwelt. Auch auf dieser Ebene muss die Wahrnehmung akustischer Informationen nicht bewusst stattfinden, sondern kann (zum Teils passiv) auf der Erkennung von Mustern basieren. Hingegen handelt es sich bei der vierten Ebene, dem „Zuhören zur Konstruktion neuer Wahrnehmungsmuster“ um einen aktiven, bewussten Vorgang der Wahrnehmung, wobei aus den verarbeiteten akustischen Reizen neue Zusammenhänge zur Umwelt erstellt, neue Wahrnehmungsmuster erlernt und dadurch neue Erfahrungen bzw. neues Wissen erlangt werden kann. (Raffaseder 2010, S. 30–31)

Für die Gestaltung der Tonspur in einem Spielfilm ist entscheidend, auf welcher Ebene Geräusche oder Klänge wahrgenommen werden, um entweder die Aufmerksamkeit zu erregen oder unterbewusst Informationen und Emotionen zu vermitteln.

Nach Michael Chion gibt es grundsätzlich drei verschiedene „Listening Modes“, also verschiedene Möglichkeiten, wie die Höreindrücke der akustischen Umwelt wahrgenommen werden können. **„Casual Listening“** ist die am häufigsten zum Einsatz kommende Art und Weise. Dabei dient das Hören dazu, Informationen über die Ursache bzw. den Auslöser eines Schallereignisses herauszufinden. Ist die Ursache des Klanges oder des Geräusches sichtbar, ergänzen die Informationen den visuellen Eindruck über die Schallquelle. Andernfalls dient der Höreindruck als alleinige Quelle über die Schallquelle sowie deren Eigenschaften. Die Informationen werden zur Kategorisierung des Klanges oder des Geräusches verwendet und im selben Zug mit Erinnerungen verknüpft bzw. Erfahrungen zugeordnet (zum Beispiel wird die Stimme einer bestimmten Person dieser zugeordnet oder das Bellen eines Hundes als solches erkannt). Beim **„Semantic Listening“** wird das Hörereignis nach bekannten Mustern oder Codes, wie zum Beispiel unterschiedlichen Sprachen, interpretiert und dem Hörereignis so eine Bedeutung beigemessen. Anzumerken ist hierbei, dass Causal Listening und Semantic Listening oft gleichzeitig stattfinden können. So wird zum Beispiel bei der Wahrnehmung der menschlichen Stimme erkannt *was* gesprochen und *wie*. **„Reduced Listening“** bezieht sich auf das bewusste Hören eines Klanges oder Geräusches an sich, unabhängig von der Ursache oder der Bedeutung des Höreindrucks. Dabei wird versucht, die Eigenschaften des Klanges zu erfassen, wobei diese zu beschreiben, sich schwierig gestaltet. Im Alltag wird dieser Modus nur sehr selten angewandt. Ein Anwendungsfall für das Reduced Listening ist zum Beispiel beim Üben eines Musikstückes, oder dem bewussten Hören bzw. dem ‚Verstehen‘ einer Melodie, wenn versucht wird, die Intervalle zwischen den einzelnen Tönen zu erkennen. (Chion und Gorbman 1994, 27-34)

David Sonnenschein ergänzt diese Modi noch um einen vierten, dem „Referential Listening Mode“, in dem die auditive Wahrnehmung nicht nur Rückschlüsse auf Herkunft und Entstehung der akustischen Information zulässt (Casual Listening), sondern die Einbeziehung des Gesamtkontextes zur emotionalen und dramaturgischen

Bedeutung des Klangs oder Geräusches beiträgt (Sonnenschein 2001, S.78). Besonders wichtig ist dieser Modus aber für die Bearbeitung der Elemente der Tonspur und für das Sounddesign, da es hierbei Klänge und Geräusche bewusst zu gestalten gilt und für die Bearbeitung eine isolierte Betrachtung und Beurteilung der Klangeigenschaften notwendig ist.

2.2.3 Wahrnehmbare Eigenschaften akustischer Ereignisse

Die in Kapitel 2.1.2 beschriebenen Parameter zur Definition der Eigenschaften können auch zu einer übersichtlicheren Beschreibung der Wahrnehmung von Schallereignissen herangezogen werden.

Die Herstellung von Zusammenhängen zwischen akustischen Ereignissen bzw. deren physikalischen Eigenschaften und den dadurch ausgelösten Hörempfindungen ist Gegenstand der klassischen Psychoakustik. (Raffaseder 2010, S. 43)

Lautstärke

In erster Linie hauptverantwortlich für die Wahrnehmung der Lautstärke sind folgende Parameter: Die Amplitude, also der Wert der maximalen Auslenkung einer Schallwelle, wird bestimmt durch die Stärke der Auslösung der Schwingung und steht in direktem Zusammenhang mit der Schallintensität und dem Schalldruckpegel. Die Schallintensität beschreibt dabei den „Betrag der gesamten Energie, die pro Sekunde durch eine Fläche von einem Quadratmeter, die senkrecht zur Ausbreitungsrichtung der Schallwelle steht, fließt“. (Hellbrück und Ellermeier 2004, S. 58)

Der Schalldruckpegel (angegeben in dB_{SPL} = Sound Pressure Level) stellt das logarithmische Verhältnis des Schalldrucks zu einem Bezugswert von $p_0 = 2 \cdot 10^{-5} \text{ Pa}$ (Raffaseder 2010, S. 67) (dieser Wert entspricht der menschlichen Hörschwelle) dar und definiert somit den Unterschied des gemessenen Schalldrucks zum leisesten wahrnehmbaren Schallereignis. Nach oben hin wird der vom Menschen hörbare Bereich auf einen

Schalldruck von 20 Pa angegeben, wobei dieser Wert als Schmerzgrenze gilt. Diese Werte sind selbstverständlich nur näherungsweise durch Experimente definiert worden und variieren von Person zu Person. (Raffaseder 2010, S. 88) Ein weiterer wichtiger Parameter in der Wahrnehmung der Lautstärke sind, neben dem durch die Stärke der Anregung bestimmten Schalldruckpegel, die Frequenzen, die in einem Schallereignis enthalten sind.

Schallpegel	Schallquelle
0	unhörbar (Hörschwelle)
10	Schneefall
20	leichter Wind, Ticken einer Taschenuhr
30	Flüstern
40	Kühlschrank
50	Ruhiger Bach oder Fluss, leises Gespräch
60	normales Gespräch
70	Lautes Gespräch, Rasenmäher in 7m Entfernung
80	Laute Radiomusik, starker Straßenverkehr
90	Presslufthammer in 1 m, schwerer Lkw in 5 m Entfernung
100	Diskotheek (innen)
110	Propellerflugzeug in 7 m Entfernung
120	Verkehrsflugzeug in 7 m Entfernung

Abbildung 4: Zur Veranschaulichung von im Alltag vorkommenden Schallereignissen und deren Schalldruckpegel in dBSPL können diese Beispiele dienen. (Umweltbundesamt)

Das menschliche Gehör ist für Frequenzen von 2000-4000 Hz am empfindlichsten, unterhalb von 500 Hz und oberhalb von 10000 Hz nimmt die Empfindlichkeit kontinuierlich ab. (Raffaseder 2010, S. 109) Verantwortlich dafür ist die anatomische Beschaffenheit des Gehörganges, der bei normalen Luftdruckverhältnissen eine Resonanzfrequenz im Bereich zwischen 3 und 3,5 KHz aufweist, was bewirkt, dass diese Frequenzen verstärkt werden. (Raffaseder 2010, S. 104) Schallereignisse mit verschiedenen Frequenzen müssen also unterschiedliche Schalldruckpegel aufweisen, um als gleich laut empfunden zu werden. Veranschaulicht wird die Frequenzabhängigkeit der empfundenen Lautstärke im Bezug zum Schalldruckpegel in den von Fletcher und Munson ermittelten *Kurven gleicher Lautstärke* (Friesecke 2007, S. 124) und den aktuell mit überholten Messwerten in der ISO 226 definierten „Equal Loudness Level Contours“.

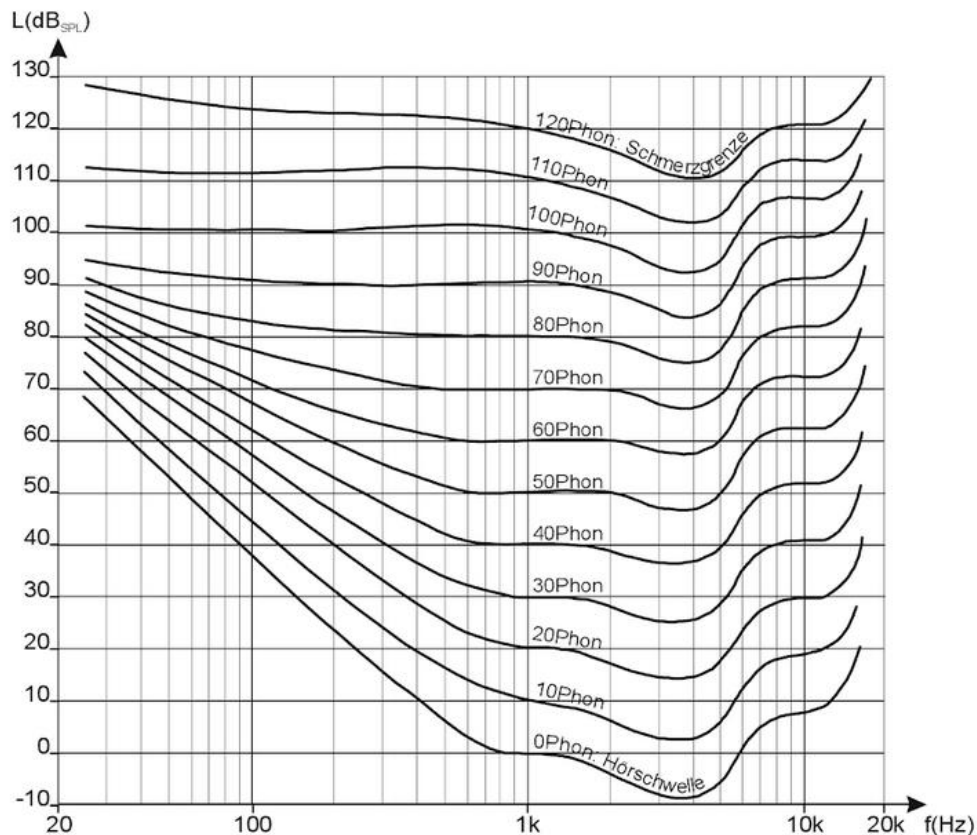


Abbildung 5: Kurven gleicher Lautstärke (auch Fletcher- Munson Curves genannt) (Friesecke 2007, S. 125)

Die Kurven zeigen ebenso, dass nicht nur die Frequenz die Lautstärke beeinflusst, sondern auch umgekehrt. Mit zunehmendem Schalldruckpegel wird die Kurve flacher, die Unterschiede in der Empfindlichkeit des Gehörs gegenüber verschiedenen Frequenzen werden daher geringer.

Da der Schalldruckpegel allein somit keine absoluten Werte für die wahrgenommene Lautstärke liefert, wird in der Psychoakustik die Einheit *Phon* verwendet, die den Lautstärkepegel im Bezug zur Lautstärke eines Sinustones mit einer Frequenz von einem Kilohertz angibt. Schallereignisse mit unterschiedlichen Frequenzen und der gleichen empfundenen Lautstärke haben so den gleichen Wert in *Phon*, auch wenn sich deren Schalldruckpegel unterscheiden, womit die Frequenzabhängigkeit des Gehörs in die Messung miteinbezogen wird. (Raffaseder 2010, S. 110) Darauf aufbauend wurde der Wert der empfundenen Lautheit in *Sone* eingeführt der einen direkteren Zusammenhang zwischen den gemessenen Werten und der empfundenen

Lautstärke liefert. Ein *Sone* entspricht einem Sinuston mit einer Frequenz von 1KHz und 40 dB_{SPL} und die Steigerung der Lautstärke verläuft (im Gegensatz zur logarithmischen Steigerung der Werte bei dB_{SPL} und Phon) linear zur Hörempfindung. (Sengpiel 2017b)

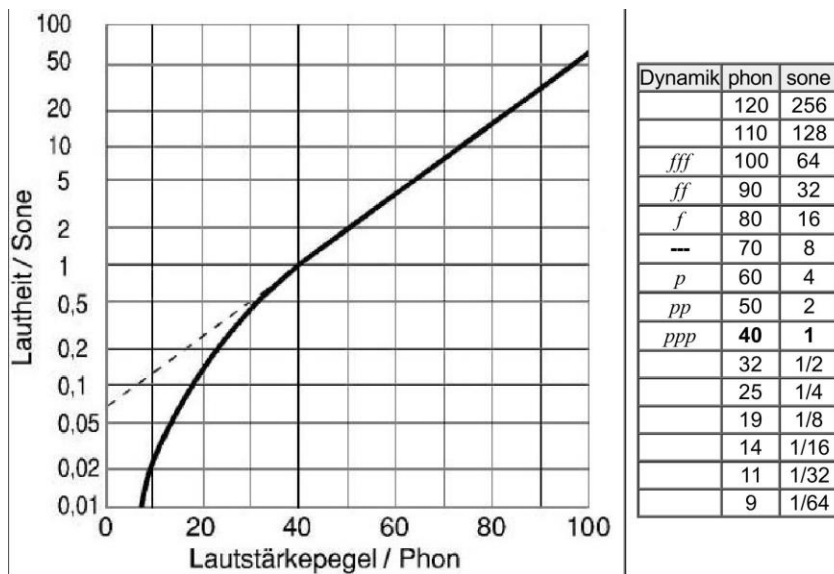


Abbildung 6: Zusammenhang zwischen den Einheiten Phon und Sone (Sengpiel 2017b)

Bei Phon und Sone handelt es sich aber stets nur um Vergleichswerte in Bezug auf den Schalldruckpegel. Die Empfindung der Lautstärke ist jedoch immer eine subjektive Empfindung. (Sengpiel 2017b) Hinzukommt, dass sich die Werte auf einen Sinuston, also ein Schallsignal mit nur einer einzigen Frequenz beziehen. Für die Lautstärkeempfindung ist ebenso essentiell, wie die Frequenzen eines Schallereignisses über den Frequenzbereich verteilt sind. Sind die Frequenzen über einen großen Bereich verteilt, leisten diese einen voneinander unabhängigen Beitrag zur Empfindung der Lautstärke. (Raffaseder 2010, S. 110) In der akustischen Gestaltung sowie bei der Abmischung der verschiedenen Elemente der Tonspur ist es also wichtig, einen möglichst großen Teil des zur Verfügung stehenden Frequenzbereiches zu nutzen, sofern eine Maximierung in der Lautstärke der Hörempfindung erzielt werden soll. Treten mehrere akustische Ereignisse im gleichen Frequenzbereich auf, kommt es zur Maskierung, das heißt Schallereignisse mit einem größeren Schalldruckpegel verdecken Schallereignisse deren Schalldruckpegel niedriger sind. Je ähnlicher sich die Frequenzspektren dabei sind, desto stärker ist der Effekt. So werden oft leisere

Geräusche die in einer ruhigen Umgebung gut hörbar sind bei größerem Umgebungslärm nicht wahrgenommen. Neben dem Maskierungseffekt im Frequenzbereich, trifft dieser auch in der Zeit auf. Dabei wird ein leiseres Schallereignis, das zeitlich sehr kurz auf ein lauterer folgt, von diesem verdeckt. (Raffaseder 2010, S. 111–112)

Neben der Frequenz und der Verteilung der Frequenzen im Spektrum spielt auch die Dauer eines Schallereignisses eine nicht unwesentliche Rolle in der Wahrnehmung der Lautstärke.

Länger anhaltende Töne oder Geräusche werden als lauter empfunden als ein kürzeres akustisches Ereignis mit demselben Schalldruckpegel. (Raffaseder 2010, S. 113) Dieser Effekt ist jedoch nur bei sehr kurzen, impulsartigen Schallereignissen wirksam. Die Schwelle liegt dabei bei 300 Millisekunden (also etwa ein Drittel einer Sekunde). Übersteigt die Länge zweier akustischer Ereignisse diese Schwelle, spielt die Dauer in der Lautstärkewahrnehmung keine Rolle mehr. (Görne 2017, S. 57) Das erklärt, warum bei einer technischen Messung der Effektivwert eines Signals größere Gemeinsamkeiten mit der menschlichen Wahrnehmung der Lautstärke aufweist als die Spitzenwerte der Amplituden. (Raffaseder 2010, S. 43) Der Effektivwert entspricht dem Mittelwert der quadrierten Amplitudenwerte eines Signals, aus dem die Wurzel gezogen wird. Würde der normale Mittelwert eines Signals verwendet werden, heben sich positive und negative Amplitudenwerte gegenseitig auf, wodurch das Ergebnis keine sinnvollen Werte liefern würde. (Raffaseder 2010, S. 66)

Bei einer konstanten Reizeinwirkung bzw. Stimulierung des Hörnervs über einen längeren Zeitraum sinkt die empfundene Lautstärke auch bei gleichbleibendem Schalldruckpegel. David Sonnenschein bezeichnet diesen Vorgang als Habituation (=de.: „Anpassung“), welcher bewirkt, dass viele Umgebungsgeräusche nach einer gewissen Zeit schwächer oder gar nicht mehr wahrgenommen werden. In diesem Zusammenhang spielt auch die sogenannte *ear fatigue* eine Rolle, die bei der Einwirkung von hohen Schalldruckpegeln über einen längeren Zeitraum auftritt. Diese beruht allerdings nicht auf psychischen Vorgängen der auditiven Wahrnehmung, sondern resultiert aus der physiologischen Ermüdung des Hörnervs. Akustische Ereignisse müssen dabei immer höhere Schalldruckpegel aufweisen, um noch

wahrgenommen zu werden. Entscheidend für die auditive Wahrnehmung und Erhaltung der Aufmerksamkeit ist daher der Kontrast in der Lautstärke. Das Prinzip der Adaption lässt sich neben der Lautstärke auch auf andere Parameter (zum Beispiel Tonhöhe oder Geschwindigkeit) anwenden. (Sonnenschein 2001, S. 94–95)

Tonhöhe

Durch die Beschaffenheit der Gehörschnecke werden Schwingungen mit unterschiedlicher Wellenlänge jeweils an bestimmten Stellen auf die Basilarmembran übertragen, wodurch je nach Wellenlänge unterschiedliche Nervenzellen angeregt werden. So ist es möglich, verschiedene Tonhöhen wahrzunehmen. (Raffaseder 2010, S. 105–106) Dieser Aspekt der Tonhöhenwahrnehmung wird als die Ortstheorie bezeichnet. Demnach bestimmt die sich aus der Wellenlänge der Schwingung abzuleitende Frequenz die wahrgenommene Tonhöhe. (Raffaseder 2010, S. 115)

Die Frequenzauflösung des Gehörs verläuft über den wahrnehmbaren Bereich von 20Hz bis 20KHz annähernd logarithmisch und gibt an, wie gering die Unterschiede in der Tonhöhe zweier Töne sein dürfen, damit diese wahrgenommen werden können. Bei einer langsamen Veränderung der Tonhöhe ist die Frequenzauflösung relativ ungenau und die kleinsten wahrnehmbaren Unterschiede bewegen sich bei 100Hz im Bereich von 3%, bei 350Hz im Bereich 1% und von 500 bis 3000 Hz bei 0,6%. Über 5000 Hz verschlechtert sich die Tonhöhen-Wahrnehmung zunehmend. (Friesecke 2007, S. 118) In musikalischen Tonleitern etwa sind daher für gewöhnlich keine Töne mit Grundfrequenzen größer als 4,2 KHz enthalten. (Raffaseder 2010, S. 115) Zur Unterscheidung der Tonhöhe muss der Unterschied in der Wellenlänge zweier Schwingungen einen bestimmten Wert übersteigen. Liegt die Frequenz der Schwingungen zu nahe beisammen, werden dieselben Bereiche an der Membran und somit dieselben Nervenzellen angeregt und die zwei Schwingungen werden als ein Schallsignal wahrgenommen. Dieser Aspekt der Wahrnehmung ist im Bassbereich am stärksten ausgeprägt. Mit zunehmender Tonhöhe verbessert sich die Frequenzauflösung, wobei im Bereich zwischen 2000 und 3000 Hz die Auflösung am feinsten funktioniert. Die Differenz, ab denen zwei Töne ähnlicher Frequenz als eigene Töne wahrnehmbar sind, wird als „kritische Frequenzbandbreite“ bezeichnet.

(Raffaseder 2010, S. 105–107) Die Fähigkeit zur Unterscheidung der Tonhöhe wird stark von der Geschwindigkeit der Veränderung beeinflusst. Im direkten Vergleich, also bei einer abrupten Änderung der Tonhöhe, ist die Frequenzauflösung des Gehörs um bis zu 30 Mal besser. (Frießecke 2007, S. 118) Zusätzlich hat auch die Lautstärke einen Einfluss auf die Tonhöhe, wobei dieser in der Regel vernachlässigbar ist. Beispielsweise wird ein Ton mit einer Frequenz 200 Hz mit steigender Lautstärke zunehmend tiefer, ein Ton mit 6000 Hz als noch höher empfunden. (Hellbrück und Ellermeier 2004, S. 117)

Wie bereits in Kapitel 2.2.2 beschrieben, kann die Tonhöhe jedoch nicht bei allen Schallereignissen eindeutig bestimmt werden. Voraussetzung für die Wahrnehmung der Tonhöhe ist, dass die durch eine Schallquelle ausgelösten Schwingungen periodisch sind, das heißt, dass sie in sich wiederholenden Mustern auftreten.

Natürlich vorkommende, komplexe Klänge setzen sich in der Regel aber nicht aus einer einzigen Frequenz, sondern aus mehreren Teiltönen mit unterschiedlicher Frequenz zusammen, die unterschiedliche Bereiche der Basilarmembran anregen. (Raffaseder 2010, S. 115) Verantwortlich für die Wahrnehmung der Tonhöhe ist dabei der Grundton bzw. die Grundfrequenz. Diese ist die tiefste im Schallereignis enthaltene Frequenz und lässt sich aus der Periodendauer des Signals ableiten, das heißt der Dauer eines sich wiederholenden Schwingungsmusters (siehe Abbildung 7). (Frießecke 2007, S. 118) Die Grundfrequenz entspricht dem Reziprokwert (=Kehrwert) der Periodendauer. (Raffaseder 2010, S. 58)

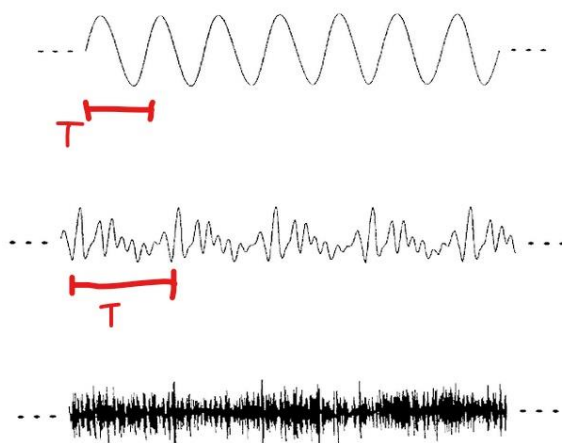


Abbildung 7: Zeitfunktionen von einem Sinuston, einem Klang und einem Geräusch

Die Periodendauer $T = 1/f$

(Hellbrück und Ellermeier 2004, S. 55)

Die restlichen Teiltöne werden als Obertöne bzw. Harmonische bezeichnet und sind mathematische Vielfache der Grundfrequenz (die erste Harmonische ist das Doppelte, die zweite Harmonische das Dreifache, die dritte Harmonische das Vierfache der Grundfrequenz usw.). (Hellbrück und Ellermeier 2004, S. 122).

Um alle Aspekte der Tonhöhenwahrnehmung zu erfassen, ist die Ortstheorie jedoch nicht ausreichend. (Hellbrück und Ellermeier 2004, 118,122) Die Grundfrequenz eines Schallereignisses wird auch dann wahrgenommen, wenn diese gar nicht im Schallsignal enthalten ist. Daraus lässt sich schließen, dass das Gehirn von den Nervenzellen nicht nur Informationen über den Ort der Reizauslösung, sondern auch über die Wellenform in Abhängigkeit von der Zeit erhält und dass daraus die Periodendauer bzw. die Grundfrequenz des Schallsignales interpretiert werden kann. Hannes Raffaseder bezeichnet diese Aussage als die Periodizitätstheorie. (Hellbrück und Ellermeier 2004, S. 122) Ein anderer Name für diesen Vorgang ist zeitliche Kodierung, wobei die genaue Erklärung für diesen Aspekt der tonalen Wahrnehmung noch Schwierigkeiten bereitet. Es wird jedoch angenommen, dass die Rate, mit der die Nervenzellen des Hörnervs bei Anregung durch die Schwingung der Basilarmembran elektrische Impulse abgeben, als Grundlage dafür dient. Sicher ist jedoch, dass die zeitliche Kodierung oberhalb von 5000 Hz keine Rolle mehr für die Tonhöhenempfindung spielt, da die Rate der Nervenimpulse zu schnell wird, um entsprechend interpretiert werden zu können. Unterhalb von 1000 Hz dürfte dagegen die Ortstheorie keinen Einfluss auf die Wahrnehmung der Tonhöhe mehr haben, da die Auslenkung der Basilarmembran zu breit ist. Im Bereich von 1000 Hz bis 5000 Hz besitzen beide Vorgänge hinsichtlich der Wahrnehmung der Tonhöhe von Bedeutung. (Hellbrück und Ellermeier 2004, S. 119–120)

In einem musikalischen Kontext ist die genaue Frequenz in der Wahrnehmung der Tonhöhe nicht wesentlich. Entscheidend sind die Unterschiede und Verhältnisse zwischen den verschiedenen Tönen. Obwohl eine wesentlich genauere Unterscheidung möglich wäre, wird der Frequenzbereich in fast allen Kulturen in relativ wenige ganzzahlige Stufen unterteilt. (Raffaseder 2010, S. 119) Die Frequenz 440Hz des Kammertones A, die in den meisten Fällen als Referenz bzw. als Ausgangspunkt für die Stimmung von Instrumenten dient, ist willkürlich gewählt.

In der westlichen Musik wird eine Oktave in zwölf Halbtonschritte unterteilt. Die Abstände zu anderen Tönen auf dieser Skala werden als Intervalle bezeichnet und können als mathematische Frequenzverhältnisse dargestellt werden. Eine Oktave entspricht dem Frequenzverhältnis zweier Töne von 1:2 und ist somit die Verdopplung der Grundfrequenz im Vergleich zum Ausgangston. (Raffaseder 2010, S. 116) Der für den Menschen hörbare Bereich (20Hz - 20 KHz) kann näherungsweise in zehn dieser Oktaven unterteilt werden. (Leonhard 1999, S. 1253) Die Darstellung als Verhältnis ist insofern zielführend, da die logarithmische Auflösung der Frequenzen durch das Gehör berücksichtigt wird und die Intervalle unabhängig von der Tonlage angegeben werden können. Der Unterschied zwischen zwei Tönen mit 500 und 1000 Hz wird als gleich groß empfunden wie der Unterschied zwischen 1000 und 2000 Hz.

In der heute gebräuchlichen gleichstufigen Stimmung (oft auch die *wohltemperierte Stimmung* genannt) liegen im Vergleich zur *reinen Stimmung* alle zwölf Halbtöne um den gleichen Faktor auseinander. Dies hat den Vorteil, dass die Intervalle für alle Tonarten gleich sind und es somit keine reiner oder unreiner klingenden Tonarten gibt. Es handelt sich jedoch um einen Kompromiss, da so alle Intervalle außer der Oktave minimal verstimmt sind (zum Beispiel ist das Frequenzverhältnis der Quinte in der reinen Stimmung mit 2:3 angegeben, in der wohltemperierten aber mit 1,49830707). Die Hörgewohnheit in westlich geprägten Kulturkreisen hat sich aber längst daran angepasst. (Raffaseder 2010, S. 121)

Je nach Kultur und individueller Prägung werden verschiedene Intervalle als mehr oder weniger dissonant bzw. konsonant empfunden. Als Dissonanz wird der „Zusammenklang von Tönen, der als nicht harmonisch, nicht als Wohlklang empfunden wird [und nach der überlieferten Harmonielehre eine Auflösung fordert]“ bezeichnet. (Duden | Dis-so-nanz | Rechtschreibung, Bedeutung, Definition, Synonyme, Herkunft 2018) Die Unterscheidung zwischen konsonanten und dissonanten Intervallen ist dabei fließend. Für gewöhnlich werden die Oktave mit dem Verhältnis 1:2, die Quinte mit dem Verhältnis 2:3 und die Quarte mit dem Verhältnis 3:4 als die konsonantesten Intervalle empfunden, gefolgt von der kleinen und der großen Terz mit dem Verhältnis 5:6 und 4:5, sowie der kleinen und großen Sext mit dem Verhältnis 5:8 und 3:5. Nach der

Proportionstheorie von Pythagoras werden zwei Töne umso konsonanter empfunden, je einfacher das Frequenzverhältnis zwischen den beiden ist.

Die Klangverwandtschaftstheorie nach Helmholtz besagt, dass zwei Töne dann konsonant sind, wenn beide Klänge gemeinsame Obertöne aufweisen. Interessant ist, dass die Unterscheidung zweier Töne im Abstand einer Oktave schwerer fällt, als bei weniger konsonanten Intervallen. Neuere Erkenntnisse zeigen, dass ein Intervall dann als dissonant empfunden wird, wenn sich die Frequenzen innerhalb einer „kritischen Frequenzbandbreite“ befinden, und nah beisammen liegende Bereiche der Basilarmembran anregen. Ebenso spielen nicht nur die Grundtöne, sondern auch die enthaltenen Teiltöne eine Rolle. Komplexe Klänge mit vielen unterschiedlichen Teiltönen werden umso konsonanter empfunden, je weniger Teiltöne sich in einem kritischen Frequenzbereich befinden. Im Bassbereich klingen Intervalle schneller dissonant als im Mitten- und Hochmittenbereich zwischen 1000 und 3000 Hz, wodurch zum Beispiel auch die sonst eher als konsonant empfundene Terz als dissonant wahrgenommen werden kann. (Raffaseder 2010, S. 116–119)

Intervall	Reine Stimmung	Temperierte Stimmung Formel	Temperierte Stimmung Multiplikator	Temperierte Stimmung Prozent
1Cent		$\cdot \sqrt[1200]{2}$	1,00057779	0,06%
1Halbton		$\cdot \sqrt[12]{2}$	1,05946309	5,95%
kleine Sekunde	$\cdot 16 \div 15$	$\cdot \sqrt[12]{2}$	1,05946309	5,95%
große Sekunde	$\cdot 9 \div 8$	$\cdot \sqrt[6]{2}$	1,12246204	12,25%
kleine Terz	$\cdot 6 \div 5$	$\cdot \sqrt[4]{2}$	1,18920712	18,92%
große Terz	$\cdot 5 \div 4$	$\cdot \sqrt[3]{2}$	1,25992105	25,99%
Quarte	$\cdot 4 \div 3$	$\cdot \sqrt[12]{32}$	1,33483985	33,48%
verminderte Quinte	$\cdot 7 \div 5$	$\cdot \sqrt{2}$	1,41421356	41,42%
Quinte	$\cdot 3 \div 2$	$\cdot \sqrt[12]{128}$	1,49830707	49,83%
kleine Sexte	$\cdot 8 \div 5$	$\cdot \sqrt[3]{4}$	1,58740105	58,74%
große Sexte	$\cdot 5 \div 3$	$\cdot \sqrt[4]{8}$	1,68179283	68,18%
kleine Septime	$\cdot 16 \div 9$	$\cdot \sqrt[6]{32}$	1,78179744	78,18%
große Septime	$\cdot 15 \div 8$	$\cdot \sqrt[12]{2048}$	1,88774863	88,77%
Oktave	$\cdot 2$	$\cdot 2$	2,00000000	100,00%

Abbildung 8: Musikalische Intervalle und deren Verhältnisse (Frieesecke 2007, S. 122)

Klangfarbe

Laut Hannes Raffaseder wird die Klangfarbe als „jene primäre Eigenschaft eines andauernden Schallereignisses verstanden, die es der menschlichen Wahrnehmung ermöglicht, zwei akustische Ereignisse zu unterscheiden, die die gleiche Lautstärke- und Tonhöhenempfindung auslösen“ (Raffaseder 2010, S. 121). Bestimmt wird die Klangfarbe im Wesentlichen von der Beschaffenheit (Material und Größe) der Schallquelle und der Art der Anregung dieser. (Raffaseder 2010, S. 42)

Hinsichtlich einer physikalischen bzw. mathematischen Beschreibung der Klangfarbe können die Unterschiede auf zwei Eigenschaften von Schallereignissen zurückgeführt werden.

Die Zusammensetzung bzw. die Form des Signales im Frequenzbereich bestimmt im Wesentlichen die Unterscheidung von Klängen und Geräuschen und weiterführend ob diesen Attribute wie zum Beispiel *hell* oder *dunkel* zugeordnet werden. (Raffaseder 2010, S. 121) Dargestellt wird diese Zusammensetzung im Frequenzspektrum. Im Gegensatz zur zeitlichen Darstellung von Schallsignalen werden im Spektrum die Amplituden der im Schallsignal enthaltenen Frequenzen in Abhängigkeit von der Frequenz angegeben. (Raffaseder 2010, S. 60) Wie bereits in Bezug auf die Wahrnehmung der Tonhöhe im vorigen Kapitel erwähnt, sind Klänge die Resultate von periodischen Schwingungen. (Friessecke 2007, S. 119) Nach Jean Baptiste Fourier besteht jeder Klang aus mehreren harmonischen Schwingungen bzw. Teiltönen, wobei die Frequenzen dieser Teiltöne ganzzahlige Vielfache der Grundfrequenz sind. Jeder harmonische Teilton kann als einzelne Linie im Spektrum dargestellt werden und besitzt so eine einzige, eindeutig bestimmbare Frequenz. (Raffaseder 2010, S. 59–60)

Von einem harmonischen Ton wird dann gesprochen, wenn das Signal durch die Winkelfunktionen Sinus und Kosinus beschrieben werden kann. Ein harmonischer Ton ist immer periodisch. In der Natur kommen harmonische Töne selten vor, [...] allerdings können sie [...] als Grundbausteine jeder beliebigen anderen Signalform angesehen werden und sind dabei von fundamentaler Bedeutung. (Raffaseder 2010, S. 59)

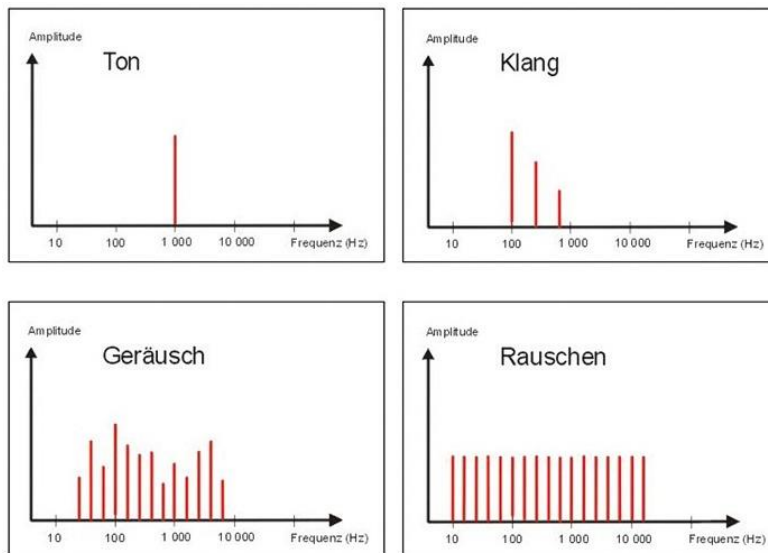


Abbildung 9:
Darstellung eines
harmonischen Tones,
eines Klanges, eines
Geräusches sowie
von weißen
Rauschen im
Frequenzbereich
(Lärmfibel)

Für die Empfindung der Klangfarbe eines komplexen Klanges ist entscheidend, aus welchen Teiltönen sich dieser zusammensetzt sowie das Lautstärkeverhältnis der Teiltöne. (Görne 2017, S. 65) Als Beispiel dienen hierfür die Vokale „o“ und „i“. Beim dem als „dunkler empfundenen „o“ sind ausgeprägte Maxima im Bereich der zweiten und fünften harmonischen des Grundtones zu finden, beim „i“ ist es lediglich die 16. Harmonische, die einen lauten Pegel im Verhältnis zum Grundton aufweist. (Friesecke 2007, S. 119) Bei Geräuschen, also aperiodischen Signalen, ist eine Zusammensetzung aus mehreren harmonischen Schwingungen nicht möglich. (Raffaseder 2010, S. 61–62)

Der zeitliche Verlauf eines geräuschhaften Signals verläuft rein zufällig, wodurch keine Periodendauer und somit keine Grundfrequenz festgestellt werden. Anstatt einiger weniger ausgeprägter Maxima verläuft das Spektrum annähernd kontinuierlich, das heißt Geräusche verfügen über sehr viele Informationen, die über einen großen Frequenzbereich verteilt sind. Wie die harmonische Schwingung für Klänge stellt das *weiße Rauschen* den Extremfall eines Geräusches dar. Anstatt einer einzigen Frequenz (harmonische Schwingung) sind im weißen Rauschen alle Frequenzen gleich stark vertreten, das heißt das Spektrum verläuft frequenzkontinuierlich. (Görne 2017, S. 55–57) Dennoch kann das Spektrum Informationen über die Klangfarbe eines Geräusches liefern, da je nach Gewichtung der Frequenzanteile dennoch zwischen hoch und tief unterschieden werden kann. Je schmalbandiger dabei das Geräusch ist (sich somit die spektrale Zusammensetzung nicht über einen weiten Frequenzbereich erstreckt,

sondern sich auf einen kleineren Bereich konzentriert), desto eher kann es eine Tonhöhenempfindung auslösen. (Raffaseder 2010, S. 61–62)

Die Grenze zwischen Tönen und Geräuschen kann nicht klar definiert werden und ist somit fließend. (Görne 2017, S. 57) Zur Veranschaulichung dafür eignen sich akustische Ereignisse wie Tierlaute oder auch manche Schlaginstrumente, die gleichzeitig Eigenschaften von Klängen und Geräuschen aufweisen, demnach deren Spektrum zwar deutliche Hervorhebungen bestimmter Frequenzen oder Frequenzbereiche besitzt, die aber keine ganzzahligen Vielfachen der Grundfrequenz darstellen. (Raffaseder 2010, S. 122) Für die Ausprägung verschiedener Maxima, das heißt der für die Zusammensetzung des Frequenzspektrums verantwortlichen Hervorhebung bestimmter Teiltöne, sind die Resonanzkörper der Schallquelle verantwortlich. Aufgrund deren Größe und Schwingungseigenschaften werden unabhängig von der Tonhöhe bzw. der Grundfrequenz manche Frequenzbereiche mehr verstärkt als andere. Diese Frequenzbereiche werden als Formanten bezeichnet und sind für eine bestimmte Schallquelle charakteristisch. (Raffaseder 2010, S. 122) Beispielsweise klingt so die Saite einer Violine bei gleicher Tonhöhe anders als die Saite einer Gitarre. Die Hervorhebung von unterschiedlichen Obertönen bei der Aussprache von den Konsonanten „o“ und „i“ bei gleichbleibender Tonhöhe ist durch die Veränderung der Form der Mundhöhle (also dem Resonanzkörper der menschlichen Stimme) bedingt. Daraus folgt, dass zwischen der Beschaffenheit der Schallquelle und den damit verbundenen charakteristischen Formanten ein von der Tonhöhe und Lautstärke unabhängiger Zusammenhang mit der Klangfarbe eines akustischen Ereignisses besteht. (Raffaseder 2010, S. 122)

Das Frequenzspektrum allein reicht zur Beschreibung der Klangfarbe eines Schallsignales jedoch nicht aus. (Raffaseder 2010, S. 43) Die zweite wesentliche Eigenschaft ist der, auch als Transient bezeichnete, Einschwingvorgang eines akustischen Ereignisses. Wird dessen Einschwingphase zu stark verfremdet, kann die Identifikation bestimmter Klänge oder Geräusche Probleme bereiten oder auch gar unmöglich werden, und die gehörten Schallereignisse werden als unnatürlich empfunden. (Friesecke 2007, S. 119) Beispielsweise kann eine rückwärts abgespielte Aufnahme eines Instrumentes auch von erfahrenen Musikern nicht mehr eindeutig

identifiziert werden, obwohl sich das Spektrum nicht verändert, da die Einschwingphase des Klanges variiert. (Raffaseder 2010, S. 122)

Mehr noch als bei der Lautstärke und der Tonhöhe kann die Beschreibung der Klangfarbe anhand technischer Messwerte nur näherungsweise einen Zusammenhang zur Wahrnehmung der Klangfarbe herstellen (Görne 2014, S. 67), da die Klangfarbe keine zweidimensionale Skala wie die Lautstärke (die Empfindungen hinsichtlich der Lautstärke von akustischen Ereignissen lassen sich als Werte zwischen *laut* und *leise* wiedergeben) oder die Tonhöhe (Empfindungen können auf einer Skala zwischen *hoch* und *tief* dargestellt werden) ist. (Raffaseder 2010, S. 123)

Verschiedene Attribute wie die **Schwankungsstärke**, welche die „Variation von Amplitude und Frequenz im Zeitbereich der Transienten beschreibt (Modulationseffekte wie das Tremolo oder das Vibrato können auf die Schwankungsstärke zurückgeführt werden) und welcher Parameter wie *pulsierend* oder *brodelnd* zugeordnet werden können, werden von Hannes Raffaseder zur Beschreibung der Klangfarbe herangezogen. Die **Rauigkeit** wird durch „die Anzahl jener ausgeprägter Maxima im Spektrum [...] die in ein gemeinsames kritisches Frequenzband fallen“ (Raffaseder 2010, S. 123) definiert. Das **Volumen** eines Schallsignals, welches zum Beispiel „ermöglicht, zwischen zarten oder kräftigen Stimmen oder dem vollen Klang einer Pauke und dem dünnen Klang eines Triangels zu unterscheiden, ohne dabei Lautstärke oder Tonhöhe berücksichtigen zu müssen“, wird durch die Anzahl der spektralen Maxima bestimmt, was zur Beschreibung von Klängen hilfreich ist. Hinsichtlich der Beschreibung der Klangfarbe von Geräuschen wird statt des Volumens meist von der **Dichte** gesprochen, die näherungsweise der Bandbreite des Signales im Frequenzspektrum entspricht, wobei der Dichte die Parameter *breit* oder *schmal* zugeordnet werden können. Als Beispiel für ein schmales Geräusch könnte das Zirpen einer Grille dienen, für ein breites Geräusch ein eingeschalteter Fön oder das Rauschen eines Wasserfalles. Zur Beschreibung der **Schärfe** eines Klanges dient der Schwerpunkt der Frequenzen im Spektrum bzw. die „Balance zwischen hohen und tiefen Frequenzen“ und ist ähnlich der Unterscheidung zwischen *hellen* und *dunklen* Klängen oder Geräuschen. (Raffaseder 2010, S. 123)

Räumlichkeit und Richtung

Die Richtungswahrnehmung basiert, neben einer sehr eingeschränkten Fähigkeit zur Lokalisation von Schallereignissen mit nur einem einzelnen Ohr (monaurales Hören), beim binauralen Hören (=mit beiden Ohren) auf Differenzen im Hörereignis zwischen linkem und rechtem Ohr. (Lensing 2006, S. 149) Verantwortlich dafür sind Laufzeitunterschiede (liegt etwa die Schallquelle zum Beispiel auf der rechten Seite, kommen Schallwellen am linken Ohr durch die größere Distanz leicht verzögert an), sowie durch die Schallbeugung verursachte Unterschiede in der Zusammensetzung des Frequenzspektrums (hohe Frequenzen mit einer Wellenlänge λ kleiner dem Durchmesser des menschlichen Kopfes werden durch den Kopf blockiert, Schallwellen mit λ größer als dieser Durchmesser können dagegen von beiden Ohren gehört werden). (Sonnenschein 2001, S. 85–86) Daher ist die Richtungswahrnehmung von hochfrequenten akustischen Ereignissen (etwa ab 2-3 KHz) durch Intensitätsunterschiede bedingt, bei tiefen Frequenzen hingegen durch Zeitdifferenzen. (Lensing 2006, S. 149) Die Ortung von Schallquellen oberhalb und unterhalb sowie vor und hinter dem Rezipienten wird durch die spezielle Form der Ohrmuscheln ermöglicht. Aufgrund der vergleichbar geringen Dimensionen der Wölbungen der Ohrmuschel im Verhältnis zu den Wellenlängen von Frequenzen im hörbaren Bereich und der Frequenzabhängigkeit der Schallbeugung ist die Richtungswahrnehmung bei hohen Frequenzen besser ausgeprägt. (Sonnenschein 2001, S. 85–86) Für die menschliche Wahrnehmung gilt in Bezug auf die Empfindung von Richtung und Distanz im Raum das „Prinzip der ersten Wellenfront“ (= engl. precedence effect). So werden vom Gehör für die Richtungswahrnehmung nur der Direktschall, das heißt die ersten am Ohr eintreffenden Schallwellen, und die frühesten eintreffenden Reflexionen innerhalb von 40 Millisekunden nach dem Direktschall für die Richtungswahrnehmung herangezogen. Alle nachfolgenden Reflexionen sind für die Wahrnehmung von Rauminformationen und Nachhall verantwortlich. (Görne 2017, S. 39)

Für die Wahrnehmung von **Räumlichkeit** durch den Gehörsinn sind zum einen der Frequenzbereich und die Intensität eines akustischen Ereignisses verantwortlich. Aufgrund der geringen Energie hochfrequenter Schallwellen nimmt die Intensität

dieser im Vergleich zu tiefen Frequenzen schneller ab, je weiter eine Schallquelle entfernt ist. (Sonnenschein 2001, S. 83–84) Dies wird zusätzlich dadurch verstärkt, dass die Absorption bzw. Transmission von Schallereignissen frequenzabhängig ist (hohe Frequenzen werden beim Auftreffen auf Objekte besser absorbiert, wohingegen tieffrequente Anteile auch durch Wände oder geschlossene Türen dringen) und ebenso, dass das Phänomen der Schallbeugung (=treffen Schallwellen auf ein Objekt, werden Frequenzen mit einer Wellenlänge größer als das Hindernis um dieses gebeugt und sind dadurch auf der anderen Seite des Objektes hörbar (Raffaseder 2010, S. 94)) bei hohen Frequenzen zunehmend seltener auftritt. Zum anderen beeinflussen sowohl der zeitliche Abstand von den Reflexionen an Oberflächen und Strukturen des Raumes als auch das Lautstärkeverhältnis der Reflexionen zum Direktsignal die Wahrnehmung der akustischen Umgebung, wodurch von der menschlichen Wahrnehmung Rückschlüsse auf die Entfernung bzw. die Position der Schallquelle im Raum gezogen werden können. (Raffaseder 2010, S. 93–94) Stimmen diese physikalischen Gegebenheiten bei zwei oder mehreren akustischen Informationen überein und ändern sich simultan, werden diese Informationen als eine Bewegung der Schallquelle im Raum interpretiert. (Sonnenschein 2001, S. 93–94)

Die Wahrnehmung von **Bewegung**, sowohl objektiv (die Änderung der Position eines anderen Objektes) als auch subjektiv (die Änderung der eigenen Position relativ zu einem Schallereignis), erfolgt somit durch die Veränderung der Werte jener Parameter, die auch für die Wahrnehmung von Richtung und Räumlichkeit verantwortlich sind. (Sonnenschein 2001, S. 87–88) Zusätzlich kann auch die Tonhöhe Informationen über die Bewegung einer Schallquelle liefern. Bekannt ist dieses Phänomen als der sogenannte Doppler-Effekt, wobei sich die Tonhöhe bei Bewegung der Schallquelle mit hoher Geschwindigkeit in Abhängigkeit von der Distanz zwischen Schallquelle und dem Empfänger verändert. (Lensing 2006, S. 21)

2.2.4 Wahrnehmung von Sprache

Sound in film is voco- and verbocentric, above all, because human beings in their habitual behavior are as well. When in any given sound environment you hear voices, those voices capture and focus your attention before any other sound (wind blowing, music, traffic). (Chion und Gorbman 1994, S. 6)

Dies ist zum einen dadurch bedingt, dass die menschliche Wahrnehmung im Allgemeinen auf die menschliche Stimme fokussiert ist und zum anderen dadurch, dass die Erzählung der Handlung in erster Linie auf Sprache und Dialogen aufbaut und die Konzepte meist in Worten in Drehbüchern festgehalten werden. (Lensing 2006, S. 48)

Grundsätzlich kann man die vom menschlichen Sprachorgan produzierten Laute in stimmhafte und stimmlose unterteilen. Stimmhafte Laute entstehen durch die von einem Luftstrom in Schwingung versetzten Stimmbänder, wodurch „ein quasiperiodisches, sehr obertonreiches Signal, dem eine eindeutige Grundfrequenz zugeordnet werden kann“ entsteht. (Raffaseder 2010, S. 84) Die Grundfrequenz der menschlichen Stimme ist bei Männern bei etwa 120 Hz angesiedelt, bei Frauen bei etwa 240 Hz und bei Kindern bei 400 Hz, wobei diese Werte Mittelwerte darstellen, die je nach Länge der Stimmlippen variieren. (Altmann und Ziegenhain 2007, S. 25) Stimmlose Laute sind geräuschartige Laute, denen keine Grundfrequenz zugeordnet werden kann. Weiters können die von der Stimme produzierten Laute in sogenannte Vokale und Konsonanten unterteilt werden. (Raffaseder 2010, S. 84) Ein Vokal bedeutet in der Sprachwissenschaft ein „deutlich erklingender Laut, bei dessen Artikulation die Atemluft verhältnismäßig ungehindert ausströmt; Selbstlaut“. (Duden | Vo-kal | Rechtschreibung, Bedeutung, Definition, Herkunft 2018) Bei geöffnetem Mund, das heißt wenn der Luftstrom ungehindert ausströmen kann, entstehen die stimmhaften Vokale a, e, i, o und u. Trotz annähernd gleichbleibender Grundfrequenz entsteht je nach Form des Mund- Nase- Rachen- Raumes sowie der Haltung der Zunge ein Resonanzraum, der die für den jeweiligen Vokal charakteristischen Obertöne hervorhebt, welche als Formanten bezeichnet werden. Wesentlich für die Bildung von Vokalen sind die ersten beiden Obertöne bzw. Formanten. Eine übersichtliche Darstellung des Zusammenhanges zwischen der Haltung der Zunge im Resonanzraum

und den entstehenden Vokalen stellt das sogenannte Vokalviereck dar (siehe Abbildung 10). (Raffaseder 2010, S. 84)

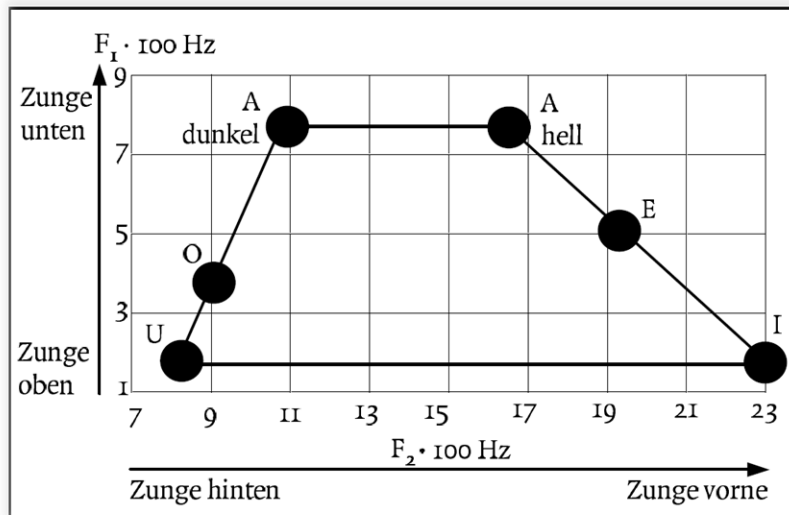


Abbildung 10:
Vokalviereck
(Raffaseder
2010, S. 84)

Ein entscheidender Unterschied in der Produktion von Vokalen durch das menschliche Sprachorgan, im Vergleich zur Produktion von Klängen durch Musikinstrumente besteht darin, dass die verschiedenen Laute der Sprache durch Veränderung des Resonanzraumes und den damit verbunden Formanten gebildet werden, wobei der Grundton gleich bleibt. Bei Klängen von Musikinstrumenten ändert sich hingegen der Grundton, während sich der Resonanzraum nicht verändert. Kann der Luftstrom nicht ungehindert ausströmen, wird dieser zum Beispiel von der Zunge blockiert, entstehen die turbulenten Luftströmungen, die für die geräuschhaften Anteile der Konsonanten verantwortlich sind. Dennoch können Konsonanten, je nachdem ob die Stimmbänder eingesetzt werden, auch stimmhaft sein, das heißt einen Grundton besitzen. (Raffaseder 2010, S. 83–84)

Untersucht wird die Lautproduktion der Sprachlaute in der artikulatorischen Phonetik, während sich die auditive Phonetik mit der Wahrnehmung der Sprachlaute beschäftigt. (Karl Heinz Wagner 2018)

In Bezug auf die Wahrnehmung ist der Hörsinn darauf programmiert, diese Laute zu Einheiten zusammenzufassen, wobei die im vorherigen Kapitel besprochene

Ähnlichkeit zur Gestalt-Wahrnehmung zum Tragen kommt. Es ist zum Beispiel fast unmöglich, die eigene Muttersprache zu hören, ohne diese Kombination von Lauten als zusammenhängende Silben und Wörter wahrzunehmen. (Sonnenschein 2001, S. 137) Dabei werden die vom Sprachorgan produzierten Vokale und Konsonanten zu Lauteinheiten (sog. Phonemen = „kleinste bedeutungsunterscheidende sprachliche Einheit“ (Duden | Pho-nem, Fo-nem | Rechtschreibung, Bedeutung, Definition 2018)) zusammengefasst, diese wiederum zu Silben und in weiterer Folge zu Phrasen und Sätzen kombiniert, denen durch Konvention eine semantische Bedeutung beigemessen wird. Besonders ist dabei, dass diese Laute vom menschlichen Gehirn schneller verarbeitet werden und andere Regionen des Gehirnes anregen als andere Geräusche oder Töne. (Sonnenschein 2001, S. 134)

Die Strukturierung der einzelnen Laute einer Sprache erfolgt dabei durch die kontinuierliche Veränderung von bestimmten Parametern, was unter dem Begriff *Prosodie* zusammengefasst wird. Bei den Parametern handelt es sich um die Intonation (=„der Verlauf der Grundfrequenz in den stimmhaften Abschnitten des Sprachsignals“) (Tillmann, Schiel 1995), die Pausen zwischen den Lauten sowie die relative Lautstärke der einzelnen Laute zueinander. Die Veränderung dieser kann zu einem großen Teil bewusst gesteuert werden und so Informationen über die Satzintention (ob es sich somit um einen Aussage- oder Befehlssatz oder um eine Frage handelt) vermitteln, sowie Gefühle oder die körperliche Verfassung des Sprechers oder der Sprecherin ausdrücken (=Prosodie A). Fehlen diese Veränderungen in der Sprache, klingt diese unnatürlich (wie zum Beispiel bei Computerstimmen). Eine Rolle in der Prosodie spielt auch die Silbenfolge bzw. der Silbenrhythmus, welcher „die rhythmische Abfolge von stimmhaften und stimmlosen Abschnitten“ definiert (=Prosodie B). Die intrinsische Struktur der Laute (=Prosodie C) beschreibt subtile Vorgänge wie zum Beispiel „die korrekten Formantübergänge zwischen benachbarten Lauten“ (Tillmann, Schiel 1995) oder die Kombination von der Anregung der Stimmbänder und der Verengung des Resonanzraumes bei stimmhaften Konsonanten. Die semantischen Inhalte, die mit der Sprache transportiert werden, sind durch die B- und insbesondere die C- Prosodie bestimmt. (Tillmann, Schiel 1995)

3 Wirkung akustischer Ereignisse

Ein wichtiger Aspekt in Bezug auf die Gestaltung der Tonspur im Spielfilm ist, dass die Klangeigenschaften und Parameter nicht zwingend einen Zusammenhang mit der Wirkung von Hörereignissen bzw. deren Interpretation durch die RezipientInnen besitzen. Hannes Raffaseder bezeichnet dies als die *Semantische Lücke*. Dabei spielt der zeitliche, räumliche, kulturelle und der universelle Kontext, in dem ein akustischer Reiz verarbeitet wird sowie persönliche Erfahrungen, Erinnerungen, Erwartungen, ebenso wie die aktuelle Stimmung und Befindlichkeit eine wesentliche Rolle, wie der Reiz interpretiert wird, welche Wirkung er hat und welche Informationen oder Emotionen er transportiert. (Raffaseder 2010, S. 33–34) Im Sinne der akustischen Kommunikation transportieren Schallereignisse daher keine Emotionen oder Informationen. Diese werden erst durch den Prozess der auditiven Wahrnehmung bedeutungsvoll, also durch die auf der kognitiven Stufe der Wahrnehmung stattfindende Umwandlung eines physischen Schallereignisses in ein Hörereignis. (Görne 2017, S. 29)

Eine große Schwierigkeit in der akustischen Gestaltung besteht somit nicht nur darin, physikalische Eigenschaften wie die spektrale Zusammensetzung oder die räumliche Eigenschaften entsprechend der vorangehenden Intention (das heißt, welche Wirkung erzielt werden soll) zu bearbeiten. Das geschaffene Schallereignis muss von den RezipientInnen wiederum in ein Hörereignis übersetzt werden, wobei die wahrnehmbaren physikalischen Eigenschaften (zum Beispiel Tonhöhe oder Klangfarbe) interpretiert werden müssen, wie auch der semantische oder informative Inhalt des akustischen Ereignisses richtig gedeutet werden muss. Damit die intendierte Wirkung bei den RezipientInnen eintritt, müssen persönliche Erfahrungen und Erwartungen, gesellschaftliche Konventionen sowie die akustischen und narrativen Rahmenbedingungen im Sinne der erzielten Wirkung vorhanden sein und auch bei der ursprünglichen Gestaltung berücksichtigt werden.

Ursprünglich ist jedes Schallereignis abstrakt, erst durch die Umwandlung in ein Hörereignis durch die auditive Wahrnehmung wird daraus ein konkreter Klang, dem eine Schallquelle (oder auch eine semantisch weitreichendere Bedeutung) zugewiesen

wird. Dabei unterscheidet sich das wahrgenommene Hörereignis bei jeder wiederholten Wahrnehmung des gleichen Schallereignisses von den bisherigen Hörereignissen (auch wenn vielleicht nur geringfügig), ebenso wie sich das gebildete Hörereignis eines gleichen Schallereignisses individuell von Mensch zu Mensch anders darstellt. (Görne 2017, S. 85) Interessant ist auch, dass akustische Ereignisse mit annähernd gleichen technisch erfassbaren Messwerten hinsichtlich ihrer semantischen Bedeutung unterschieden werden können und somit zu unterschiedlichen Hörereignissen mutieren. Am Vergleich der Aufnahme eines Gebirgsbaches und einer Autobahn aus einiger Entfernung zeigt Hannes Raffaseder auf, dass es sich bei beiden Signalen um ein breitbandiges Rauschen handelt, der Gebirgsbach aber beispielsweise mit unberührter Natur in Verbindung gebracht wird, die Autobahn aber für „moderne, hektische, lärmende Mobilität“ steht. (Raffaseder 2010, S. 50–51) Somit ist der jeweilige Kontext ganz entscheidend, welche Bedeutung das Rauschen für die HörerInnen besitzt.

Entscheidend in der akustischen Gestaltung ist die Wirkung bzw. Bedeutung, die die eingesetzten Klänge und Geräusche auf die RezipientInnen haben. Dabei gibt es unterschiedliche Möglichkeiten, wie die verschiedenen Elemente der Tonspur eine Wirkung erzielen oder eine Bedeutung erlangen können.

3.1 Bedeutung erster Ordnung

Klänge und Geräusche beeinflussen sowohl die Psyche als auch den Körper des Menschen. Schallwellen haben direkte physikalische Einwirkungen auf den menschlichen Organismus. Die von Schallereignissen ausgelösten Luftdruckschwankungen versetzen nicht nur das Trommelfell, sondern auch den Körper in Schwingung. Deutlich wird dies bei tiefen Signalanteilen mit hoher Intensität, deren Resonanz besonders im unteren Rücken, im Becken und in den Beinen zu spüren ist. Höhere Frequenzanteile fühlt der Rezipient dagegen eher im Brustbereich, Hals und Kopf. Hörempfinden kann biologische Funktionen wie Herzfrequenz und

Körpertemperatur, die Atmung oder die Schweißproduktion beeinflussen, sowie im Extremfall Schmerzen auslösen. (Sonnenschein 2001, S. 70–71) Auch können Körperreaktionen wie Übelkeit und Erbrechen durch die Einwirkungen von Schallereignissen mit bestimmten Frequenzmustern hervorgerufen werden. Naturereignisse wie zum Beispiel Donner lösen ebenso wie bestimmte, mit grundlegenden Bedürfnissen in Verbindung stehende Laute von Menschen oder Tieren instinktiv und unabhängig von sozialen und kulturellen Konventionen eine Reaktion aus. (Raffaseder 2010, S. 44) Gleichzeitig beeinflussen Klänge und Geräusche auch Emotionen. Dies geschieht zum einen durch die Bedeutung, die einem Hörereignis zugeordnet wird, (Lensing 2006, S. 197) aber auch unterbewusst und im Zusammenspiel mit den physikalischen Einwirkungen auf den menschlichen Organismus. Dabei lässt sich ein Bezug zur ersten Wahrnehmungsebene (Wahrnehmung zur Überwachung der Umwelt) herstellen, da eine bekannte, gleichbleibende akustische Umgebung ein Sicherheitsgefühl hervorrufen und somit entspannend bzw. beruhigend sein kann, „während ungewohnte, starken und raschen Veränderungen unterworfenen Sounds zu einer erhöhten Aufmerksamkeit führen und aktivierend wirken“ (Raffaseder 2010, S. 45). In diesem Zusammenhang können Hörereignisse auch zur Stimmungsregulierung eingesetzt werden. Sehr deutlich ist dies beim Hören von Musik zu beobachten da unterschiedliche Musikstile bzw. –Titel beispielsweise fröhliche oder melancholische Gefühle hervorrufen können, auch wenn die Musik den HörerInnen nicht bekannt ist.

In Zusammenhang mit der zweiten Wahrnehmungsebene, dem „Hören zur Orientierung in der Umwelt“ oder dem von Chion beschriebenen Casual Listening, gibt es mit dem „Informationsgehalt akustischer Ereignisse“ eine weitere Möglichkeit, wie Schallereignisse eine Wirkung auf das Filmpublikum ausüben können. (Raffaseder 2010, S. 46) Die auditive Wahrnehmung versucht evolutionär bedingt allen akustischen Ereignissen ein Objekt (also eine Schallquelle) zuzuordnen und so die an sich abstrakten Klänge bzw. Geräusche zu kategorisieren und entsprechende Reaktionen zu setzen. Dies schlägt sich auch in der verbalen Beschreibung von akustischen Ereignissen nieder, wo meist die Schallquelle, deren Beschaffenheit oder die Art der Anregung miteinbezogen wird sowie lautmalerische Begriffe verwendet werden. Die verbale Beschreibung funktioniert allerdings nur unzureichend, da sie die nahezu

unbegrenzten Möglichkeiten bzw. Abstufungen der Eigenschaften (insbesondere der Klangfarbe) nicht abdecken können. Auch die Kategorisierung von Schallereignissen in Sound-Bibliotheken erfolgt oft anhand verbaler Beschreibungen der Schallquelle sowie deren materieller Beschaffenheit. Diese stehen aber nicht oder nur bedingt in Zusammenhang mit der Bedeutung des Hörereignisses, da die Klänge oder Geräusche je nach Verwendung vom Filmpublikum in einen anderen räumlichen, zeitlichen und konventionellen Kontext gesetzt werden. (Flückiger 2007, 78-81, 102-103)

„Wir nehmen keinen Schall wahr sondern eine Hypothese über den Ursprung des Schalls“ (Görne 2017, S. 35) Akustische Ereignisse sind in ihrer Natur flüchtig und nicht greifbar (Raffaseder 2010, S. 21) und es ist erst seit den technischen Errungenschaften der Schallaufzeichnung und -Wiedergabe möglich, die Gleichzeitigkeit der Schallentstehung und dessen Auslösung aufzuheben, also den Klang oder das Geräusch von seiner ursprünglichen Schallquelle zu trennen. Da akustische Ereignisse immer das Resultat von physikalischen Vorgängen sind, haben die als Schall wahrgenommenen Luftdruckschwankungen gewisse, durch den Auflösungsprozess bestimmte Eigenschaften. Hannes Raffaseder spricht von der ikonischen Verknüpfung von Schall- und Lautereignis, also einer Abbildung des physikalischen Ereignisses im Hörereignis. So vermittelt jedes Schallereignis Informationen über die Eigenschaften der Schallquelle, also über die Größe, das Material, den Ort und den Bewegungszustand. Der Auslösungsprozess spiegelt die Art, Stärke und Geschwindigkeit der Anregung wiederum sowie in der Beschaffenheit des Auslösers (wobei hier wiederum Größe, Form und Material hörbar werden). Ebenfalls werden durch akustische Ereignisse Informationen zum Ort der Auslösung und die Entfernung zu den EmpfängerInnen, sowie über die Größe und Beschaffenheit des Raumes geliefert (siehe Kapitel 2.2.3 – Räumlichkeit und Richtung). (Raffaseder 2010, S. 46–49)

Die Identifizierung der Quelle anhand deren Informationsgehalt ist in der Regel problemlos möglich, solange das grundlegende Vorwissen bei den EmpfängerInnen vorhanden ist. Beispielsweise wird in den meisten Kulturkreisen ein vorbeifahrendes Auto als solches durch das dabei entstehende Geräusch erkannt. (Flückiger 2007, 80, 109) Durch die ikonische Verknüpfung von Schall und Lautereignis (also der Neigung der auditiven Wahrnehmung, einem Schallereignis eine entsprechende Schallquelle

zuzuordnen) eröffnen sich für die akustische Gestaltung weitreichende Möglichkeiten, Geräusche zu verändern und zu bearbeiten oder in anderen akustischen Szenen bzw. in einen anderen Kontext zu überführen, ohne dass diese vom Filmpublikum als unnatürlich wahrgenommen werden. (Flückiger 2007, 80, 109) So ist es für die FilmzuseherInnen nicht relevant, ob es sich beim Geräusch eines umfallenden Baumes im Wald tatsächlich um die Aufnahme eines solchen handelt oder um die Aufnahme zerbrechender Spaghetti und eines raschelnden Magnetbandes (Görne 2017, S. 83–84) oder ob die akustische Untermalung der auf der Bildebene sichtbaren Faustschläge durch Geräusche erfolgt die beim Zerhacken von Gemüse entstehen. (Flückiger 2007, S. 80) Begünstigt wird dies durch den Effekt der Synchrese (mehr dazu in Kapitel 4.2.1) welcher das Phänomen beschreibt, dass durch die Synchronität von Ereignissen auf der Bild- und Tonebene das visuelle und das akustische Ereignis von der menschlichen Wahrnehmung miteinander verknüpft werden. (Flückiger 2007, S. 141) So werden die akustischen Sinnesreize automatisch mit einer im Bild sichtbaren Schallquelle verbunden, auch wenn diese möglicherweise nicht an dessen Entstehung beteiligt war. „In diesem Sinne gibt es also keine akustische ‚Realität‘ und keine ‚authentischen‘ Sounds.“ (Raffeseder 2010, S. 237)

Den Informationsgehalt akustischer Ereignisse bezeichnet Thomas Görne als Semantik 1. Ordnung und baut diesen Begriff in Bezug auf die Wirklichkeit 1. Ordnung von Paul Watzlawick auf. (Görne 2017, S. 106–107) Die Wirklichkeit 1. Ordnung „bezieht sich auf die rein physischen und daher weitgehend objektiv feststellbaren Eigenschaften von Dingen und damit entweder auf Fragen des sogenannten gesunden Menschenverstands oder des objektiven wissenschaftlichen Vorgehens.“ (Watzlawick 2004, S. 182) Die Semantik 1. Ordnung beschreibt also die *Dinghaftigkeit* von den durch die auditive Wahrnehmung geschaffenen Hörereignissen, dessen Kategorisierung und Verknüpfung mit einem Objekt (also der Schallquelle). Schallereignisse, die darüber hinaus eine Bedeutung erlangen, fallen unter den Begriff der Semantik 2. Ordnung, welche dadurch gekennzeichnet ist, dass akustischen Ereignissen erst durch den Menschen bzw. durch den Vorgang der Wahrnehmung eine Bedeutung beigemessen wird, also vom Menschen geschaffen wird. (Görne 2017, S. 106–107)

3.2 Bedeutung höherer Ordnung

Die Wirklichkeit 2. Ordnung „beruht ausschließlich auf der Zuschreibung von Sinn und Wert an diese Dinge und daher auf Kommunikation.“ (Watzlawick 2004, S. 182) Zur Semantik 2. Ordnung zählen die „kommunikative Bedeutung, [...] Zeichenhaftigkeit oder Symbolik“ (Görne 2017, S. 106) von akustischen Ereignissen, die Textsemantik, nach der menschlichen Lauten durch Konventionen (im Fall der Sprache den Regeln der Grammatik) eine Bedeutung beigemessen wird, sowie der Kontextsemantik. Der explizite Kontext beschreibt den Umstand, dass akustischen Ereignissen je nach Ort, Zeit sowie je nach Konvention (beispielsweise dem Klang einer Klingel) eine unterschiedliche Bedeutung zukommt. Der narrative Kontext bezieht sich auf die Bedeutung, die ein Hörereignis durch vorangehende Informationen in der Erzählstruktur oder in einem Genre erlangt. (Görne 2017, S. 106-107)

Unter der Zeichenhaftigkeit versteht man Klangobjekte mit „expliziter Bedeutung und expliziter kommunikativer Funktion“, es wird also bewusst zur Kommunikation (zwischenmenschlich oder immer öfter auch zur Kommunikation mit technischen Geräten) eingesetzt, wobei die Bedeutung kulturell verankert ist oder sich aus dem Kontext erschließt. (Görne 2017, 116–117) Barbara Flückiger spricht in diesem Zusammenhang von Signalen, wobei diese häufig einen Warnhinweis darstellen und als Handlungsaufforderung zu verstehen sind. Die Zusammensetzung der Eigenschaften von Signalen ist meist sehr einfach und es handelt sich oft um Klänge und Geräusche, die sich in einem für das menschliche Ohr empfindlichen Frequenzbereich abspielen. (Flückiger 2007, S. 159) Musterbeispiele sind Klingeln, Hupen oder Sirenen, also Alarmer oder Signaltöne mit denen auf bestimmte Ereignisse hingewiesen oder Reaktionen hervorgerufen werden sollen. Obwohl sich der Klang mitunter nicht wesentlich voneinander unterscheidet, ist beispielsweise die Bedeutung einer Klingel sehr unterschiedlich, je nachdem ob es sich um eine Tür- oder um eine Fahrradklingel handelt. Die Bedeutung ist also entweder durch kulturelle Standards definiert oder erfolgt durch Metakommunikation. (Görne 2017, S. 116–117) Bei akustischen Ereignissen, welchen durch Konvention eine bestimmte Bedeutung beigemessen wird, besteht laut Hannes Raffaseder „eine abstrakte Verknüpfung, die bewusst erlernt

werden muss, um verstanden zu werden.“ (Raffaseder 2010, S. 51–52) Dieser Kommunikationsprozess setzt also entweder eine Metakommunikation voraus, welche die Rahmenbedingungen, in denen der Symbolgehalt interpretiert werden muss, definiert, sofern diese nicht bereits durch kulturelle Konvention oder durch genrespezifische Normen vorgegeben sind. (Görne 2017, S. 106) Die festgelegten Konventionen sind daher in der überwiegenden Anzahl der Fälle nur innerhalb eines bestimmten Kulturkreises gültig. Über die Bedeutung von meist aus einzelnen Signalen bestehenden Warnhinweisen hinausgehend gibt es auch akustische Zeichensysteme, denen durch Konvention eine Bedeutung beigemessen wird, wie beispielsweise Morsezeichen oder Alarmsignale wie Feuerwehirsirenen, wobei durch die Anzahl und die Länge der einzelnen Signale festgelegt ist, welche Bedeutung diese transportieren. Wesentlich komplexere Systeme, bei welchen durch Konvention die Bedeutung definiert ist, sind menschliche Sprachen. (Raffaseder 2010, S. 52) Thomas Görne bezeichnet dies als die Textsemantik (Görne 2017, S. 106), wobei zwischen einzelnen Lauten, Silben und Wörtern bis hin zu ganzen Sätzen – und wiederum zwischen diesen Sätzen – ein komplexes Netzwerk an mehrstufigen Verknüpfungen besteht, das von den Hörer- oder Leser/Innen nach den Regeln der Grammatik der jeweiligen Sprache interpretiert werden muss, um schließlich die transportierte Bedeutung zu erschließen, wodurch die einzelnen Laute eine Semantik höherer Ordnung erlangen. (Raffaseder 2010, S. 52) Abgesehen von ihrer semantischen Bedeutung in Bezug auf die Sprache vermittelt die menschliche Stimme noch viele weitere Informationen. Unterschiedlichste Laute wie Lachen, Weinen, Seufzen oder Angstschreie werden unabhängig von Konventionen, wie der Beherrschung einer Sprache oder kulturell bedingter Interpretation von Verhaltensmustern, in Bezug auf die Bedeutung und den emotionalen Informationsgehalt verstanden. Dies betrifft aber nicht nur jene Laute, die keinen semantischen Mehrwert besitzen, sondern es werden Informationen hinsichtlich der emotionalen Bedeutung auch durch Sprache vermittelt. Dabei ist es unwesentlich, ob die einzelnen Wörter und Sätze verstanden werden bzw. ob die jeweilige Sprache überhaupt beherrscht wird. Die Wahrnehmung von Klangfarbe, Rhythmus sowie die Variation der Töne transportieren dennoch, ob der oder die SprecherIn beispielsweise aufgeregt, ängstlich oder verärgert ist. Zieht man zu diesen Eigenschaften noch die Grundfrequenz hinzu, können akustisch problemlos Rückschlüsse über das Alter oder

das Geschlecht der sprechenden Person gezogen werden, sowie mehrere sprechende Personen voneinander unterschieden werden. (Raffaseder 2010, 44-45, 240-241) Es gibt Situationen in denen die Bedeutung, die dem sprachlichen Ausdruck (insbesondere der Prosodie) beigemessen wird, dem semantischen Informationsgehalt überwiegt. So lässt sich an der Sprachmelodie beim Ausruf eines Namens meist erkennen, warum die Person gerufen wird, bzw. ob der Sprecher beispielsweise besorgt oder verärgert über die gerufene Person ist. (Sonnenschein 2001, S. 138)

Die Bedeutung von Symbolen dagegen ist schwieriger fass- und nur unmittelbar erlebbar. (Görne 2017, S. 116–117) Nach David Sonnenschein erlangen Klänge oder Geräusche dann einen Symbolcharakter, wenn diese, abgesehen von ihrer kommunikativen Funktion und physischen Empfindungen, bestimmte Gefühlszustände oder Gedanken hervorrufen und so mit ihrer Bedeutsamkeit in die Tiefen unserer Psyche vordringen. (Sonnenschein 2001, S. 205) Für Barbara Flückiger sind „Symbole in außerfilmischen, mehrheitlich religiösen, mythischen oder sozialen Traditionen verankert“, besitzen eine Bedeutung bzw. Semantik höherer Ordnung und stehen stellvertretend für ein abstraktes Konzept, das von den HörerInnen interpretiert werden muss. Dabei ist es ein Merkmal von Symbolen, dass sie nicht restlos gedeutet werden können, sondern sich ein Teil davon an das Unterbewusstsein wendet. (Flückiger 2007, S. 164) In Symbolen werden das Bewusste und das Unterbewusste miteinander vereint. (Atmanspacher et al. 1995, S. 115) Hannes Raffaseder spricht in Zusammenhang mit der Symbolik von einer metaphorischen Verknüpfung zwischen Schall- und Lautereignis. So werden Schallereignisse mit Erinnerungen an bereits erlebte Erfahrungen oder kulturell überlieferte Ereignisse verknüpft und es entstehen Assoziationen der Schallereignisse mit diesen Erfahrungen und Ereignissen. „Akustische Ereignisse werden so zu Metaphern für diese Vorgänge und erhalten übergeordnete Bedeutung.“ (Raffaseder 2010, 49, 51) Eine genaue Abgrenzung zwischen Zeichen und Symbolen ist aber schwierig, da beispielsweise das Donnergerollen in verschiedenen Kulturen als göttliche Stimme galt, ebenso aber auch nur die Ankündigung eines nahenden Gewitters signalisieren kann, wobei wiederum der explizite und der narrative Kontext entscheidet, welche Funktion das Geräusch ausübt. (Görne 2017, S. 117)

3.3 Formen der semantischen Verknüpfung

Der eigentliche Wahn liegt in der Annahme, daß es eine »wirkliche« Wirklichkeit zweiter Ordnung gibt und daß »Normale« sich in ihr besser auskennen als »Geistesgestörte«. (Watzlawick 2004, S. 184)

Da ein Hörereignis von jeder Person individuell gebildet wird, welche jeweils individuelle „Erinnerungen und Assoziationen, subjektive Erfahrung und Vorwissen“ in Bezug auf das Schallereignis haben, sowie sich die momentane Befindlichkeit unterscheiden kann, ist auch die Bedeutung, die das akustische Ereignis erlangt, niemals gänzlich die gleiche. Dementsprechend gibt es auch akustische Zeichen und Symbole, die nur für einen einzigen Menschen eine tiefere Bedeutung haben. Als Beispiel hierfür können Musiktitel dienen, die von einer Person in einem bestimmten Moment oder Lebensabschnitt gehört wurden und so jedes Mal Erinnerungen daran hervorrufen. (Raffaseder 2010, S. 49–50) Praktisch kann also je nach persönlicher Erfahrung und Entwicklung alles zu einem akustischen Symbol werden. (Görne 2017, S. 118) Wie nachfolgend gezeigt, gibt es abgesehen von der individuellen Bedeutungszuweisung mehrere Möglichkeiten, wie die Schallereignisse Symbolcharakter bzw. weiter gefasst, Bedeutung hinsichtlich einer Semantik höherer Ordnung, erlangen können. (Sonnenschein 2001, S. 205)

Bedeutung durch kulturelle Verknüpfungen

Das kulturelle Umfeld des Rezipienten bestimmt maßgeblich, welche Bedeutung ein akustisches Ereignis für die HörerInnen erlangt. Sonnenschein beschreibt dies unter anderem am Beispiel des Namens einer US-amerikanischen Baseball-Mannschaft (die *Dogers*), zu dem die meisten Menschen in deren Heimatstadt Los Angeles einen anderen Bezug haben als die Einwohner einer beliebigen anderen Stadt in den USA. Dieser unterscheidet sich wiederum vom Verhältnis, den eine Person aus London (die „möglicherweise schon einmal von diesem amerikanischen Sport-Team gehört hat“) sowie vom Bezug einer Person aus Brasilien, die den Namen möglicherweise noch nie gehört hat. So wird aus einem Wort mit einer symbolischen Bedeutung ein Wort, mit

dem keine Assoziationen verbunden sind, oder auch nur eine Aneinanderreihung von Lauten, die aufgrund sprachlicher Differenzen nicht als Wort verstanden werden kann. (Sonnenschein 2001, S. 205–212) Der Klang von Glocken wiederum wurde hinsichtlich seiner klanglichen Eigenschaften (Lautstärke, Klangfarbe etc.) sowohl von japanischen als auch von deutschen Probanden als sehr ähnlich eingestuft, die Bedeutung höherer Ordnung war jedoch sehr unterschiedlich. Bei deutschen Testpersonen riefen die Glocken eher angenehme Assoziationen hervor. Sie stellten sich Glocken einer Kirche vor, während der gleiche Klang von den japanischen Testpersonen mit Warnsignalen (wie einem Feuersalarm oder Signalen an einem Bahnübergang) „in Verbindung gebracht und dadurch mit Gefahr assoziiert wurden“. (Flückiger 2007, S. 159) Als weiteres Beispiel für die Bedeutung, die durch das kulturelle Umfeld entsteht, kann das Meeresrauschen dienen. Das Geräusch weckt bei vielen Menschen aus Mitteleuropa angenehme, entspannende Gefühle und Erinnerungen an Urlaube, während es von einem einheimischen Fischer mit seiner täglichen anstrengenden Arbeit in Verbindung gebracht wird. (Raffaseder 2010, S. 49)

Bedeutung durch historische Verknüpfungen

Eine weitere Möglichkeit Bedeutung zu erlangen ist der historische Kontext von akustischen Ereignissen. So können sich die Assoziationen, die Zeichen oder Symbole durch kulturelle Bezüge erhalten, über die Zeit verändern. Es gibt Klänge und Geräusche, die für eine bestimmte Epoche charakteristisch waren und von den Menschen als Teil des täglichen Lebens wahrgenommen wurden, mittlerweile bei den HörerInnen jedoch nostalgische Gefühle hervorrufen. Beispiele dafür sind die Geräusche einer Schreibmaschine, die noch vor wenigen Generationen mit technologischem Fortschritt und Arbeitserleichterung in Verbindung gebracht wurden, mittlerweile jedoch fast verschwunden sind. Als weiteres Beispiel dienen etwa die Geräusche eines Dreschflegels, welche vor der Industrialisierung die akustische Szenerie in der Landwirtschaft geprägt haben, in der heutigen Zeit jedoch den meisten Menschen als Gerät und somit aber auch nicht mehr als Quelle von charakteristischen Geräuschen bekannt sind. In weniger fortschrittlichen Ländern hingegen gehört dieses Geräusch möglicherweise noch zur alltäglichen Arbeitswelt, womit es für die dort

lebenden Menschen wiederum eine andere Bedeutung aufweist. (Raffaseder 2010, S. 34) Einen besonderen Stellenwert haben kulturelle und historische Bezüge bei der Gestaltung der Atmo, da hier sehr wirkungsvoll die Zeit und der Ort der Szene etabliert werden können. Dadurch muss aber auch ein besonderes Augenmerk darauf gelegt werden, dass die akustische Umgebung die richtigen Referenzen enthält. Beispielsweise sind die akustischen Umgebungen von verschiedenen Städten durch unterschiedliche, charakteristische Geräusche gekennzeichnet (Sonnenschein 2001, S. 182-183) , wodurch die Montage einer beliebigen Atmo einer Stadt bei unbedachtem Einsatz unglaublich oder irritierend sein kann. Wichtig ist, dass sich die Atmos aus verschiedenen historischen Epochen in deren Klangeigenschaften nicht unwesentlich unterscheiden. Vor der sogenannten Industriellen Revolution setzte sich die akustische Umgebung vorrangig aus Geräuschen zusammen, die meist impulsartig waren, deren Position deutlich ortbar und das Spektrum durch sehr wenige tieffrequenten Anteile ausgezeichnet war. Mit der zunehmenden Verbreitung von Industrie und Automobilität wurden die Umgebungsgeräusche eher konstant, ungerichtet und somit schwieriger ortbar sowie durch viele frequentierte Anteile gekennzeichnet. In den letzten Jahrzehnten werden die akustischen Umgebungen zunehmend durch weniger mechanische, mehr elektronische Geräuschen definiert, die ungerichtet, aber mit mehr hohen Frequenzanteilen auftreten. (Sonnenschein 2001, S. 96–97)

Bedeutung durch universelle Verknüpfungen

Neben kulturellen und historischen Assoziationen gibt es auch Schallereignisse mit einer universellen Bedeutung unabhängig von der kulturellen oder persönlicheren Prägung oder historischen Bezügen. So wird beispielsweise das quietschende Geräusch von Kreide an einer Tafel von den meisten Menschen als unangenehm bzw. schmerzhaft empfunden. (Sonnenschein 2001, S. 205) Ebenso können menschliche oder auch tierische Laute, die mit grundlegenden Bedürfnissen in Verbindung stehen (wie zum Beispiel Angstschreie oder Laute zur Äußerung von Schmerzen), unabhängig vom Kulturkreis oder von der persönlichen Erfahrung verstanden werden. Ähnlich verhält es sich mit Lachen oder Weinen. (Raffaseder 2010, S. 44) Eine besondere Rolle spielen hierbei sogenannte archetypische Ereignisse, die unbewusst bei allen

Menschen Emotionen auslösen. Thomas Görne baut diese Theorie auf dem Archetypenbegriff des Begründers der analytischen Psychologie C. G. Jung auf (Görne 2017, S. 117), wobei es sich bei einem Archetypen um universell vererbte seelische Strukturen in allen Menschen handelt. Nach Jung sind diese im Gegensatz zum Bewusstsein und zum persönlichen Unbewussten Bestandteile des kollektiven Unbewussten, welches „nicht mehr persönlicher Erfahrung und Erwerbung entstammt, sondern angeboren ist“. (Jung 2001, S. 7–9) Bei archetypischen Klangereignissen handelt es sich oft um elementare Naturereignisse wie das Donnergrollen. (Görne 2017, S. 117). Diese lösen, ebenso wie zum Beispiel das Gebrüll eines Löwen, unabhängig von der kulturellen oder individuellen Prägung durch evolutionär verwurzelte Vorgänge der Wahrnehmung bestimmte Emotionen, sowie teils lebenserhaltende Reaktionen aus, wie Kampfhaltung oder Flucht. (Sonnenschein 2001, S. 205)

Bedeutung durch filmspezifische Verknüpfungen

Die vierte Möglichkeit nach Sonnenschein ist die filmspezifische Verknüpfung. (Sonnenschein 2001, S. 210) Die Bedeutung von akustischen Ereignissen wird auch vom jeweiligen Genre des Films, sowie vom narrativen Kontext beeinflusst. (Görne 2017, S. 106) Es gibt zum einen genreübergreifende Klischees, wie sie beispielsweise in sehr vielen Filmen in Form des entfernten Bellen eines Hundes sowie das Zirpen von Grillen zur Etablierung der Tageszeit Nacht verwendet werden (Raffaseder 2010, S. 243), oder zum anderen genrespezifische Klischees, die vom Filmpublikum in diesem Genre erlernt wurden und dementsprechend auch erwartet werden.

Bei genreübergreifenden und -spezifischen Klischees handelt es sich um sogenannte Stereotypen. Aus sozialwissenschaftlicher Sicht beschreibt der Stereotypen-Begriff „die Diskrepanz zwischen vereinfachten inneren Vorgängen und komplexer äußerer Welt zum Zweck der Orientierung.“ Das heißt, die Bildung von Stereotypen erfolgt zur Vereinfachung der Kommunikation, Interaktion und Orientierung mit der Außenwelt. (Flückiger 2007, S. 176–178) Im Bereich der Medientheorie ist der Stereotypen-Begriff von der negativ behafteten Verwendung in der Sozialpsychologie zu unterscheiden, in der dieser in Bezug auf Vorurteile verwendet wird, welche die zwischenmenschliche Kommunikation und Urteilsbildung negativ beeinflussen. (Wuss 1993, S. 162)

Filmstereotypen sind charakterisiert durch „die massenhafte Wiederholung von erzählerischen Konstellationen und damit einhergehenden optischen und akustischen Repräsentationsformen [...], das heißt Strukturen, die über eine größere Anzahl von Filmen während längerer Zeit unverändert bleiben“. Die Bildung von Stereotypen erfolgt somit durch einen Lernprozess, bei dem durch häufige Wiederholung des gleichen Musters schließlich eine filmische oder genrespezifische Konvention entsteht. Werden diese über einen längeren Zeitraum nicht mehr bedient, verschwinden die Stereotypen schließlich wieder. Die Nützlichkeit der Verwendung von Stereotypen für den Einsatz in der Narrationsstruktur im Spielfilm beruht auf deren Unauffälligkeit, da diese kaum mehr bewusst registriert werden, je besser sie etabliert sind. Stereotypen erleichtern die Kommunikation mit dem Filmpublikum und ermöglichen es, komplizierte und verschlungene Zusammenhänge einfach darzustellen und so die Wahrnehmung für die Aufnahme von anderen Informationen zugänglich zu machen. Es kommt also zu einer Komplexitätsreduktion. (Flückiger 2007, S. 176–182) Jörg Lensing spricht in diesem Zusammenhang von einem kollektiven Gedächtnis des Filmpublikums, das heißt von Erwartungen der FilmzuseherInnen, die bedient werden können, um Emotionen oder Informationen zu kommunizieren. (Lensing 2006, S. 198) Bei einem unreflektierten Umgang mit symbolbehafteten Klängen und Geräuschen (insbesondere Stereotypen) können diese schnell übertrieben oder lächerlich wirken, wenn sie beispielsweise im falschen Moment eingesetzt werden. (Sonnenschein 2001, S. 206) Es sollte daher in der akustischen Gestaltung ein vorsichtiger Gebrauch von symbolhaltigen Elementen gepflegt werden. (Raffaseder 2010, S. 50) Es kann aber bewusst mit Stereotypen gespielt werden. Durch einen übertriebenen, plakativen Einsatz von stereotypischen Klängen und Geräuschen, oder gegenteilig, dem Brechen von genrespezifischen, filmischen, oder auch außerfilmisch-kulturellen Konventionen, können komödiantische Elemente entstehen (Sonnenschein 2001, S. 209). Eine gängige Methode für den Einsatz von Stereotypen ist „die Betonung von Ähnlichkeiten“, wodurch ein Wiedererkennungswert geschaffen wird, oder „Die Betonung von Unterschieden“ (Flückiger 2007, S. 178–182), welche es dem Filmpublikum erleichtert, Kontraste wahrzunehmen. Eine weitere Funktion ist es, dem Filmpublikum die Rahmenbedingungen des Films bzw. des Genres und dadurch ein Gefühl von Ordnung und Sicherheit vorzugeben, wodurch es ihm möglich ist, negative Emotionen und

Gefühle dennoch als lustvoll zu erleben. So ist in Action- und Horrorfilmen eine Häufung dieser Stereotypen festzustellen. Viele akustische, filmische Stereotypen unterscheiden sich nicht unwesentlich von deren echten Erscheinungen in Bezug auf deren klangliche Eigenschaften, wodurch es bis zu einer vorherrschenden Verzerrung der Realität in der Gesellschaft kommt. So sind beispielsweise Faustschläge und Schüsse in den meisten Filmen sehr deutlich überzeichnet. Wird eine realistische, unbearbeitete Aufnahme dieser Geräusche in der Tonspur verwendet, wirkt diese für das Publikum sehr unspektakulär oder sogar unglaubwürdig. (Flückiger 2007, S. 178–182) Des weiteren schreibt Flückiger in Bezug auf filmische Stereotypen, dass bei der Gestaltung der Tonspur mit einem höheren gestalterischen Anspruch diese nur dann direkt gebraucht werden, wenn sie „unmittelbar zum automatisierten Bedeutungstransfer“ eingesetzt werden (Flückiger 2007, S. 182). Andernfalls dienen sie nur als Ausgangsbasis und werden entsprechend den Anforderungen des jeweiligen Filmes verändert oder ergänzt.

Darüber hinaus gibt es auch unabhängig von diesen filmischen Stereotypen die Möglichkeit, durch geschickten Einsatz akustische Ereignisse im Handlungsverlauf zu etablieren, so dass diese einen Symbolcharakter erlangen. (Sonnenschein 2001, S. 210)

Symbole sind wie Indizien, die im Diskurs auftauchen und durch strategische Platzierung im narrativen Kontext oder durch Wiederholung eine gewisse Aufmerksamkeitsoptimierung erfahren. Wenn der Zuschauer in einem Western von Sergio Leone, der die Abrechnung mit einem Verräter thematisiert, mitten in der Stille einen Hahn dreimal krähen hört, so sind diese Bedingungen hinreichend, um einen symbolorientierten Interpretationsprozess in Gang zu setzen. (Flückiger 2007, S. 164)

Wurde ein akustisches Symbol im narrativen Konstrukt bereits etabliert und/oder bestimmten Schauplätzen oder Objekten im Film zugeordnet, muss dieses (bzw. dessen klangliche Eigenschaften) auch beibehalten werden, da die auditive Wahrnehmung in diesem Zusammenhang sehr empfindlich auf Veränderungen reagiert. Sofern nicht anderes gewünscht und durch Wendepunkte in der Handlung oder der Entwicklung von Charakteren gefordert wird, kann bei einer unbedachten Änderung beim

Filmpublikum Verwirrung entstehen. Auf der anderen Seite ergeben sich weitreichende Möglichkeiten zum Einsatz als narratives Gestaltungsmittel, da ein symbolhaltiges akustisches Ereignis, sobald es einmal etabliert wurde, alleinstehend und unterbewusst Informationen und Emotionen vermitteln kann. (Sonnenschein 2001, S. 210–212) Barbara Flückiger führt in Bezug auf Klänge und Geräusche, die innerhalb eines Filmes (durch deren exponierten Platzierung an dramaturgisch wichtigen Stellen, sowie deren wiederholten Einsatz im Verlauf der Handlung) eine Bedeutung erlangen, den Begriff *Key Sounds* ein. „Der Rezipient erkennt in der Strukturbildung eine Regelmäßigkeit, die zur Hypothese führt, dass mit dem Key Sound eine tiefere Bedeutung verbunden sei.“ (Flückiger 2007, S. 174–175) Darüber hinaus unterscheidet sie Key Sounds aufgrund des Prozesses, wie diese Bedeutung erlangen, von Symbolen, deren Bedeutung aus einem außerfilmischen Kontext, also aus einem bereits bestehenden kulturellen oder persönlichen Wissens- oder Erfahrungsschatz entstammt. Hannes Raffaseder wiederum definiert Key Sounds folgendermaßen:

Als Keysounds werden Sound-Effekte bezeichnet, die nahezu klischeehaft schon nach einem Bruchteil einer Sekunde mit bestimmten Zeiten, Orten, Situationen oder Handlungen assoziiert werden und dadurch die Szenerie und die Narration eines Medienprodukts verdeutlichen können. Eher selten stehen Keysounds als Hard-Effect im akustischen Vordergrund. Meist sind sie als Soft-Effekt oder auch als Teil der Atmo zu hören. (Raffaseder 2010, S. 245)

Diese Definition unterscheidet sich insofern von der Flückigers, nach der „die Bedeutungsmodifikation von Key Sounds nicht auf die Relation zu bestimmten Ereignissen, Orten, Figuren oder Ideen“ zurückgeht. (Flückiger 2007, 174–175)

Zusammenfassend sind in der Gestaltung der Tonebene die verschiedenen Möglichkeiten, wie akustischen Ereignissen Bedeutung (insbesondere hinsichtlich der Semantik 2. Ordnung) beigemessen wird, zu berücksichtigen, um die gewünschte Wirkung zu erzielen. Eine hundertprozentige Vorhersage der erzielten Wirkung ist jedoch nicht möglich. Auf jeden Fall erforderlich ist aber das Wissen über kulturelle, historische, universelle oder filmspezifische Konventionen, damit Verknüpfungen hergestellt oder Assoziationen ausgelöst werden, die Hörereignisse im Sinne der

Erzählstruktur bilden. Zu beachten ist, dass kulturelle, historische universelle und filmspezifische Bezüge nicht individuell voneinander für den Symbolgehalt verantwortlich sind, sondern die unterschiedlichen Verknüpfungen zusammenwirken und somit jede für sich das aus dem Schallereignis entstehende Hörereignis beeinflusst.

4 Die Tonspur als narratives Gestaltungsmittel

4.1 Grundlagen

Im folgenden Kapitel wird die Fragestellung behandelt, wie durch den gezielten Einsatz der unterschiedlichen Elemente der Tonspur ein narrativer Mehrwert geschaffen werden kann. Im Zuge dessen werden die notwendigen Rahmenbedingungen innerhalb des Mediums Film definiert, sowie die Konzepte, die der akustischen Gestaltung im Spielfilm zu Grunde liegen und welche bezüglich der Fragestellung zum Ziel führen, erarbeitet.

Ein Ziel in der filmischen Gestaltung ist es, mit der erzählten Geschichte bzw. den dargestellten Ereignissen das Filmpublikum zu berühren, das heißt Emotionen auszulösen und Stimmungen zu erzeugen. Dabei ist die immersive Wirkung des Mediums eine sehr dankbare Voraussetzung, aber obwohl das Kino und die technischen Errungenschaften der letzten Jahrzehnte die filmische Immersion vor allem in Bezug auf das körperliche Erleben durch bessere optische und auditive Bedingungen begünstigen, entscheidet schlussendlich die „imaginative Kraft der Geschichte“ und die „imaginative Kraft der audiovisuellen Gestaltung“, ob das Filmpublikum bewegt bzw. emotional berührt wird. (Görne 2017, 197-202) Der Spielfilm, wie auch andere audiovisuelle Medien, ermöglicht dem Filmpublikum in die filmische Wirklichkeit einzutauchen. Diese Empfindung wird als Immersion bezeichnet und beschreibt zum einen die kognitive Erfahrung, „sich ganz in einer Geschichte zu verlieren“ (Görne 2017, S. 205), was auch beim Lesen eines Buches oder im Theater möglich ist. Zum anderen beschreibt der Begriff aber die emotional noch intensivere, teils auch körperliche Erfahrung, also dem Erleben der filmischen Realität.

„Der Film als narratives System muss die Fragmente so aufbereiten, dass sie eine kohärente Illusion schaffen. Die Diegese als sinnvolles Ganzes entsteht nur durch die konstruktive Tätigkeit des Zuschauers.“ (Flückiger 2007, S. 77) Nach dem Philosophen Noel Carroll ist dem Filmpublikum bewusst, dass die filmische Wirklichkeit eine Illusion

darstellt, also nicht real ist. Allerdings wird diese Illusion akzeptiert und unsere Gedanken vermögen so, obwohl uns bewusst ist, dass es sich um Fiktion handelt, echte Emotionen zu erzeugen, wie sie auch von realen Ereignissen hervorgerufen werden. Ein beeindruckendes Beispiel hierfür ist der Film *Dogville* (2003) vom dänischen Regisseur und Autor Lars von Trier, in der die Bildebene ganz klar von der Realität abstrahiert ist. Die Kulisse, in der sich die Schauspieler bewegen, ist nichts weiter als eine Fläche, auf der eingezeichnete Bodenmarkierungen Häuser und Straßen darstellen, sowie Türen und Requisiten auf der ansonsten freien Fläche stehen (siehe Abbildung 11). Die Tonebene transportiert jedoch ein sehr realitätsgetreues Abbild der akustischen Umgebung und der Räume, in der sich die Charaktere in der Geschichte gerade befinden. Von den FilmzuseherInnen wird das abstrakte visuelle Erscheinungsbild jedoch nach wenigen Minuten ausgeblendet und so die Illusion akzeptiert. (Görne 2017, S. 197–198)



Abbildung 11: Die Kulisse von Dogville (Lars von Trier, 2003)

In der akustischen Gestaltung muss also, wie auch in der restlichen filmischen Gestaltung, darauf geachtet werden, dass die geschaffene filmische Wirklichkeit vom Publikum als solche akzeptiert wird. Nur dann ist es möglich, die gewünschten Reaktionen beim Filmpublikum hervorzurufen. „Gutes Sounddesign vermittelt Stimmungen und Bedeutungen, die wesentliche Beiträge zur Dramaturgie, zur Narration etc. liefern.“ (Raffaseder 2010, S. 237)

Um das volle Potential der akustischen Gestaltung hinsichtlich der Narration und Dramaturgie im Spielfilm ausnutzen zu können, ist eine intensive Auseinandersetzung mit dem Skript und der narrativen Struktur der Geschichte erforderlich. Im besten Fall ist es möglich, über die Tonspur unabhängig von der Bildebene bestimmte Ereignisse zu erzählen oder Bedeutungen zu kommunizieren, insbesondere aufgrund der, wenn richtig eingesetzten unterschweligen, subtilen Natur der akustischen Kommunikation. Jedoch kann auch eine nüchterne akustische Repräsentation der Geschehnisse der Bildebene zum Ziel führen. Die Voraussetzung muss immer sein, der Geschichte bzw. dem Film zu dienen. Dafür gilt es, den Gesamtkontext zu berücksichtigen, um die richtigen Entscheidungen zur Unterstützung der Dramaturgie treffen zu können und nicht nur die Tonspur um ihrer selbst willen zu gestalten. Eine gute Geschichte wird meist durch eine über den Zeitverlauf des Spielfilms aufrechterhaltene Spannung, Charakterentwicklung und emotionale Einbindung des Filmpublicums charakterisiert. Bei einem Spielfilm mit einer Länge von über zwei Stunden wechseln Konflikte, Konfrontationen und Lösungen einander ab, verändern sich und es tauchen neue Probleme oder Ziele auf, um die Spannung aufrecht zu erhalten. Im oft zitierten *Hero's Journey Model*, das im Spielfilm weit verbreitet ist, handelt es sich beispielsweise um eine Struktur mit einem klar ausgeprägten Anfang, einem Mittelteil und Schluss. Der Protagonist ist unzufrieden mit dem Status Quo (seinem Alltag) und bricht auf (oft unfreiwillig), um auf ihm unbekannten Pfaden ein bestimmtes Ziel zu erreichen. Im Laufe des Weges muss er sich diversen Gegnern oder Problemen, sowie auch seinen inneren Dämonen stellen und kommt am Schluss mit dem hart erkämpften Ziel als Held zurück in sein gewohntes Umfeld. In diesem Model sind deutliche Höhen, Tiefen und Wendepunkte ersichtlich, die essentiell für den weiteren Handlungsverlauf sind. (Sonnenschein 2001, S. 173–174)

Die gewählte Form soll außerdem neue Blickwinkel auf den Handlungsverlauf ermöglichen, Überraschungsmomente vorbereiten und den Rezipienten Freiräume für Phantasie, gedankliche Weiterentwicklung und Ausgestaltung lassen. [...] Oberstes Ziel der Dramaturgie sollte ein in sich schlüssiges, stimmiges, abwechslungsreiches, spannendes Produkt sein. Die richtige Balance zwischen Spannung und Entspannung, Weiterentwicklung und Wiederholung, Unbekanntem und Bekanntem, Erfüllen und Brechen von Erwartungshaltungen ist dafür entscheidend. (Raffaseder 2010, S. 279)

Egal welches Model der Narration bei einem Film zum Einsatz kommt: Ob es sich am Beginn der Konzeption der Tonspur um ein Skript handelt oder bereits um einen fertig geschnittenen Film, es müssen die dramaturgischen Gestaltungsmittel, Höhe - und Wendepunkte der Handlung und der Charakterentwicklung identifiziert werden, um eine sinnvolle, der Narration dienliche akustische Gestaltung zu erarbeiten. (Sonnenschein 2001, S. 174)

„Der Spielfilm, im Besonderen das Kino, ist ein hinsichtlich seiner eigenen Konventionen gewachsener Rahmen, der sich seine eigenen Regeln und Grenzen vorgibt“ (Flückiger 2007, S. 77), was insbesondere deutlich wird, je näher sich ein Film am Mainstream orientiert. Das Medium Film besitzt insgesamt drei Kommunikationskanäle, wobei jeder davon wiederum verschiedene Eigenschaften bzw. Möglichkeiten aufweist, dessen kommunikative Wirkung zu beeinflussen. Der erste Kanal ist die Narration, worunter unter anderem das Schauspiel, Dramaturgie oder die Montage fallen. Der zweite ist die visuelle Gestaltung, die unter anderem in Bezug auf die Farbstimmungen und Lichtverhältnisse oder die Perspektive kommuniziert. Der dritte ist die Gestaltung der Tonspur, deren kommunikative Funktion in weiterer Folge erarbeitet wird. „Sowohl die visuelle als auch die auditive Gestaltung vermitteln Information, unabhängig von der erzählten Geschichte“. (Görne 2017, S. 236) Die Elemente der Tonspur als Teil dieses Rahmens transportieren nicht nur wie in Kapitel 3 beschrieben eine bestimmte Wirkung, also eine mögliche Bedeutung aufgrund historischer, kultureller, sozialer oder filmspezifischer Konvention. Sie beeinflussen sich untereinander gegenseitig, ebenso wie deren Bedeutung in ständiger Wechselwirkung mit der auf der Bildebene vermittelten Information steht. Weiters besteht die Wechselbeziehung auch mit der Bedeutung, die durch den narrativen Kontext, der Handlung bzw. der Geschichte vermittelt wird. (Raffaseder 2010, S. 278) Flückiger betont dabei die Souveränität der Tonebene:

Mit Souveränität ist gemeint, dass die Tonspur aus einer Position der Eigenständigkeit in einen Dialog mit den anderen Elementen des Films tritt. Sie formuliert einen eigenständigen kommunikativen und emotional wirksamen Output, der unter anderem dadurch zustande kommt, dass ausformulierte klangliche Elemente zueinander in Beziehung treten. [...] Die anderen Elemente

[...] werden nicht allein durch die visuelle Repräsentation abgedeckt. (Flückiger 2007, S. 133)

Es bestätigt sich also die Annahme, dass die Tonspur ein eigenständiges narratives Element im Gesamtkonstrukt des Films darstellt. Es werden nicht zwangsweise alle wichtigen Aspekte der Ereignisse durch die Bildebene kommuniziert. Das gewöhnliche Geräusch einer Türe kann so durch feine Abstufungen der klanglichen Eigenschaften ganz unterschiedliche Bedeutungen vermitteln. So wird ein lautes, dumpfes und tiefes Geräusch beim ins Schloss fallen (einer nach dem Geräusch zu schließen sehr schweren Tür) eine andere symbolische Bedeutung vom Filmpublikum zugewiesen bekommen als ein leises Klicken derselben Türe beim Einrasten. Jedoch sind immer der Gesamtkontext sowie die Beziehung von Bild, Ton und Narration zu beachten, da diese Wechselwirkung die Bedeutung der Kanäle ganz wesentlich beeinflussen kann.

4.2 Bild- Ton Beziehungen

4.2.1 Grundlagen

Im Folgenden wird die Wechselwirkung zwischen Bild- und Tonebene behandelt, die Voraussetzungen definiert, und der daraus resultierenden Nutzen hinsichtlich der dramaturgischen Funktion der akustischen Gestaltung erarbeitet.

Schon in den Anfängen des Tonfilms kritisieren die russischen Filmtheoretiker Sergei Eisenstein und Wsewolod Pudowkin die genaue Korrespondenz des Tons mit den Geschehnissen auf der Leinwand, da das bloße Verdoppeln der visuellen Eindrücke eine Redundanz erzeugt. (Flückiger, 2007, 133-135; Görne 2017, S. 17) Es kann laut Eisenstein „nur eine kontrapunktische Verwendung des Tons in Beziehung zum visuellen Montage-Bestandteil [...] neue Möglichkeiten erlauben“. (Görne 2017, S. 17) Die damalige Kritik am Tonfilm ist, auch wenn, wie im Folgenden beschrieben, einige

Aspekte nach heutigem Forschungsstand widerlegbar sind und die Terminologie nicht sehr ausgereift ist, grundsätzlich sehr fortschrittlich. Auch heute gilt, insbesondere in TV- und Mainstream-Filmen, die oft zitierte Prämisse *See a dog, hear a dog*. Gemeint ist damit eben jene redundante Information, die beim Geräuscheinsatz zum Zweck der Verdeutlichung der Geschehnisse auf der Bildebene entsteht. So werden Geräusche laut Flückiger in der Hollywood- Klassik fast ausschließlich synchron mit im Bild sichtbaren Objekten und Handlungen montiert. Bedingt ist dies unter anderem historisch, da in den Anfängen des Tonfilms die technischen Voraussetzungen für die Aufnahme und Wiedergabe noch sehr dürftig und auch auf die menschliche Stimme zugeschnitten waren, so dass geräuschhafte Elemente mit komplexeren Frequenz-Zusammensetzungen oft schlichtweg nicht als das entsprechende Geräusch wahrgenommen wurden. Somit war für den Geräuscheinsatz die auf der Bildebene sichtbare Schallquelle notwendig, um richtig interpretiert bzw. erkannt werden zu können. Als Beispiel nennt Flückiger das Prasseln von Regen und den Applaus, die nur anhand der Zuordnung zur Quelle unterschieden werden konnte. Eine andere Ursache ist die Vereinfachung des Regelwerks, das beim Filmpublikum vorhanden sein muss, um der Handlung eines Films folgen zu können. So ist von den RezipientInnen kaum mehr ein Vorwissen oder eine vorangehende Auseinandersetzung mit dem Medium Film nötig, um die meisten Spielfilme konsumieren zu können, was auch heute noch ein relevanter Aspekt in der Gestaltung von Spielfilmen ist. (Flückiger 2007, S. 135–137)

Die psychoakustische Grundlage für die Wechselwirkung zwischen Bild und Ton stellt die Verknüpfung der visuellen und auditiven Sinnesreize bei gleichzeitiger Wahrnehmung dieser dar. Michael Chion spricht in diesem Zusammenhang vom „audiovisual contract“, also einem *audiovisuellen Vertrag*. (Chion und Gorbman 1994, S. 1) Damit meint er, dass Bild und Ton bei gleichzeitiger Darbietung nicht getrennt voneinander wahrgenommen werden, sondern im kognitiven Prozess der Wahrnehmung zu einer Einheit zusammengefasst werden, wobei die gestaltpsychologischen Vorgänge der Wahrnehmung einen wesentlichen Beitrag leisten. Des Weiteren führt Chion hinsichtlich der Beschreibung dieses Phänomens die Bezeichnung *Synchrese* ein, welche er aus den Wörtern Synchronität und Synthese zusammensetzt. Die Synchrese beschreibt laut Chion die Verschmelzung zwischen einem spezifischen akustischen und optischen Phänomen bei deren zeitgleichem

Auftreten durch die menschliche Wahrnehmung, jenseits von jeglicher rationaler Logik. (Chion und Gorbman 1994, S. 34) Auch nach Thomas Görne haben Bild und Ton dadurch „eine nahezu magische Anziehungskraft“ und die menschliche Wahrnehmung versucht immerzu visuelle und auditive Eindrücke zu verbinden. (Görne 2017, S. 17) Barbara Flückiger bezeichnet diesen Vorgang als intermodale Assoziation.

„Es gehört zu den genuinen Verfahren der menschlichen Wahrnehmung, dass die verschiedenen Empfindungen assoziativ miteinander verknüpft werden. Diese Assoziationsleistung wird [...] als intermodale Assoziation (cross-modal association) bezeichnet.“ Diese ist an die gleichzeitige Darbietung gebunden, jedoch ist laut Flückiger die „zeitliche Übereinstimmung [...] nur ein notwendiger, kein hinreichender Faktor für die visuell-auditive Assoziation.“ (Flückiger 2007, S. 138)

Hinsichtlich der Genauigkeit der zeitlichen Auflösung ergibt sich eine Toleranz durch die schlechtere Auflösung des Sehsinnes. So beträgt das zeitliche Auflösungsvermögen der auditiven Wahrnehmung etwa drei Millisekunden, das Auflösungsvermögen der visuellen mit 20 bis 30 Millisekunden etwa das Zehnfache. Diese vergleichsweise Trägheit des Auges ist auch die Voraussetzung für die Unterteilung des optischen Datenflusses in eine Abfolge von Standbildern, die trotzdem als flüssige Bewegung wahrgenommen werden. Erst um 1930 waren die technischen Möglichkeiten hinreichend, um die Wiedergabe von Bild und Ton innerhalb dieser zeitlichen Toleranz zu halten und so die Synchronität zu gewährleisten. (Flückiger 2007, S. 139–140) Die Intermodale Assoziation ist eine Grundvoraussetzung für das Medium Film (sowie auch für andere audiovisuellen Medien). Optische und akustische Informationen werden ab dem Zeitpunkt der Aufnahme technisch voneinander getrennt. Diese Trennung bleibt über den gesamten Produktionsprozess und auch bei der anschließenden Wiedergabe aufrecht. Die visuellen und auditiven Sinnesreize werden erst vom Filmpublikum bzw. deren kognitiver Leistung bei der Wahrnehmung wieder zu zusammengehörigen Objekten und Ereignissen verknüpft. Welche Auswirkungen die intermodale Assoziation von visuellen und auditiven Sinneseindrücken hat und welche Möglichkeiten für eine sinnvolle Gestaltung im Spielfilm sich dadurch ergeben, wird im Folgenden beschrieben:

Visuelle Sinneseindrücke beeinflussen die Bedeutung eines Hörereignisses oder können diese gänzlich verändern. So kann beispielsweise das Geräusch, das beim Zerdrücken einer Wassermelone entsteht, als sehr unangenehm empfunden werden und vom Hörer als brechende oder zerquetschte Körperteile interpretiert werden, wenn die Bildebene zeitgleich einen Unfall darstellt. Dabei ist nicht zwingend der Realismus des Klanges oder des Geräusches entscheidend, sondern die Gleichzeitigkeit des Geschehens der visuellen und akustischen Sinneseindrücke. (Chion und Gorbman 1994, S. 22–23) Dieser Aspekt wird besonders im Spielfilm von Foley- Artist und im Sounddesign genutzt, um Geräusche zu erstellen, die andernfalls nicht oder nur schwer produzierbar wären. Darüber hinaus ist es dadurch möglich, unterschiedliche Klänge oder Geräusche zu mischen und so hyperrealistische Soundeffekte (häufig wird dabei Begriff *larger than life* verwendet) zu gestalten, die in der Realität nicht zu finden sind, aber zusammen mit einem korrespondierenden Ereignis auf der Bildebene nicht als unglaublich empfunden werden. (Sonnenschein 2001, S. 170) Aber nicht nur beeinflussen visuelle Sinneseindrücke die Hörempfindung, sondern auch umgekehrt. Es kann also auch die Tonspur die Bedeutung von Handlungen oder Ereignissen auf der Bildebene ganz wesentlich beeinflussen bzw. verändern. Sonnenschein vergleicht dies mit der Einblendung von Text oder einer Erzählerstimme aus dem Off über einem Bild, wodurch die Interpretation dieses Bildes unterbewusst immer in Bezug auf die semantische Bedeutung der Worte stattfindet. (Sonnenschein 2001, S. 169) Michael Chion bezeichnet dieses Phänomen als *added value*.

By added value I mean the expressive and informative value with which a sound enriches a given image so as to create the definite impression, in the immediate or remembered experience one has of it, that this information or expression „naturally“ comes from what is seen, and is already contained in the image itself. Added value is what gives the (eminently incorrect) impression that sound is unnecessary, that sound merely duplicates a meaning which in reality it brings about, either all on its own or by discrepancies between it and the image. (Chion und Gorbman 1994, S. 5)

Es entsteht also ein Mehrwert, der durch die alleinige Darstellung der Handlung auf der Bildebene nicht möglich wäre. Damit widerspricht Chion ganz deutlich den Kritikern

des Tonfilmes, nach denen die gleichzeitige Verwendung von Bild und Ton zu vermeiden sei. Im besten Fall entsteht zwischen dem Ton und dem Bild eine *konzeptuelle Resonanz*.

The sound makes us see the image differently, and then this new image makes us hear the sound differently, which in turn makes us see something else in the image, which makes us hear different things in the sound, and so on. (Chion und Gorbman 1994, S. 1)

Der Mehrwert durch die intermodale Assoziation von visuellen und auditiven Sinneseindrücken entsteht demnach reziprok. Dass bedeutet, dass zwischen dem Ton und dem Bild, insbesondere bei synchroner Wahrnehmung der beiden Sinnesmodalitäten eine Wechselwirkung entsteht, die bewirkt, dass sich die Bedeutungen der Sinneseindrücke gegenseitig beeinflussen und verändern.

Fehlt nun einer der beiden Sinnesreize oder beeinflusst der Ton die gewünschte Bedeutung des Bildes (oder umgekehrt) in einer der Wirkung nicht dienlichen, oder sogar schädlichen Art und Weise, geht das Potential der multisensuellen Wahrnehmung des Spielfilms verloren. Hannes Raffaseder fordert für die Mediengestaltung eine Gleichstellung der verschiedenen Sinnesreize (auch in Bezug auf den Geschmackssinn und den Geruchssinn, was jedoch für den Spielfilm in seiner heutigen Form noch nicht relevant ist). Dabei muss diese Gleichstellung „noch keine zwingende Verbesserung bringen, denn nicht das Nebeneinander, sondern ausschließlich das Miteinander unterschiedlicher Sinnesreize kann zu einer nachhaltigen Qualitätssteigerung führen.“ (Raffaseder 2010, S. 19) Es ist also notwendig, die Wechselwirkung mit der Bildebene und die daraus resultierende Bedeutungsmodifikation in der akustischen Gestaltung von Spielfilmen nicht nur zu berücksichtigen (so wie dies auch für die visuelle Gestaltung und deren Einfluss durch und auf die Tonebene gelten muss), sondern die intermodale Assoziation der visuellen und auditiven Sinnesreize in der Konzeption und Produktion einzusetzen. Nur so kann ein Mehrwert hinsichtlich der dramaturgischen Funktion sowie eine Steigerung des immersiven Erlebens vom Filmpublikum erzielt werden. Zu berücksichtigen ist jedoch immer auch der narrative Gesamtkontext des Films sowie die in Kapitel 3 beschriebenen außerfilmischen Konventionen, die ebenso einen Einfluss auf die Wahrnehmung der Bild- und der Tonebene besitzen.

4.2.2 Bedeutung der Bild- Ton Beziehungen für die Narration

Es gibt verschiedene Verhältnisse, mit der die Bedeutungen, die über die Bild- und die Tonebene kommuniziert werden, in Beziehung stehen können, wobei jedes Verhältnis eine andere Wirkung im Gesamtkontext vermittelt. Oft zitierte, ursprünglich aus der Analyse von Stummfilm-Musik stammende Begriffe zur Beschreibung der Bedeutungszusammenhänge zwischen Bild und Ton sind Paraphrase, Polarisierung und Kontrapunkt.

Die *Paraphrase*, die häufigste Konstellation zwischen der Bedeutung von visuellem und auditivem Kommunikationskanal, bezeichnet die Kongruenz von Ereignissen auf der Bild- und Tonebene (see a dog, hear a dog). Die Tonebene spiegelt die vorhandene Information auf der Bildebene und es vermitteln beide Kanäle die gleiche Bedeutung. Aus informationstechnischer Sicht kommt es dabei zu einer Redundanz, wofür die Paraphrase oft kritisiert wird. Die Bedeutung ist unmissverständlich und benötigt keine weitere Interpretationsleistung bei der Wahrnehmung des Filmpublikums. Die kongruente Ton- und Bildsprache hat, sofern mit Bedacht eingesetzt, in der Filmtongestaltung seine Berechtigung, allerdings kann der unüberlegte Einsatz schnell kitschig wirken. Ebenso ist die Paraphrase, wenn über mehrere Kameraeinstellungen und/oder Szenen aufrecht, aufgrund der Neutralität der Bedeutungen ein dramaturgisch ungeeignetes Mittel, um Spannung zu erzeugen und aufrechtzuerhalten. Es fehlt also der Mehrwert hinsichtlich der dramaturgischen Bedeutung. Dieser Mehrwert entsteht demnach insbesondere dann, wenn Bild und Ton unterschiedliche, oder sogar zueinander widersprüchliche Bedeutungen aufweisen. (Görne 2017, S. 236–240)

Als *Polarisierung* wird die Bedeutungsgebung eines für sich neutralen Bild- oder Toninhaltes durch die Information bzw. die Bedeutung, die der andere Kommunikationskanal vermittelt, bezeichnet. Diese Technik wird sowohl im Mainstream-Film als auch in der künstlerisch anspruchsvolleren akustischen Gestaltung oft genutzt, um „neutrale Bilder zu dramatisieren“. (Görne 2017, S. 238) Dabei fällt auf, wie unterschiedlich die Eindrücke des völlig gleichen Bildes bei unterschiedlicher akustischer Untermalung durch verschiedene Musikkrichtungen, Umgebungsgeräusche

oder emotional aufgeladene Soundeffekte sein können. (Raffaseder 2010, S. 277) Die Innenaufnahme ein- und derselben Autofahrt wird so vom Filmpublikum völlig anders interpretiert werden, je nachdem ob ein tiefes, ruhiges Motorengeräusch und leise klassische Musik zu hören sind, oder dröhnender Verkehrslärm, hektisches Hupen in Kombination mit schneller, elektronischer Musik mit einem ausgeprägten Rhythmus.

Für die unterschiedliche bzw. gegensätzliche Bedeutung von Bild- und Tonebene verwendet Hannes Raffaseder den Begriff *Dissonanz*. Eine völlige Widersprüchlichkeit bezeichnet er als *Kontrapunkt*. Im Zuge der Wahrnehmung einer Gegensätzlichkeit von zwei Kommunikationskanälen findet in der Wahrnehmung der RezipientInnen ein aktiver kognitiver Prozess statt, um die Gegensätzlichkeit aufzulösen bzw. zu interpretieren. Werden die Gegensätze zu groß, besteht aber die Gefahr, dass die visuellen und auditiven Sinnesreize nicht mehr miteinander verknüpft werden, also vom Filmpublikum kein Zusammenhang zwischen Bild- und Tonebene hergestellt wird, wodurch der gewonnene Mehrwert verloren geht. (Raffaseder 2010, S. 278) Thomas Görne erklärt diesen Fall anhand der zwischenmenschlichen Kommunikation. Es gibt zum einen die kongruente Kommunikation, bei der verbale und non-verbale Kommunikation miteinander in Einklang stehen bzw. die nonverbale Kommunikation die Bedeutung der verbalen unterstützt. Im umgekehrten Fall, wenn die beiden Kommunikationsarten sich widersprechen, also unterschiedliche Bedeutungen bzw. Informationen vermitteln, kommt es in der Wahrnehmung der EmpfängerInnen zu einer „kognitiven Dissonanz“. Damit wird bei den EmpfängerInnen ein „komisches Gefühl“ ausgelöst, auch wenn sich diese mitunter der Ursache nicht bewusst sind. Es werden innere Spannungen erzeugt und in den EmpfängerInnen wird das Bedürfnis geweckt, diese aufzulösen. „Im täglichen Leben [...] gilt die kognitive Dissonanz als Kommunikationsstörung. Im Film ist sie dagegen ein Schlüssel für spannungsvolles, im wahrsten Sinne des Wortes vielschichtiges Schauspiel“. (Görne 2017, S. 236)

Thomas Görne erweitert das Model zur Beschreibung der Bedeutungszusammenhänge. So ist der Begriff Kontrapunkt für die Beschreibung der Beziehung von Musik zur Bildebene zwar sehr treffend, für die vielfältigen Konstellationen der Bedeutungen zwischen Bild und Ton jedoch nicht ausreichend. Für die verschiedenen Möglichkeiten der Dissonanz bzw. des Kontrapunktes führt er die Bezeichnungen *Parallelisierung* und

Inkongruenz ein. Die Parallelisierung beschreibt den Fall, dass einer der Kommunikationskanäle Information liefert, die durch den anderen nicht vermittelt wird. Das können zum einen diegetische Geräusche oder Musik sein, welche die Zeit und den Ort des Geschehens durch charakteristische Umgebungsgeräusche etablieren, wenn diese Informationen über die Bildebene nicht transportiert werden. Die Tonspur kann so einen wichtigen Beitrag zur Narration liefern. Ein hinsichtlich der Dramaturgie noch wirksamerer Einsatz der Parallelisierung sind nicht diegetische Elemente wie Soundeffekte (wie beispielsweise Symbole oder Keysounds), sowie Erzählerstimmen und innere Monologe aus dem Off, die „ein klassisches Stilmittel zur Verdichtung der Handlung“ darstellen. (Görne 2017, S. 238) Die Bedeutungen der Kommunikationskanäle sind bei der Parallelisierung unterschiedlich, aber nicht widersprüchlich wie bei der seltener eingesetzten Inkongruenz, bei welcher es zu der in Bezug auf die verbale und nonverbale Kommunikation beschriebenen kognitiven Dissonanz kommt, die zu einer Spannungsaufladung führt und eine Auflösung fordert. Ein klassisches, sehr plakatives Beispiel für die Inkongruenz wären idyllische Bilder, zu denen tiefe, basslastige Klänge oder Musik ein drohendes Unheil ankündigen. Insbesondere im Thriller oder im Horrorfilm findet die Inkongruenz oft Verwendung.

Die schon von Raffaseder beschriebene, nicht mehr stattfindende Verknüpfung von Bild und Ton aufgrund einer zu großen Gegensätzlichkeit, bezeichnet Görne als *Irritation*. Diese ist als Stilmittel selten zu finden, und wird meist nur bei diegetischen Klängen, die nicht zum Bild passen, als solche empfunden. Als Beispiel hierfür nennt Görne den vorgezogenen Schnitt der zur der nächsten Szene gehörigen Tonelemente, die dementsprechend nicht zum aktuellen Bild passen. Dabei entsteht in der Wahrnehmung des Filmpublikums eine starke Spannung, die sich aber nach wenigen Sekunden auflöst, sobald die Synchronität zwischen Bild und Ton nach dem Bildschnitt wiederhergestellt ist. Wird die Spannung nicht aufgelöst, empfinden BetrachterInnen die Irritation als technischen Fehler, der die Immersion stören kann. Bei extra- oder metadiegetischen Klängen und Geräuschen entsteht die Irritation im Normalfall nicht. Diese Elemente eröffnen in der Regel einen eigenen Kommunikationskanal (Görne 2017, S. 238–239), deren Informationen zwar die auf der Bildebene stattfindenden Ereignisse beeinflussen, aber deren gegensätzliche Bedeutung durch die vom Filmpublikum erlernten Konventionen eher als Inkongruenz empfunden wird.

Abschließend ist es wichtig zu beachten, dass das Modell der Bedeutungsverhältnisse stark vereinfacht ist. Die Zusammenhänge betreffen nicht nur Bild und Ton, sondern alle drei Kommunikationskanäle. So kann eine Dissonanz auch zwischen der Tonebene und der Narration entstehen, wobei die Bildebene gleichzeitig durch die Narration polarisiert wird. Darüber hinaus lässt sich das Model auch auf die verschiedenen Eigenschaften der Kommunikationskanäle erweitern, wodurch ein komplexes Netzwerk von Zusammenhängen entsteht. Beispielsweise kann die Gestaltung der Lichtsetzung auf der Bildebene die gleiche Stimmung vermitteln wie die Musik, oder deren Wirkung völlig gegensätzlich sein. Die so entstehende Komplexität der Bedeutungszusammenhänge eröffnet eine Vielzahl an Gestaltungsmöglichkeiten beziehungsweise entstehen dadurch sehr vielfältige Positionen, die die Tonspur gegenüber der Bildebene und der Narration einnehmen kann. Wichtig ist dabei letztlich die entstehende Wirkung in Kombination mit den anderen Kommunikationskanälen. (Görne 2017, S. 238)

4.3 Hörperspektive

Einen essentiellen Aspekt für den Stil der Narration, also der Art wie eine Geschichte erzählt wird, stellt die Perspektive dar.

„In the same way that a camera can have a point of view, the perspektive of sound can have a point of audition“. (Sonnenschein 2001, S. 163) Ausgehend vom Begriff Blickpunkt, der sich grundsätzlich auf die visuelle Wahrnehmung bezieht, gibt es zwei mögliche Bedeutungen der Perspektive. Zum einen kann mit dem Blickpunkt der physikalische Standpunkt, also der Ort, von dem aus ein Ereignis räumlich wahrgenommen wird, gemeint sein. Zum anderen kann sich die Perspektive aber auch auf die subjektive Wahrnehmung von Ereignissen aufgrund der individuellen Befindlichkeit sowie der Rahmenbedingungen beziehen. (Flückiger 2007, S. 366–367) Damit ist gemeint, dass manche Ereignisse, die für die meisten Personen völlig irrelevant sind und nicht bewusst wahrgenommen werden, für jemand anderen eine

große Bedeutung haben können, und sich dadurch im Zentrum seiner/ihrer Aufmerksamkeit befinden. Hannes Raffaseder nennt dafür als Beispiel die akustische Umgebung in einer U-Bahn, wo die Durchsage der nächsten Station für eine unter Zeitdruck stehende Person, die dort aussteigen will, völlig anders wahrgenommen wird, als für eine Person, die in ein Gespräch vertieft ist und so die Durchsage möglicherweise gar nicht bewusst wahrnimmt, da ihre Aufmerksamkeit ganz auf das Gespräch bzw. auf den Gesprächspartner gerichtet ist. In der Wahrnehmung der ersten Person wird aber wiederum das Gespräch höchstwahrscheinlich ausgeblendet, bzw. zusammen mit den restlichen Umgebungsgeräuschen in den Hintergrund gerückt sein. (Raffaseder 2010, S. 261)

In Bezug auf die räumlich-physikalische Bedeutung vermittelt der Filmtone laut Flückiger jedoch keine hundert Prozent eindeutige Hörperspektive, wie dies beim Bild der Standpunkt der Kamera vorgibt. Zwar sind die Entfernung vom Mikrofon und der Raumanteil hörbar, allerdings kann die exakte Position bei der Aufnahme mitunter auch von geübten Hörern nicht genau festgestellt werden. Ein Grund dafür ist, dass Mikrofone, ebenso wie das menschliche Gehör, zwar eine höhere Empfindlichkeit bei frontaler Schalleinwirkung aufweisen, aber dennoch Schall aus allen Richtungen aufnehmen. Verstärkt wird dies zusätzlich noch durch unzählige Reflexionen im Raum aus allen Richtungen. Dieser Fakt ist ein wesentliches Unterscheidungsmerkmal zwischen der auditiven und der visuellen Wahrnehmung (bzw. darauf aufbauend, zwischen Mikrofon und Kamera), da diese im Gegensatz ein ganz klar begrenztes Blickfeld aufweist. Ereignisse außerhalb des Blickfeldes können nicht wahrgenommen werden. (Flückiger 2007, S. 366–370) Jedoch ist es sehr wohl möglich, aufgrund der Klangeigenschaften eine Wahrnehmung der Perspektive abzuleiten. (Sonnenschein 2001, S. 163) Die physikalischen Eigenschaften zur Vermittlung von Nähe oder Distanz sind das Frequenzspektrum, die Lautstärke, sowie das Verhältnis der Reflexionen zum Direktschall (siehe Kapitel 2.2). Je weiter die Schallquelle entfernt ist, desto weniger Bässe hat das Schallereignis. Bei noch größerer Entfernung nimmt zusätzlich auch der Höhenanteil ab, ebenso wie der Raumanteil zunimmt. Ein nahes Schallereignis mit einem ausgeprägteren Frequenzspektrum erscheint so, analog zur Bildperspektive, auch größer als ein dünner klingendes, weiter entferntes Schallereignis. (Görne 2017, S. 171) So klingt die O-Ton Aufnahme eines Schauspielers gedämpft, leiser und die

Aufnahme hat mehr Raumanteil, wenn dieser mit dem Rücken zur Kamera gedreht ist, und das Mikrofon aus der gleichen Richtung wie die Kamera aufnimmt. Das Mikrofon nimmt nicht den Direktschall der Stimme, wie bei frontaler Einsprechrichtung, sondern hauptsächlich die Reflexionen durch den Raum auf. Dagegen ist bei einem Mikrofon welches der Schauspieler am Körper trägt, diese Art der Perspektive nicht vorhanden, da sich die Einsprechrichtung nicht abhängig von seiner Position zum Mikrofon verändert. (Sonnenschein 2001, S. 163) In der O-Ton Aufnahme bzw. der anschließenden Mischung der Dialoge wird im Spielfilm zugunsten der Sprachverständlichkeit jedoch oft auf diese Art der räumlich-perspektivischen Gestaltung verzichtet oder diese nur sehr abgeschwächt eingesetzt.

Ganz allgemein ist die Tonperspektive in der Filmtongestaltung meist gleich der Bildperspektive. Der visuelle Raum gibt dabei die Distanzen vor. „Erst durch einen Verstoß gegen diese Regel wird die akustische Perspektive bedeutungsvoll“. (Görne 2017, S. 173) Allerdings bleibt die Tonperspektive oft über eine Szene hinweg gleich, während die Bildperspektive wechselt. Dies geschieht zum einen auch aufgrund der besseren Sprachverständlichkeit, die in der akustischen Gestaltung von Spielfilmen gefordert wird, zum anderen auch um mittels der Tonspur die verschiedenen Bildeinstellungen der Szene als eine Einheit wirken zu lassen. Häufig kommt es dabei zu einer Subjektivierung der Perspektive, wenn also die Bildperspektive nicht mit der Tonperspektive übereinstimmt. Sind die Schauspieler in der Totalen zu sehen, also weiter von der Kamera entfernt, ist bzw. bleibt die Tonperspektive der Stimmen jedoch meistens nahe. So können gewisse akustische Ereignisse durch die Platzierung im Raum sowie in der Mischung bewusst in den Vordergrund, das heißt in den Fokus der Aufmerksamkeit des Filmpublikums gelenkt werden. Kommt es dennoch zu einer Transformation der Tonperspektive, wird diesem Perspektivenwechsel vom Filmpublikum eine Bedeutung zugewiesen, da dadurch andere Schallereignisse in den Vordergrund gerückt werden. Die akustische Perspektive kann also, in Abhängigkeit von den Rahmenbedingungen, hinsichtlich der Dramaturgie bedeutungsvoll sein -oder es durch einen Wechsel der Perspektive werden. (Görne 2017, S. 171–173)

In Zusammenhang mit der Hörperspektive steht auch die Distanz zu einer Schallquelle. Diese ist nach Thomas Görne ebenfalls informationshaltig, und kommuniziert so

verschiedene Bedeutungen. (Görne 2017, S. 171–172) Hannes Raffaseder beschreibt die Informationshaltigkeit in Zusammenhang mit räumlicher Entfernung als die soziale Beziehung zwischen Schallquelle und HörerIn. Die *intime Nähe* ist die Distanz, in der sehr leise körperliche Geräusche wie das Atmen oder der Herzschlag einer anderen Person hörbar sind. (Raffaseder 2010, S. 263–264) Die *intime Distanz* beträgt laut Görne 40–50 cm und ist der Abstand, in der „Körperkontakt möglich und gewollt ist“, wobei das Bedürfnis der intimen Nähe sexueller Natur sein oder aufgrund von Aggression bestehen kann, wobei die intime Distanz einer anderen Person verletzt wird. Insbesondere die Nähe eines akustischen Ereignisses hat eine direkte emotionale Wirkung, was vor allem beim Einsatz von extra- oder metadiegetischen Stimmen (also eines Erzählers oder zur Darstellung von Gedanken der Charaktere) ausgenutzt wird, die sehr nah und oft auch fast flüsternd ins Mikrofon eingesprochen werden und so durch die vermittelte Nähe und metaphorische Größe in die intime Distanz der HörerInnen eindringen. (Görne 2017, 172)

4.4 Konzepte und Techniken

4.4.1 Subjektivierung

Es gibt in der akustischen Gestaltung grundsätzlich zwei völlig unterschiedliche Herangehensweisen. Zum einen gibt es einen technischen Ansatz, bei dem das Ziel eine möglichst originalgetreue Nachbildung der physikalischen Realität ist. Dies ist beispielsweise bei der Aufzeichnung von klassischen Konzerten gewünscht, bei denen eine möglichst unveränderte Aufnahme und Wiedergabe der Performance des Orchesters oder der einzelnen Musiker, bzw. der Konzertsituation erzielt werden soll. Zum anderen geht ein weiterer Ansatz von der menschlichen Wahrnehmung und der daraus folgenden, von den HörerInnen gebildeten, individuellen Realität aus. Die Aufmerksamkeit wird bewusst auf bestimmte Aspekte gelenkt, um diese individuelle

Realität zu beeinflussen. „In der Filmgestaltung verspricht der zweite Ansatz eher Erfolg [...]; es wird genau das zu Gehör gebracht, was wichtig ist“. (Görne 2017, S. 162)

„Gutes Sounddesign strebt daher nicht zwingend nach einer bestmöglichen Reproduktion vermeintlicher Wirklichkeiten. Vielmehr geht es darum, ein (Nach-) Erleben von Ereignissen zu ermöglichen bzw. bestmöglich zu unterstützen.“ (Raffaseder 2010, S. 237) Davon ausgehend und auf die Erkenntnisse hinsichtlich der Bedeutung der Hörperspektive aufbauend, sind Subjektivierungen ein in so gut wie jedem Film eingesetztes Stilmittel der Tongestaltung.

Die indirekte Subjektivierung der Tonspur stellt die akustische Szene aus der Sicht bzw. aus der physikalisch-räumlichen Hörperspektive eines Charakters dar, der sich in der Szene befindet. Die Subjektivierung ist also mit einem Subjekt verknüpft. Das Filmpublikum erlebt die Geschehnisse vom *point of audition* der Figur und kann so in die Diegese eintauchen, die FilmzuseherInnen werden mitten ins Geschehen versetzt. Darüber hinaus ist es möglich, eine indirekte Subjektivierung auch hinsichtlich der subjektiven Wahrnehmung betreffende Hörperspektive zu schaffen und so bestimmte Klänge oder Geräusche für das Filmpublikum in den Vordergrund zu rücken, die für die entsprechende Figur eine besondere Rolle aufgrund des individuellen Gefühlszustandes und der Situation spielen. Durch die indirekte Subjektivierung ist es möglich, zwischen den FilmzuseherInnen und der Figur bzw. deren Erleben eine starke emotionale Bindung aufzubauen, insbesondere wenn auch die Bildebene diese Perspektive einnimmt. (Görne 2017, S. 212–222) Der klassische Fall, bei dem die indirekte Subjektivierung zum Einsatz kommt, ist die Darstellung von Traumzuständen, Wahrnehmungsverzerrungen oder von für die Figur besonders intensiven oder emotional erlebten Momenten. Dabei wird neben der selektiven Auswahl der zu hörenden Elemente auch auf die Asynchronität bzw. die Inkongruenz zwischen Bild und Ton zurückgegriffen. Durch die Verzerrung der Synchronität zwischen dem visuellen und dem auditiven Kommunikationskanal wird „ein feststehendes Wahrnehmungssystem in Frage gestellt. [...] Wahrnehmungsverschiebungen können sich auch im wirklichen Leben durch Auseinanderklaffen von Sinnesempfindungen aus den verschiedenen Modalitäten äußern.“ (Flückiger 2007, S. 395)

Des Weiteren gibt es noch die direkte Subjektivierung. Diese beschreibt die „Subjektivierung ohne Subjekt“. Die Hörperspektive nimmt dabei „die wahrnehmende Position des unsichtbaren Beobachters“ ein, die gleichzusetzen ist mit der Position der FilmzuseherInnen. (Görne 2017, S. 219) „Indem die klassische filmische Narration ihre Erzählinstanz sorgfältig maskiert, gibt sie dem Zuseher die Möglichkeit, eigene Verbindungen zur Diegese zu schaffen, vermittelt ihm das Gefühl selbst Urheber der phantasierten Welt zu sein.“ (Flückiger 2007, S. 371) Die direkte Subjektivierung ist ein Grundprinzip der akustischen Gestaltung und entspricht der Bildmontage, also der „Lenkung der Aufmerksamkeit des Publikums“ durch Schnitte auf andere Einstellungen. Es werden also bewusst nur jene Geräusche auf der Tonspur montiert, die hinsichtlich der Narration und der Dramaturgie sinnvoll sind und vom Filmpublikum gehört werden sollen. (Görne 2017, S. 219)

Gutes Audiodesign überlässt diese Selektion nicht den Rezipienten. Es zählt zu den größten Fehlern, alle denkbaren Ereignisse und Objekte in einem Medienprodukt mit mehr oder weniger passenden Sounds zu vertonen. Wichtig ist, in jedem Moment zu entscheiden, was gehört werden soll und wie dadurch die Wahrnehmung gelenkt wird. Gerade in einer immer lauter, dichter und hektischer werdenden akustischen Umwelt gilt die bekannte Wendung: „Weniger ist mehr“. Für jedes einzelne akustische Ereignis muss dessen Funktion im Gesamtkontext klar sein. Liefert ein Sound keinen nennenswerten Beitrag zu Wirkung und Bedeutung eines Medienproduktes, soll er am besten entfallen. (Raffaseder 2010, S. 237)

In der Regel wird das Filmpublikum die aufmerksameitslenkende Gestaltung der Tonspur durch die vorhandenen filmischen Konventionen und durch den Gesamtkontext mit „großer Wahrscheinlichkeit als 'echter', 'realistischer' einschätzen als den Ton, der versucht die Gesamtheit einer physikalischen Realität darzustellen“. Bedingt ist dies durch die in Kapitel 2 beschriebene Filterung der akustischen Umwelt durch die auditive Wahrnehmung, nach der nur wenige Ereignisse bewusst wahrgenommen werden können. Dazu kommt, dass eine unbearbeitete Originalaufnahme alleine durch die technischen Prozesse bei Aufnahme und Wiedergabe soweit von der tatsächlichen Realität abstrahiert wird, dass diese ohnehin

nicht mehr dieselbe Empfindung auslöst wie bei der Wahrnehmung des ursprünglichen Schallereignisses. (Görne 2017, 219, 221)

Neben der grundsätzlichen, fast immer vorhandenen direkten Subjektivierung, also der aufmerksamkeitslenkenden Gestaltung der Tonspur gibt es auch einen wesentlich seltener zu findenden, extremeren Einsatz der direkten Subjektivierung. Dabei kommuniziert der auditive Kommunikationskanal direkt mit dem Publikum, die Hörperspektive kann eindeutig keiner Person der Diegese zugeordnet werden. Thomas Görne verwendet zur Verdeutlichung dieser Technik eine Szene aus Peter Jacksons „Der Herr der Ringe: Die Gefährten“ (2001). Darin bekämpft eine der Hauptfiguren (Gandalf) in den Minen von Moria einen schier unbesiegbaren Gegner (den Balrog), um seine Gefährten und ihre Mission zu retten. Nach der Bezwingung des Gegners fällt der sich kurz in Sicherheit wählende Gandalf schließlich dem Balrog hinterher in die Tiefe. Die auf der Bildebene eingesetzte Zeitlupe der übrigen Gefährten wird von einem in „der Komplexität und Lautstärke“ (Görne 2017, S. 222) sehr reduzierten und räumlich abstrahierten Sounddesign begleitet. Für die immer leiser werdenden, durch ein Echo in die Länge gezogenen Schreie der übrigen Gefährten lässt sich eindeutig die Hörperspektive der Charaktere ausschließen. Die verhallenden Schreie können als Metapher für das Ableben Gandalfs gedeutet werden und unterstreichen die allgemeine Betroffenheit des Filmpublikums. „Die Sounddesigner verlassen sich offensichtlich darauf, dass das Publikum emotional so intensiv am filmischen Geschehen beteiligt ist, dass es diese Subjektivierung als Metapher der eigenen Erschütterung akzeptiert, und dass sich die emotionale Wirkung dadurch möglichst noch steigert“. (Görne 2017, S. 222)

„Potenziell besteht bei subjektivierenden Eingriffen, die nicht konventionell kodiert sind, die Gefahr, dass der Diskurs störend an die Oberfläche tritt und den Film als Artefakt entlarvt.“ (Flückiger 2007, S. 382) Es kann daher bei übertriebenem Einsatz die Immersion gestört werden. Diese Gefahr wird bei der indirekten Subjektivierung jedoch entschärft. Falls beim Publikum die emotionale Reaktion nicht ausgelöst wird, „so wird durch die indirekte Kommunikation die Emotion als Erleben des Protagonisten erklärt. Das Publikum hat damit die Möglichkeit der distanzierten Wahrnehmung der Szene.“

(Görne 2017, S. 222) Laut Flückiger wird diese Form der Subjektivierung als *Enunziationsmarkierung* bezeichnet. (Flückiger 2007, S. 220)

Im Zusammenhang mit der Subjektivierung zu erwähnen sind die von David Sonnenschein beschriebenen „primary and secondary emotions“. *Primäre Emotionen* beschreiben die von einem Klang oder einem Geräusch ausgelösten Empfindungen einer Figur in der Geschichte, somit einer Person in der filmischen Wirklichkeit. Der montierte Klang oder das montierte Geräusch sind dabei in den meisten Fällen diegetisch. Es ist aber auch denkbar, dass eine Figur von einem extra- oder metadiegetischen Element auf der Tonspur beeinflusst wird, beispielsweise wenn diese sprachlich zum Ausdruck gebrachte Gedanken sind. *Sekundäre Emotionen* sind die Reaktionen des Filmpublikums. Als Gedankenexperiment dieser Unterscheidung nennt Sonnenschein eine mögliche Szene in einem Thriller- oder Horrorfilm, in der sich eine Frau alleine in einem Wald befindet, und auf der Tonspur das verdächtige Knacken eines Astes zu hören ist. Dreht sich die Frau erschrocken um, ist das Knacken ein primär-emotionales Geräusch. Hört sie das Geräusch jedoch nicht und spaziert unbekümmert weiter, löst dies beim Filmpublikum (in Abhängigkeit von dessen emotionaler Bindung zu diesem Charakter) höchstwahrscheinlich eine Emotion in Form von Angst um deren Wohlbefinden aus und das Geräusch vermittelt so eine sekundäre Emotion. (Sonnenschein 2001, S. 181–182)

4.4.2 Die Beziehung zur Diegese

Das vorhergehende Gedankenexperiment von David Sonnenschein stellt auch ein hervorragendes Beispiel für das Verhältnis zwischen der Subjektivierung und der Diegese dar.

Nach Thomas Görne ist eine klare Trennung zwischen diegetischen und nicht-diegetischen Elementen der Tonspur nicht möglich. „Die Beziehung eines Klangobjektes zur filmischen Realität wird darin definiert in Abhängigkeit von der *Wahrscheinlichkeit*, mit der es als diegetisch, meta- oder extradiegetisch empfunden

wird.“ (Görne 2017, S. 231) Die diskrete Einteilung von Geräuschen hinsichtlich ihrer Zuordnung zur Diegese ist also nicht möglich, vielmehr gibt diese die Grenzen vor, innerhalb welcher die Sounds als diegetisch empfunden werden. Die Bedeutung von Klängen und Geräuschen in Bezug auf die Narration und Dramaturgie lässt sich auch anhand der Einordnung innerhalb dieser Grenzen erklären. Laut Thomas Görne können akustische Ereignisse, welchen vom Filmpublikum keine Semantik höherer Ordnung zugewiesen wird, nur in der filmischen Realität vorhanden bzw. diegetisch sein. Dies sind zumeist Geräusche, die lediglich dazu dienen, eine Szene normal und glaubhaft wirken zu lassen, bzw. „die filmische Welt zum Leben zu erwecken“. (Görne 2017, S. 231) Allerdings ist eine Abgrenzung selbst hier schwierig, da die meisten Umgebungsgeräusche von den FilmzuseherInnen weitere Bedeutungen zugewiesen bekommen als bloß jene über ihre Entstehung. So werden diese, auch durch den narrativen Kontext, mit bestimmten Orten oder der erzählten Zeit in Verbindung gebracht. Deshalb sollten auch die Umgebungsgeräusche sorgfältig ausgewählt werden. Sind die Klänge oder Geräusche nicht mehr eindeutig der Diegese zuzuordnen, werden diese demnach automatisch bedeutungshaltig. Zwar ist es möglich, dass ihnen von den FilmzuseherInnen nicht zwangsläufig eine Bedeutung aufgrund außerfilmischer, kultureller, historischer oder universeller Natur zugewiesen wird, jedoch werden diese, aufgrund innerfilmischer Konvention (also durch deren Verwendung im Gesamtkontext der Narration) immer bedeutungsvoll. Diese Relevanz steigert sich, die akustischen Ereignisse werden demnach bedeutungsvoller mit der Wahrscheinlichkeit, mit der sie nicht als diegetisch empfunden werden. „Sie öffnen zusätzliche Kommunikationskanäle im Film, und sie können unmittelbare emotionale Wirkung entfalten“. (Görne 2017, S. 231)

Ein dramaturgisch interessantes Gestaltungsmittel stellen insbesondere Klänge und Geräusche dar, die zwischen den beiden Extremen (also diegetischen und der nicht-diegetischen Sounds) liegen, sich somit nicht genau zuordnen lassen. Das können zum einen Soundeffekte wie Signale oder Symbole sein, oder „semantisch überladene“ Klänge und Geräusche. (Görne 2017, S. 231) Als *semantisch überladen* bezeichnet Görne „ein Klangobjekt oder ein audiovisuelles Objekt, das in der Tongestaltung mehr kommunikative Inhalte transportiert als im alltäglichen Leben“. (Görne 2017, S. 147) Als Beispiele bieten sich filmische Stereotypen wie Faustschläge und Schüsse an. Diese

werden, wie erwähnt, in der Regel deutlich überzeichnet, da sie in der Realität wesentlich unscheinbarer klingen als im Film und so die gewünschte Wirkung auf das Filmpublikum ausbleiben würde. Der Grund für die Entstehung dieser filmischen Stereotypen liegt unter anderem in der Produktionslogistik, da diese Geräusche in der Regel nicht Teil der O-Ton Aufnahmen waren und erst im Nachhinein der Tonspur hinzugefügt wurden, wodurch bereits sehr früh die Möglichkeit gegeben war, diese meist bedeutungsvollen Geräusche durch Bearbeitung eindrucksvoller klingen zu lassen. (Flückiger 2007, S. 178–182) Semantisch überladene Klänge oder Geräusche können daher nicht gänzlich der Diegese zugeordnet werden, da sie nicht einer physikalisch-realistischen Abbildung der filmischen Wirklichkeit entsprechen.

Ebenfalls nicht eindeutig zuzuordnen sind Off- Screen Sounds, da diese grundsätzlich sowohl diegetisch als auch nicht-diegetisch sein können. (Sonnenschein 2001, S. 153) Michael Chion bezeichnet solche Klänge oder Geräusche ohne sichtbare Schallquelle im Bild als akusmatisch. Dabei kann entweder eine sichtbare Schallquelle erst durch einen Wechsel der Bildperspektive aus dem Bild verschwinden, wodurch das akustische Ereignis klar als diegetisch eingeordnet wird (sofern das Verschwinden nicht durch den Schnitt auf eine andere Szene bedingt ist - mehr dazu unter *Transformation* weiter unten). Die Schallquelle kann aber auch am Beginn gänzlich verborgen bleiben. Dabei wird beim Filmpublikum die Aufmerksamkeit erhöht bzw. Neugierde geweckt, da durch den Drang der Wahrnehmung, ein Schallereignis mit einem Objekt zu verknüpfen, Fragen über den Ursprung des Klangs oder des Geräusches auftauchen. Ein solches akusmatisches Element erzeugt eine Spannung, die durch das Preisgeben der Schallquelle aufgelöst werden kann, was Chion als *De-akusmatisierung* bezeichnet. (Chion und Gorbman 1994, S. 71–72) Thomas Görne spricht in diesem Zusammenhang von *semantischer Ambivalenz*, wobei er aber nicht von der Unsichtbarkeit der Schallquelle im Bild ausgeht. Die semantische Ambivalenz bezieht sich auf die Mehrdeutigkeit, das heißt ein Klang oder Geräusch kann gleichzeitig unterschiedliche Bedeutungen kommunizieren.

Weiter gefasst kann man auch solche Klangobjekte als semantisch ambivalent bezeichnen, deren Bedeutung sich nicht eindeutig entschlüsseln lässt, [...] Das Kalkül dabei ist natürlich, dass die durch unklare Klänge befeuerte Vorstellung von dem was sein *könnte* emotional stärker wirkt als die explizierte Erklärung was

ist; die Unverständlichkeit [...] ist dabei ein dramaturgisches Werkzeug. (Görne 2017, S. 154–155)

Ein Stilmittel der Tongestaltung ist die Veränderung die Beziehung eines akustischen Ereignisses zur Diegese, also deren Transformation. Dabei wird ein diegetisches Element der Tonspur über den zeitlichen Verlauf zu einem extra- oder metadiegetischen Element, oder umgekehrt. Thomas Görne nennt als Beispiel das Vorlesen eines Briefes, wobei anfänglich der sprechende Charakter im Bild zu sehen ist, dann jedoch andere Szenen oder Ereignisse dargestellt werden, wodurch die Stimme aus dem Off zu hören ist und so als Erzählerstimme fungiert. Häufig wird diese Technik auch bei der Filmmusik verwendet, die durch eine anfänglich im Bild sichtbare Schallquelle als Teil der Atmo etabliert wird, aber die Musik durch den Schnitt auf einen anderen Schauplatz extradiegetisch wird und sich die Bedeutung aus der Sicht des Publikums verändert. Die Musik kann nun die Handlung kommentieren oder Gefühle der Charaktere ausdrücken. (Görne 2017, S. 232) Es lässt sich daher festhalten, dass der Grad der Diegese ein wesentliches Stilmittel der akustischen Gestaltung in Hinsicht auf die Dramaturgie des Films darstellt.

4.4.3 Konzepte der akustischen Gestaltung

Nun besteht nicht nur ein direkter Zusammenhang zwischen dem Grad der Diegese und einer Bedeutung höherer Ordnung. Auch zwischen der Beziehung der verschiedenen Elemente der Tonspur zur Diegese und dem Grad der Subjektivierung lässt sich ein Verhältnis herstellen. Demnach wird einem Sounddesign, dessen Klänge und Geräusche eindeutig nicht der Diegese zugeordnet werden können, vom Filmpublikum nicht nur eine starke Bedeutung zugewiesen, es handelt sich auch um eine starke Subjektivierung.

In Abhängigkeit vom Verhältnis zwischen dem Grad der Diegese und der Subjektivierung beschreibt Thomas Görne vier mögliche Ebenen, auf denen die Tonspur eine Funktion übernehmen kann (Görne 2017, 234). Die Gestaltung durch als

rein als diegetisch empfundene Elemente, die als „getreue, dokumentarische Abbildung einer fiktionalen Wirklichkeit“ fungieren, bezeichnet er als den „*dokumentarischen oder naturalistischen Ton*“. Dagegen basiert der „*aufmerksamkeitslenkende Ton*“ auf der direkten Subjektivierung und kann mitunter realistischer wirken als die unveränderte Abbildung der physikalischen Umgebung. Dabei werden Klänge und Geräusche auf der Tonspur montiert, die eine Funktion im Kontext der Narration ausüben. „Bei einer gelungenen aufmerksamkeitslenkenden Tongestaltung ist es für den Hörer u.U. nicht zu unterscheiden, ob seine Aufmerksamkeit gelenkt wurde oder ob er selbst intentional, bewusst, seine Aufmerksamkeit auf einen Teil der Szene gelegt hat.“ (Görne 2017, S. 234) Die dritte Ebene ist der „*überwirkliche Ton*“ und entfernt sich zunehmend von der Diegese. Die Subjektivierung ist also deutlicher, ambivalente und semantisch überladene Elemente sowie symbolische Klänge und Geräusche stehen in Kontrast mit diegetischen Elementen und werden bewusst „aus der Realität der filmischen Welt herausgehoben und als bedeutungsvoll erklärt“. Schließlich beschreibt der „*surreale oder mystische Ton*“ eine Gestaltungsweise, in der fast ausschließlich extra- oder metadiegetische Sounds verwendet werden, die als Metaphern für die Umschreibung von Gefühlszuständen der Charaktere oder des Publikums, oder von Vorgängen bzw. Ereignissen der Handlung. Dabei handelt es sich um eine starke, oft indirekte Subjektivierung, die wie schon im vorangehenden Abschnitt beschrieben zur Symbolisierung von emotionalen Ausnahmezuständen sowie Wahrnehmungs- oder Bewusstseinsveränderungen eingesetzt werden. Oft lässt sich die Tongestaltung ganzer Filme anhand einer dieser Ebenen einordnen, jedoch ist der Wechsel zwischen den Ebenen ein effizientes „Mittel, um Spannungsbögen zu erzeugen“. (Görne 2017, S. 235)

In Verbindung mit diesen Ebenen beschreibt Thomas Görne auch vier Konzepte der akustischen Gestaltung im Spielfilm. Der „*Dokumentarische Realismus*“ oder „*Naturalismus*“ bedient sich in der Regel rein diegetischen Elementen und vermeidet Inkongruenzen und Irritationen zur Bildebene. Zu beachten ist laut Görne, dass paradoxer Weise nicht jene Gestaltung der Diegese, die der menschlichen Wahrnehmung der physikalischen Realität am nächsten kommt, als authentisch empfunden wird, sondern eine meist monophone, technisch-qualitativ nicht perfekte Aufzeichnung mit einem gewissen Lo-Fi Charakter. Bedingt ist dies mitunter von der

durch Mediennutzung entstandenen Konventionen, da die Tonspuren in dokumentarischen Werken meist Aufzeichnungen in Form monophoner O-Ton Aufnahmen sind, die oft mit technisch begrenzten Mitteln umgesetzt wurden. Der dokumentarische Realismus wird oft mit einer belanglosen Tongestaltung assoziiert und dafür kritisiert, keinen informativen Mehrwert zu bieten (siehe „Paraphrase“ im Abschnitt 4.2.2) Laut Görne kann jedoch gerade die naturalistische Gestaltung der Tonspur durch ihre Reduziertheit und ihre Realitäts-Behauptung unter Umständen emotional intensiver wirken als ein ausgefeiltes, technisch aufwändiges Sounddesign mit vielen Effekten und mehreren Surroundkanälen. (Görne 2017, S. 245) So eingesetzt kann der dokumentarische Realismus ungeschöntes Abbild der Realität vermitteln. In Horrorfilmen wird manchmal versucht, durch diese Art der akustischen Gestaltung die Illusion zu vermitteln, dass es sich bei den Inhalten tatsächlich um dokumentarische, also „echte“ Aufnahmen der Geschehnisse handelt, um so noch erschreckender zu sein.

Aufbauend auf dem aufmerksamkeitslenkenden Ton ist der „*Quasi-Realismus*“ bzw. der „*Quasi-Naturalismus*“ das zweite Konzept der akustischen Gestaltung. Wie beim dokumentarischen Realismus werden Inkongruenzen mit der Bildebene eher vermieden, allerdings finden sich sehr wohl zeichenhafte und symbolische Klänge mit einer metaphorischen Bedeutung. Der Quasi-Realismus ist das wohl am häufigsten zum Einsatz kommende Gestaltungskonzept. (Görne 2017, S. 247) Hannes Raffaseder beschreibt die dabei zum Einsatz kommende intentionale Verwendung von Klängen und Geräuschen zur direkten Subjektivierung als „selektiven Geräuscheinsatz“. Der Großteil der Umgebungsgeräusche wird dabei bewusst reduziert, um die Aufmerksamkeit auf einige wenige Soundeffekte in den Vordergrund zu rücken. (Raffaseder 2010, S. 245) Die Informationsdichte wird gering gehalten, da „der Versuch, alle im Bild sichtbaren oder zum Kontext der Handlung gehörenden [...] Schallquellen auch hörbar zu machen“ (Raffaseder 2010, S. 254) zu einer nicht mehr zu differenzierenden Klangkulisse führt, die zum einen die Sprachverständlichkeit negativ beeinflusst, sowie dazu, dass die einzelnen Geräusche an Bedeutung verlieren. Natürlich kann, in Abhängigkeit vom narrativen Kontext auch bei manchen Szenen eine hohe Informationsdichte gefordert sein. Wichtig ist dabei, dass dennoch die Transparenz nicht verloren geht, sowie dass diese nur an gewissen Stellen gezielt eingesetzt wird, da

sonst deren Wirkung aufgrund der fehlenden Dynamik nicht zur Geltung kommt. (Raffaseder 2010, S. 254)

Das dritte Konzept ist der *Expressionismus*, welcher ebenfalls auf dem aufmerksamkeitslenkenden Ton aufbaut, das Sounddesign aber oft auf die überwirkliche oder surreal-mystische Ebene wechselt. Die verwendeten Elemente der Tonspur sind grundsätzlich bedeutungsvoll, Polarisierung, Parrallelisierung und auch Inkongruenzen sind häufig eingesetzte Stilmittel. Ein expressionistisches Sounddesign ist öfter zu finden, je weniger sich die Gestaltung der Filme am Mainstream orientiert. (Görne 2017, S. 251)

Schließlich beschreibt Görne mit dem *Impressionismus* das vierte Gestaltungskonzept, wobei die indirekte Subjektivierung ein häufig eingesetztes Stilmittel darstellt. Der Impressionismus stellt die filmische Wirklichkeit aus der Hörperspektive der Charaktere dar, die Elemente der Tonspur sind dadurch metadiegetisch. Die Aufmerksamkeit des Publikums wird auf die, für den Charakter aufgrund dessen körperlicher und mentaler Befindlichkeit bedeutungsvollen Klänge oder Geräusche gelenkt. Zu finden sind alle möglichen Konstellationen in der Beziehung zwischen Bild und Ton, „die typische Funktion des Tons ist aber die Parallelisierung, indem metadiegetischer Ton mehr oder weniger diegetische Bilder ergänzt“. In der Regel gibt es nur wenige Filme mit einer rein impressionistischen Tongestaltung. (Görne 2017, S. 255)

Die Beziehung zur Bildebene, der in Zusammenhang mit der Hörperspektive stehende Grad der Subjektivierung und der Diegese, sowie die darauf aufbauenden Ebenen und Konzepte der akustischen Gestaltung stellen das theoretische Grundgerüst für die Überprüfung der aufgestellten Thesen dar.

5 Empirische Analyse

5.1 Methodik, Definition der Parameter

Anhand der Filmanalysen soll untersucht werden, was die Funktion der Tonspur in Bezug auf die Dramaturgie im fiktionalen Spielfilm ist, welche Position diese im Verhältnis zu den anderen Kommunikationskanälen (also der Bildebene und der Narration) einnimmt und auch, welche Bedeutung die im vorangehenden Kapitel erarbeiteten Konzepte und Techniken der Tongestaltung haben.

Nach David Sonnenschein muss in der Tongestaltung ein umfassendes Verständnis der Handlung, der Thematik sowie der zu vermittelnden Bedeutung von Ereignissen in der Geschichte des zu gestaltenden Films vorhanden sein, wenn ein Mehrwert in Bezug auf Narration und Dramaturgie eines Spielfilms geschaffen werden soll. (Sonnenschein 2001, S. 173–174) Um den Sachverhalt untersuchen zu können, muss dieses Verständnis auch die Voraussetzung für die Analyse sein, welche somit anhand zweier Aspekte erfolgt. Zum einen müssen die Narration und die Dramaturgie der Filme analysiert werden.

„In the fiction film, narration is the process whereby the film's *syuzhet* and style interact in the course of cuing and channeling the spectator's construction of the fabula.“ (Bordwell 1985, S. 53) Mit *syuzhet* meint Bordwell den Plot des Films, mit *style* die ästhetische Materialisierung des Plots, also die Ton- und die Bildsprache. „Erst im Stil konkretisiert sich also das Erzählen durch den systematischen Einsatz der spezifischen Ausdrucksmittel eines Mediums“. (Keutzer et al. 2014, S. 195) Der Plot des Films beschreibt die Anordnung und Strukturierung der Ereignisse zum Erzählen der Geschichte. Der ursprünglich lineare Zeitverlauf einer Geschichte wird im fiktionalen Film zumeist verändert, bestimmte Teile verkürzt, ausgelassen oder neu arrangiert, zugunsten einer interessanteren Dramaturgie. Eine Analyse des Plots beantwortet daher die Frage „Wie wird erzählt?“ (Keutzer et al. 2014, S. 195) und gibt Aufschluss über die Erzählstruktur, also die Anordnung und Zusammenhänge der Ereignisse. Durch die

Plot-Analyse können also die für die Narration wesentlichen Abschnitte, Höhe-, Tief-, und Wendepunkte ermittelt werden.

Zum anderen werden darauf aufbauend die Elemente der Tonspur (also Sprache, Geräusche und Musik) in diesen Schlüsselmomenten der Filme analysiert. Außerdem wird deren Beziehung zum Bild, deren Bedeutung im narrativen Kontext und welche der in Kapitel 4 erarbeiteten Techniken und Gestaltungsmittel dabei zum Einsatz kommen, untersucht. Die anschließende Interpretation der Ergebnisse der Analyse zeigt auf, welche Rolle die Tonspur hinsichtlich der Narration und Dramaturgie einnimmt und welchen Mehrwert diese schafft.

Der Analyse unterzogen werden im folgenden Stanley Kubricks Sci-Fi Abenteuer „2001: A Space Odyssey“ (1968), David Lynchs „Mulholland Drive“ (2001), Lars von Triers „Melancholia“ (Lars von Trier, 2011) und der mittellange Spielfilm „Der Dunkle Reigen“ (Bücheler, 2017). Mit dieser Filmauswahl wurden bewusst unterschiedliche Produktionsjahre abgedeckt und verschiedene Genres ausgewählt, sowie klassische Hollywood-Mainstreamproduktionen vermieden, bei welchen der Spielraum der Tongestaltung meist sehr begrenzt ist (Flückiger 2007, S. 136), um so die Möglichkeiten der akustischen Gestaltung ausloten bzw. erforschen zu können.

5.2 „Melancholia“ (Lars von Trier, 2011)

„Melancholia“ (2011) ist ein Drama/Sci-Fi Film vom dänischen Autor und Regisseur Lars von Trier, in dem zwei Schwestern die letzten Tage vor dem Weltuntergang erleben. Die vorrangigen Themen sind Depression, die Angst vor dem Tod, sowie der Umgang der Figuren mit dem unausweichlichen Ende. Der Aufbau des Films ist an dem einer Oper angelehnt, was eine Parallele zu Richard Wagners „Tristan und Isolde“ darstellt.

5.2.1 Narration und Dramaturgie

Der erste Teil des Films [00:00:26 - 00:08:30] ist als Prolog beschrieben. Als musikalische Untermalung dient das Vorspiel zu Wagners Opernwerk, das Ende des Films wird in diesen acht Minuten für die FilmzuseherInnen bereits vorweggenommen. Es handelt sich um in extremer Zeitlupe gezeigte im Orbit aufeinanderprallende Planeten, zusammen mit surreal anmutenden, ebenfalls in Zeitlupe abgespielten Bildern, welche als Sinnbilder für den Verlauf der Handlung und die Gefühlszustände von Justine (Kirsten Dunst) dienen.

Darauf folgt der erste Akt „Justine“ [00:08:40 - 01:09:00], der chronologisch die Ereignisse der Hochzeit zwischen Justine und Michael (Alexander Skarsgård) auf dem Anwesen von Justines Schwager John (Kiefer Sutherland) und ihrer Schwester Claire (Charlotte Gainsbourg) darstellt. Nach der verspäteten Ankunft des Brautpaares feiert die Hochzeitgesellschaft zunächst ausgelassen. Im Angesicht der im Prolog gezeigten Apokalypse erscheint die pompöse Feier jedoch bereits belanglos und aufgesetzt, was zudem noch durch die Oberflächlichkeit der Justine nahestehenden, aber wie es scheint, aus Eigeninteresse handelnden Hochzeitgäste unterstrichen wird. Im Verlauf des Abends zeigt sich nach und nach Justines Depression, und obwohl sie anfangs versucht, die glückliche Fassade aufrecht zu erhalten, sind letztlich auch Claires Bemühungen, die Hochzeit zu retten, vergebens. Schließlich scheitert nicht nur die Hochzeit, auch verliert sie ihren Job und Ihren Mann Michael.

Der zweite Akt [01:09:00-02:09:30] trägt den Namen „Claire“. Darin werden insbesondere die Gefühlswelt und die unterschiedenen Persönlichkeiten von Claire und Justine beschreiben sowie deren Verhalten im Angesicht des herannahenden Endes. Zu Beginn des zweiten Aktes hat sich Justines Depression verschlimmert und sie zieht auf das Anwesen von John und ihrer Schwester Claire, welche fortan ihre kranke Schwester pflegt. Der herannahende Planet „Melancholia“ wird nun thematisiert und Claire äußert ihre Angst, dass dieser die Erde treffen könnte. John dagegen ist fasziniert von dem außergewöhnlichen Ereignis und versichert ihr, dass es wissenschaftlichen Berechnungen zufolge zu keiner Kollision mit der Erde kommen wird. Justine reagiert mit Gleichgültigkeit auf die Befürchtungen ihrer Schwester und eröffnet ihr zudem ihre

Vorahnung, dass Melancholia die Erde zerstören wird. Justine scheint es zunehmend besser zu gehen, je näher der Planet kommt. Die Anzeichen für die Annäherung werden schließlich deutlicher. Die Pferde werden unruhig, der Strom fällt aus. In der Nacht als der Planet die Erde passiert, beobachten die Charaktere das Naturschauspiel und Claire ist schließlich erleichtert. Am nächsten Morgen bemerkt John jedoch, dass seine Annahme falsch war und sich Melancholia wieder der Erde nähert, woraufhin er sich das Leben nimmt. Claire, die Johns Verschwinden bemerkt, stellt ebenfalls fest, dass der Planet wieder näherkommt. Sie reagiert panisch und findet ihren Gatten schließlich tot im Pferdestall auf. In diesem Moment bleibt Claire aber erstaunlich ruhig und versucht Johns Tod vor ihrem Sohn und Justine zu verbergen. Schließlich entgleitet der sonst immer sehr vernünftigen Claire die Kontrolle über ihre Angst. Sie unternimmt einen erfolglosen und im Angesicht der nahenden Apokalypse sinnlos scheinenden Versuch, mit ihrem Sohn vom Anwesen zu fliehen. Justine dagegen wirkt immer gefasster, am Ende ist schließlich sie es, die ihren Neffen beruhigt und ihm Unterstützung bietet. Die Charaktereigenschaften von Justine, Claire und John wandeln sich im zweiten Akt unter den Umständen des unausweichlichen Todes in das Gegenteil. John, der in seiner Rolle als Vater und Ehemann seiner Familie normalerweise immer Unterstützung bietet, handelt egoistisch, nimmt sich selbst das Leben und überlässt seine Liebsten ihrem Schicksal. Die sonst immer im Sinne der Vernunft handelnde Claire verliert im Angesicht des Todes die Fassung, wohingegen die depressive, ansonsten gleichgültige Justine ruhig und gelassen wirkt und sich um ihren Neffen kümmert.

5.2.2 Tongestaltung

Die Gestaltung der Tonspur in „Melancholia“ lässt sich als quasi-realistisch einordnen. Die Atmos sind auf das Wesentliche reduziert, die wenigen Geräusche sind nur dann zu hören, wenn diese für die Charaktere eine Bedeutung haben. Durch diese Reduziertheit wird der Fokus der Aufmerksamkeit im Allgemeinen auf die Dialoge gelenkt, sowie Platz für die Filmmusik geschaffen. Sprache und Geräusche entstammen fast ausschließlich

der Diegese, selbst das an sich abstrakte Geräusch des herannahenden Planeten, das später noch genauer behandelt wird, kann dieser zugeordnet werden.

In einem starken Kontrast dazu steht die Musik. Dabei handelt es sich um Exzerpte der Ouvertüre von Richard Wagners Oper „Tristan und Isolde“, interpretiert vom The City of Prague Philharmonic Orchestra. Im ersten Teil, dem Prolog des Films, ist die Ouvertüre in voller Länge zu hören und ist dabei das in der Tonmischung priorisierte Element. Später werden immer wieder kurze Ausschnitte des Stücks verwendet, um bestimmte Ereignisse zu umschreiben und zu deuten. Es handelt sich um klassische, extra-diegetische Filmmusik, deren Funktion es nicht nur ist, den Verlauf der Handlung zu kommentieren, sondern auch diese voranzutreiben. Ganz wesentlich sind die, im Verlauf des Films sehr oft eingesetzten ersten Takte des Vorspiels, welche als musikalisches Leitmotiv gesehen werden können (siehe Abbildung 12).

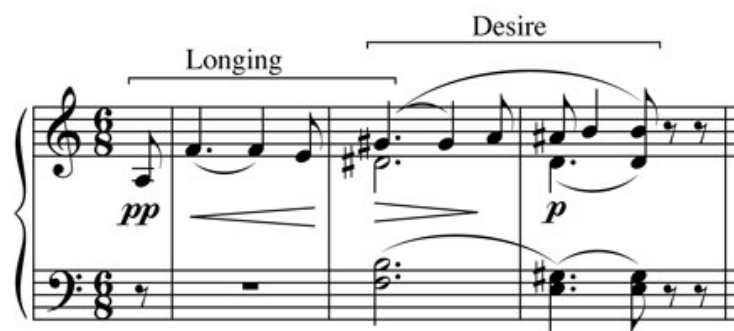


Abbildung 12: Takt 1-3 der Ouvertüre von Richard Wagners „Tristan und Isolde“ (Tristan and Isolde)

Das Intervall zwischen den ersten beiden Tönen (A und F) ist die Sext, welches David Sonnenschein als „besänftigend, gebrechlich, empfindlich und traurig“ bezeichnet. Das Intervall zum darauffolgenden e ist die kleine Sekunde, die „angespannt, unbehaglich und mysteriös“ klingt. (Sonnenschein 2001, S. 121) Der darauffolgende Vierklang wird auch als „Tristan Akkord“ bezeichnet. Dieser entspricht jedoch nicht im gewöhnlichen Sinne einem Akkord, da nicht nur eine, sondern gleich zwei Dissonanzen enthalten sind, die bei den HörerInnen eine Spannung erzeugen und nach einer Auflösung fordern. Der nächste Akkord löst eine dieser Dissonanzen auf, die zweite bleibt jedoch bestehen, wodurch den HörerInnen die komplette Auflösung der Spannung verwehrt bleibt. (Magee 2002, S. 208)

Eine nicht unwesentliche Information in Bezug auf die Handlung, die nicht über die Musik transportiert wird, ist das tiefe, konstante Grollen, das zu hören ist, wenn der Planet sich nahe der Erde befindet. Ob dieses nach physikalischen Gesetzmäßigkeiten so klingen oder überhaupt zu hören wäre, sei dahingestellt. Es handelt sich bei diesem Grollen jedenfalls um eines der wenigen semantisch überzeichneten Geräusche im Film.

Im Folgenden werden einige Schlüsselszenen des Filmes genauer beleuchtet:

Nach der in einem Desaster endenden Hochzeit reiten Justine und Claire mit ihren Pferden aus [01:07:48]. Dabei verweigert Justines Pferd Abraham sich über eine Brücke zu bewegen, die vom Anwesen wegführt, und Claire wirft durch die erzwungene Pause einen Blick in den Himmel. Dabei bemerkt sie, dass in einem Sternbild ein Stern fehlt („Der rote Stern fehlt im Skorpion. Antares ist nicht zu sehen“). Später stellt sich heraus, dass die Sicht auf den Stern durch den herannahenden Planeten Melancholia verdeckt ist.

Eingeleitet wird die Szene mit dem oben beschriebenen Leitmotiv, in diesem Fall in Solo gespielt von einem Holzblasinstrument. Das tiefe Register in Kombination mit der dunklen Klangfarbe der Holzblasinstrumente vermittelt etwas Bedrohliches und wirkt wie eine unheilvolle Ankündigung. Dazu hört man die Hufe der heranreitenden Pferde. Nach dem Verstummen des Leitmotivs [01:08:00] ist außer dem Hufgetrampel und dem Wiehern des sich sträubenden Pferdes sowie Justines Rufe zum Antreiben des Pferdes nichts zu hören. Während Justine nach ihren erfolglosen Versuchen, das Pferd über die Brücke zu bewegen, innehält und in den Himmel aufschaut [01:08:20], verstummen die Umgebungsgeräusche kurz. Es setzen leise und tiefe Streicherklänge ein. Die Auseinandersetzung Justines mit ihrem Pferd wird durch die naturalistische, diegetische Vertonung der Bildinhalte nüchtern und realistisch dargestellt. Die Sprache und Geräusche vermitteln also keine zusätzlichen Informationen oder Emotionen, außer dem Wiehern des Pferdes und Justines Unmutsäußerungen. Die Musik allerdings polarisiert das Geschehen, der Einsatz des spannungsgeladenen Leitmotivs und dessen reduzierter Instrumentation verleihen der Weigerung des Pferdes etwas Mystisches, das Ereignis erhält einen symbolischen Charakter. Die an sich objektive Feststellung Justines erhält durch die polarisierende Wirkung der Musikuntermalung eine höhere Bedeutung. Die Musik kündigt das kommende Unheil an.

Im zweiten Akt kommt es zu einer ähnlichen Situation. Wieder reiten die beiden Schwestern aus, wieder weigert sich das Pferd über die Brücke zu reiten [01:22:15]. Das einleitende Leitmotiv fehlt hier, es ist nur das Hufgetrappel und das Wiehern des Pferdes zu hören. Justine reagiert noch weniger verständnisvoll und versucht Abraham mit Schlägen der Gerte zum Weitergehen zu zwingen [01:22:22]. Die Schläge der Gerte, die Geräusche der Hufe, Justines Unmutsäußerungen und die gelegentlichen Zwischenrufe von Claire sind die einzigen hörbaren Elemente und entspringen durchwegs der filmischen Wirklichkeit. Abraham aber lässt sich nicht erweichen und die Bildebene zeigt mit wackeliger Kamera, die vermutlich die Sicht einer der Protagonistinnen darstellen soll, den Himmel, wo zum ersten Mal im Film (vom Prolog abgesehen) der Planet „Melancholia“ deutlich zu sehen ist [01:23:00], was Justine zynisch mit „Da ist er, euer Vorbeiflug“ kommentiert. Kurz davor, als das Pferd in die Knie geht, setzt wieder Filmmusik ein, während die Geräusche der Umgebung reduziert werden. Zuerst sind zurückhaltende Holzbläser zu hören, die den Ereignissen etwas Geheimnisvolles, Mystisches verleihen. Beim Anblick des Himmels kommentiert schließlich eine tiefe Melodie der Streicher unheilvoll und drohend die Anwesenheit von Melancholia. Wieder polarisiert die Musik und nimmt vorweg, dass der herannahende Planet nichts Gutes bedeutet.

In einer weiteren, für die Veranschaulichung der Charaktere und der Entwicklung wichtigen Szene, begibt sich Justine bei Nacht, wie magisch vom Planeten angezogen auf eine Waldlichtung. Claire, die Justine nachschleicht, offenbart sich der Anblick von Justine, die sich nackt im Licht von Melancholia badet [01:26:00]. Die Musik beginnt wiederum mit dem dezenten Einsatz des Leitmotivs, das sich aufbaut und anschließend, durch Holzbläser und Streicher gespielte, wie Frage und Antwort erklingende Sekunden eine starke Spannung aufbauen, während Claire ihrer Schwester folgt [01:26:45]. Die musikalische Spannung bzw. Dissonanz löst sich genau in dem Moment auf, in der die nackte, dem Planeten sehnsüchtig entgegenblickende Justine sichtbar wird. Die Auflösung deutet die Erlösung an, die Justine in der Vorstellung des herannahenden Weltuntergangs findet. Gleichzeitig eröffnet die Szene einen weiteren Interpretationsspielraum des Films, nach dem die gesamte Handlung eine Neuinszenierung der Oper Richard Wagners ist, in der Justine den verwundeten Tristan verkörpert, während Melancholia die ihm zu Hilfe eilende Isolde darstellt.

[01:45:57] Während Johns Entdeckung des erneut herannahenden, diesmal Kurs auf die Erde nehmenden Planeten schläft Claire nichts ahnend auf der Terrasse ein. Als sie wieder aufwacht, ist John verschwunden und nicht mehr aufzufinden. Auf der Bildebene sind die Bilder der einschlafenden und der aufwachenden Claire hintereinander montiert. Der Zeitsprung, während sich John das Leben nimmt, ist auf der Tonebene subtil, aber wirksam in Form eines harten Schnittes der Atmo umgesetzt [01:46:30], die sonst über die ganze Szene hinweg gleichbleibt oder durch kontinuierliche Übergänge gekennzeichnet ist.

In der Nacht, in der Melancholia an der Erde vorbeifliegt, beobachten die Charaktere das Naturspektakel von der Terrasse aus. Bevor der Planet am Horizont aufgeht, kündigt bereits ein dumpfes Grollen seine Anwesenheit an [01:39:00]. Das Geräusch wird bereits im Prolog etabliert, wo es die Bilder der kollidierenden Planeten untermalt, wodurch für das Filmpublikum klar ist, dass es sich dabei um den herannahenden Melancholia handelt. Es ist fortan über die ganze Szene zu hören. Erst nachdem sich Claire und John sicher sind, dass sich Melancholia wieder entfernt, endet mit dieser Szene auch das Geräusch des Planeten. Später im Film, nachdem Claire realisiert, dass sich Melancholia erneut auf die Erde zubewegt, kommt es zum letzten Mal zum Versuch, die Brücke ins Dorf zu überqueren. Diesmal verweigert jedoch nicht ein Pferd die Überquerung, sondern das Golfmobil, mit dem Claire gemeinsam mit ihrem Sohn vom Anwesen flüchten will [01:57:00]. Von dieser Stelle an ist das Grollen bis zum Ende des Films zuerst leiser, dann kaum merklich lauter werdend, stetig zu hören. Es kann als Metapher für die Unaufhaltsamkeit von Melancholia gedeutet werden sowie für das unausweichliche Ende. Das Geräusch des Planeten kann im Film als Keysound beschrieben werden, wobei sowohl die Beschreibung Flückigers als auch jene von Hannes Raffaseder zutrifft (siehe Kapitel 3 - Bedeutung durch filmspezifische Verknüpfungen).

In der fast vier minütigen Endsequenz [02:05:43-02:09:30], in der Justine, Claire und deren Sohn unter einer Zelt-ähnlichen Holzkonstruktion auf einer Wiese den bevorstehenden Weltuntergang erwarten, sind die natürlichen Umgebungsgeräusche verschwunden, das Grollen des Planeten ist lauter und nimmt den gesamten Platz in der Atmo ein. Dazu kommt ein überzeichnetes Windgeräusch, welches das unmittelbare Ende verdeutlicht [02:07:55]. Die Dynamik der Musik ist zu Beginn sehr leise, sie steigert

sich aber im Verlauf der Endsequenz kontinuierlich, bis sie schließlich synchron mit einem tiefen Donnern, ausgelöst durch die Kollision der beiden Planeten endet, was im Prolog bereits ähnlich gelöst ist.

5.3 „Mulholland Drive“ (David Lynch, 2001)

„Mulholland Drive“ (2001) ist ein Thriller/Mystery-Drama vom US-amerikanischen Autor David Lynch, in dem das Scheitern einer jungen Schauspielerin dargestellt wird. Im Film werden die Abgründe Hollywoods bzw. zerstörte Leben durch das Filmbusiness thematisiert, wobei der Film eine Vielzahl möglicher Interpretationen zulässt. So sind im Film eine ganze Reihe von unaufgelösten Rätseln, Handlungsstränge, die nicht im klassischen Sinne erklärt bzw. aufgelöst werden, sowie Symbole und Ereignisse, über deren Bedeutung das Filmpublikum nur mutmaßen kann, zu finden.

5.3.1 Narration und Dramaturgie

Eine weit verbreitete Interpretation der Handlung ist jene, nach welcher der Hauptteil des Films in einem Traum der gescheiterten Schauspielerin Diane Selwyn besteht, um die von ihr in Auftrag gegebene Ermordung ihrer Geliebten Camilla zu verarbeiten.

„Mulholland Drive“ ist grundsätzlich in zwei Teile gegliedert, wobei der erste unterschiedliche Handlungsstränge beinhaltet, deren Abfolge chronologisch zu sein scheint.

Dieser erste Abschnitt, der einen großen Teil des Films ausmacht, erzählt zum einen das Aufeinandertreffen von Betty (Naomi Watts), die nach dem Gewinn eines Tanzwettbewerbes nach Hollywood kommt, um Schauspielerin zu werden und einer jungen Frau (Laura Harring). Diese entgeht durch einen mehr oder weniger

unglücklichen Zufall durch einen Autounfall einem Mordanschlag, verliert dabei aber ihr Gedächtnis.

Ohne erkennbaren Zusammenhang sind am Anfang des Films zwei Männer in einem Diner (Winkie's) zu sehen [00:12:00], wobei einer davon durch einen Traum panische Angst vor einer Person mit einem fürchterlichen Gesicht hat, die sich hinter dem Diner befinden soll. Der andere überredet ihn nach draußen zugehen, um sich seiner Angst zu stellen. Als die beiden zur Rückseite des Diners gehen, kommt hinter einer Mauer tatsächlich das fürchterliche Gesicht zum Vorschein, woraufhin der ängstliche Mann zusammenbricht.

Daraufhin wird eine Reihe aufeinanderfolgender Telefonate gezeigt, bei denen allerdings nur die Identität der vier Gesprächspartner sichtbar ist. Diese Person wird in einer späteren Szene als Mr. Rock angesprochen. Er teilt jemandem mit, dass „die Kleine“ (mit der vermutlich die Frau ohne Identität gemeint ist) immer noch vermisst wird.

Betty findet die andere Frau, die sich inspiriert durch ein Filmplakat in Folge Rita nennt, in der Wohnung ihrer Tante auf, in der sich diese, vom Mordversuch und vom Unfall traumatisiert, unerlaubterweise aufhält bzw. versteckt. Obwohl Ritas Deckung auffliegt, nach der sie, entgegen Bettys Annahme, nicht zu den Freundinnen ihrer Tante zählt, ist Betty begeistert von der Idee, Ritas unbekannte Identität aufzuklären. Im Zuge dieses Unterfangens öffnen die beiden Ritas Handtasche, in der sich eine große Menge Bargeld und ein mysteriöser, blauer Schlüssel befinden. Durch eine wiedergekehrte Erinnerung Ritas können die beiden schließlich den Mullholand Drive als Unfallort identifizieren. Ebenfalls erinnert sich Rita, als sie den Namen einer Kellnerin im Winkie's sieht, an den Namen Diane Selwyn. Die beiden finden Dianas Adresse heraus, um festzustellen, ob es sich bei ihr um Rita handelt.

Betty kann durch ihre Tante an einem Casting-Termin teilnehmen, in dem sie so herausragend spielt, dass die Casting-Agentin sie zu einem weiteren Casting mitnimmt, bei dem sie auf Regisseur Adam Kesher trifft. Sie verlässt das Casting jedoch vorzeitig, da sie mit Rita verabredet ist, um Diane Selwyn zu suchen. Nach einem Gespräch mit Dianas Nachbarin können die beiden aber ausschließen, dass es sich bei Diane Selwyn um Rita handelt und erfahren außerdem, dass sich Diane offensichtlich vor der Polizei

versteckt. Sie steigen in Dianas Apartment ein und finden diese tot auf. In der folgenden Nacht schlafen Betty und Rita im gleichen Bett und werden miteinander intim. Später, mitten in der Nacht, wacht Rita aus einem Altraumtraum auf und überredet Betty mit ihr auszugehen. Sie suchen den Club Silencio auf, dessen Namen Rita vorhin im Traum laut ausgesprochen hat. Dort sehen die beiden eine Aufführung, die mit den Worten angekündigt wird, dass alles nur Illusion ist. Eine Sängerin bewegt die Lippen zu einem Song, der von einem Tonband abgespielt wird und bricht mitten im Stück zusammen. Der Song läuft weiter. Betty, sichtlich schockiert, findet in ihrer Tasche eine blaue Box, die zu dem seltsamen Schlüssel zu passen scheint. Wieder in der Wohnung angekommen, wollen die beiden die Box öffnen, aber auf einmal ist Betty verschwunden. Also Rita die Box öffnet, ist auch sie verschwunden, und Bettys Tante findet sich in der Wohnung wieder.

Ein zweiter Handlungsstrang im ersten Teil des Films erzählt die Geschichte von Adam Keshner, der sowohl vor beruflichen als auch privaten Problemen steht. Die Hauptrolle von Keshners Film „The Sylvia North Story“ muss neu besetzt werden. In einem Meeting mit seinem Manager, seinen Produzenten und den ominösen Castiglione-Brüdern wird Keshner von den Brüdern zu verstehen gegeben, dass er die Rolle mit einer Schauspielerin Namens Camilla Rhodes zu besetzen hat. Keshner widersetzt sich der Aufforderung und demoliert vor Wut das Auto der Castiglione-Brüder. Daraufhin folgt eine Szene, in der Luigi Castiglione mit Mr. Rock, der offenbar sein Vorgesetzter ist, die weitere Vorgehensweise bespricht.

Ohne zunächst erkennbaren Zusammenhang folgt eine Szene in einem Büro, in der ein Auftragskiller einen Mann tötet, um ein mysteriöses schwarzes Telefonbuch an sich zu nehmen.

Zuhause angekommen findet Keshner seine Frau mit dem Poolboy vor, bekommt einen Wutanfall und wird schließlich von diesem aus dem Haus geworfen. Er versteckt sich in einem billigen Hotelzimmer vor den Castiglione-Brüdern bzw. deren Handlangern, bis ihm vom Vermieter mitgeteilt wird, dass seine Kreditkarte gesperrt wurde. Seine Assistentin teilt ihm wenig später telefonisch mit, dass seine Konten leer sind und dass ein Mann namens „der Cowboy“ mit ihm sprechen will. Keshner stimmt dem Treffen zu und sucht den Cowboy auf seiner Ranch auf, wo ihn dieser auffordert, der Besetzung

der Hautrolle mit Camilla Rhodes zuzustimmen. Im darauffolgenden Casting, bei dem es auch zur Begegnung mit Betty kommt, gibt der Regisseur dem Druck des Cowboys nach und wählt Camilla Rhodes aus.

Der zweite Abschnitt folgt ohne Unterbrechung zum ersten und beginnt mit einer Szene, in der der Cowboy zur schlafenden Diane Selwyn sagt, dass es Zeit sei aufzuwachen. In der nächsten Szene wird Diane, bei der es sich offenbar um dieselbe Person wie Betty handelt, vom Klopfen ihrer Nachbarin geweckt, die sich einige ihrer Sachen abholt und ihr mitteilt, dass die Polizei wiederholt nach ihr gefragt hat. Wieder alleine, stottert die sichtlich verwirrte Diane, dass Camilla wieder bei ihr ist. In einer Rückblende eröffnet die Frau, die im ersten Abschnitt Rita dargestellt hatte und nun Camilla ist, Diane, dass sie nicht mehr mit ihr schlafen will. Diane fragt, ob ‚er‘ der Grund sei. Die nächste Szene spielt auf einem Filmset, in der sich Adam Kesher und Camilla küssen, was Diane, sichtlich verletzt, beobachtet. Am Schluss der Rückblende ist noch ein Streit zwischen den Frauen zu sehen.

In einem Telefonat wird Diane von Camilla auf eine Feier in Keshers Haus eingeladen, welches sich am Mullholand Drive befindet. Diane telefoniert mit jenem Telefon unter einer roten Lampe, das zu Beginn des Filmes zu sehen ist [00:18:00]. Dianas Taxi hält an eben jener Stelle, in der „Ritas“ gescheiterte Ermordung und der Unfall stattfindet. Diane fragt ebenso wie Rita im ersten Teil, warum sie an dieser Stelle halten. Camilla kommt zum Taxi und begleitet Diane zum Haus. Auf der Feier befindet sich Keshers Mutter, die im ersten Teil die Vermieterin von Bettys Tante dargestellt hat. Diane erzählt, dass ihre Tante gestorben ist und ihr etwas Geld hinterlassen hat, worauf sie nach Hollywood gekommen ist. Am Set von ‚The Sylvia North Story‘ hat sie Camilla kennengelernt, die zwar statt ihr die Hauptrolle bekommen hat, in der Folge aber ihre Freundin wurde und ihr einige kleinere Rollen verschafft hat. Im Verlauf der Feier küsst Camilla vor Dianas Augen Adam Kesher sowie eine weitere Frau. Am Ende eröffnen Adam und Camilla ihren Gästen, dass sie heiraten werden, was Diane unter Tränen beobachtet.

In der nächsten Szene übergibt Diane dem Auftragskiller aus dem ersten Teil im Winkie's ein Foto von Camilla sowie die Bezahlung für den Auftrag. Auf dem Schild der Bedienung ist der Name Betty zu sehen. Der Killer zeigt Diane einen blauen Schlüssel, der zuvor schon in Dianas Apartment zu sehen war und teilt ihr mit, dass sie den

Schlüssel am vereinbarten Treffpunkt finden wird, sobald der Auftrag ausgeführt ist. Auf Dianas Frage, was der Schlüssel denn öffne, lacht er nur. An der Bar sieht Diane den ängstlichen Mann, der zu Beginn des Films vom fürchterlichen Gesicht geträumt hat. Die nächste Einstellung zeigt eben diesen Mann mit dem fürchterlichen Gesicht als Obdachlosen, der hinter dem Diner haust. Dieser hält die mysteriöse blaue Box in den Händen, aus der in Miniaturform das alte, hysterisch lachende Ehepaar kommt, das am Anfang des Films Bettys Ankunft in L.A. begleitet hat.

In der Schlusszene ist Diane in ihrem Appartement, starrt entsetzt den blauen Schlüssel an, der vor ihr auf dem Couchtisch liegt. Diane halluziniert von dem lachenden Ehepaar. Es klopft an der Türe und Blaulicht ist durchs Fenster zu sehen. Die lachenden Alten stehen nun in voller Größe im Appartement. Diane gerät immer mehr in Panik und erschießt sich in ihrem Schlafzimmer.

5.3.2 Tongestaltung

In „Mullholand Drive“ handelt es sich um eine expressionistische Gestaltung der Tonspur. Die Basis ist ein quasi-naturalistisches Sounddesign, zu dem jedoch abstrakte Klänge, Geräusche und Musik eingesetzt werden, um die Ereignisse im Bild zu paraphrasieren oder zu polarisieren. Dadurch werden die Ereignisse und Bilder als informationshaltig bzw. bedeutungsvoll erklärt.

Weite Teile des Films sind von tiefen, mysteriös und bedrohlich klingenden Streichern untermalt. Der Einsatz dieser ist oft sehr subtil, manchmal aber auch deutlicher. Ein musikalisches Thema, das immer wieder im Film zu hören ist, ist die Tonfolge a– g– a– f– g. Dazu kommen, vermutlich durch Klangsynthese erzeugte oder synthetisch verfremdete Klänge und Geräusche, die wie die Streicher ebenfalls deutlich nicht der Diegese zuzuordnen sind. Es ist unter anderem dieser Fremdheit der Klangfarbe geschuldet, dass durch die Klänge und Geräusche immer ein unheimliches, beklemmendes Gefühl erzeugt wird.

Die Titelsequenz [0:02:50], welche die Fahrt auf dem Mullholand Drive vor dem Unfall zeigt, ist mit einer solchen Musik unterlegt, wobei hier noch recht deutlich eine Harmonie erkennbar ist. Die Streicher klingen, mystisch und unheilvoll. Als das Auto anhält, verstummt die Musik. Allerdings wird das Zücken der Waffe, die auf Rita gerichtet wird, durch die Musik akzentuiert.

Dianes Ankunft am Flughafen in L.A. [00:18:05] ist mit einem Thema unterlegt, das als euphorisch, engelsgleich beschrieben werden könnte. Die musikalische Harmonie hält aber nur kurz an, es mischen sich spannungsgeladene Intervalle bzw. Dissonanzen in das Thema. Die Bild-Ton-Beziehung ist somit eine Transformation von einer Paraphrasierung, bzw. positiven Polarisierung der Bildebene zu negativen Polarisierung. Die Musik könnte hier für die Unschuld der jüngeren Diane stehen, mit der diese nach L.A. kommt; die Transformation dafür, dass sie diese durch die Stadt bzw. das Filmbusiness verliert. Bei Dianes Ankunft im Apartment ihrer Tante [00:20:15] ist das Thema noch einmal zu hören. Wieder hält es nur kurz.

Die beiden Szenen, in der Mr. Rock zu sehen ist [00:17:06 und 00:35:00], sind beide durch die gleichen, langgezogenen und leicht pulsierenden elektronischen Klänge untermalt, die keine wiederkehrende Melodie und keinen klaren Rhythmus aufweisen. Ähnlich gestaltet ist das Treffen zwischen Kesher und dem Cowboy [01:05:00]. Unter die Windgeräusche mischen sich tiefe Streicher, die den geheimnisvollen Cowboy ankündigen. Dessen Auftreten und das Verschwinden wird durch eine brizzelnde, an- und ausgehende Lampe signalisiert.

Viele Ereignisse im Bild werden auch durch die Kommentierung, also durch den Einsatz von abstrakten Geräuschen und langanhaltenden Klängen polarisiert, also für bedeutungsvoll erklärt. Bemerkenswert ist, dass eine Unterscheidung von Effekten und Musik schwierig ist, da die Musik auch geräuschhafte Komponenten beinhaltet. Auch hat sie oft keine Melodie, keinen Rhythmus oder einen strukturierten Aufbau im musikalischen Sinne. So ergeben sich abstrakte Gemische aus verschiedenen Klängen und Geräuschen, welche die Ereignisse im Bild kommentieren.

In einer Szene am Anfang des Films, in der sich die beiden Männer im Winkie's treffen [00:12:00], ist das Sounddesign zu Beginn naturalistisch und diegetisch. Es ist nur der

Dialog zwischen den zwei Männern und die Umgebungsgeräusche zu hören. Ab [00:13:00] beginnt der ängstliche der beiden von seinem Albtraum zu erzählen. Die Umgebungsgeräusche faden langsam aus, ein nicht erkennbares, abstrakt klingendes Geräusch setzt, langsam lauter werdend, ein. Das abstrakte Element beinhaltet klangliche Komponenten, es ist aber keine Melodie und kein Rhythmus erkennbar. Durch die ansteigende Lautstärke und die Fremdartigkeit der Klangfarbe erzeugt es ein unheimliches Gefühl und baut Spannung auf. Als die Beschreibung des Albtraumes vorbei ist, lässt die Lautstärke kurz nach [00:14:40]. Der andere Mann beschließt nachzusehen und der verängstigte sieht ihm hinterher, die Lautstärke steigt wieder kurz an. Es ist eine bedrohliche, aufsteigende Melodie im Klangteppich zu vernehmen, deren Klangfarbe einem tiefen Blasinstrument ähnelt [00:14:50]. Diese Akzentuierung des Bildes vermittelt das Gefühl, dass der ängstliche Mann eben diesen Anblick aus seinem Traum bereits kennt. Die Sounds dienen dabei als Metapher für dessen anschwellende Furcht. Als die beiden nach draußen gehen, steigt die Lautstärke nochmals an, es sind wieder Umgebungsgeräusche zu hören, jedoch verhindert der abstrakte Klangteppich die Orientierung. Auf dem Weg zur Hinterseite des Diners ist wiederholt die tiefe Melodie zu hören sowie komplexe, an- und abschwellende Geräusche, deren Ursachen wiederum nicht genau identifiziert werden können. Die gesamte Lautstärke nimmt kurz vor dem Auftauchen des Gesichtes wieder etwas ab, der folgende Schockmoment [00:16:35] ist aber zusätzlich durch ein besonders lautes, unangenehmes Geräusch verstärkt, welches, nachdem der verängstigte Mann sein Bewusstsein verliert, bestehen bleibt. Die gesamte Szene verläuft von einer quasi-realistischen Darstellung der Ereignisse hin zu einer immer stärker werdenden Subjektivierung und Abstrahierung, wobei die Hörperspektive nicht eindeutig zuordbar ist. Die Dynamik der Lautstärke und der Grad der Subjektivierung erzeugt einen deutlichen Spannungsbogen.

Bemerkenswert ist, dass die fremdartigen Geräusche und die Musik im zweiten Teil des Films [01:57:00] bis zum Erteilen des Mordauftrages [02:17:00] sehr viel subtiler eingesetzt werden als im ersten Teil des Filmes. Die Elemente der Tonspur entspringen hauptsächlich der filmischen Wirklichkeit und die Klangfarbe ist weniger fremdartig. Die Musik besteht aus Songs, die sich eher an populärer Musik orientieren, da sie durch den prägnanten Rhythmus eines Schlagzeugs bzw. eines Drumcomputers sowie durch elektrische Gitarren charakterisiert sind. Sehr subtil sind doch ab und zu Soundeffekte

zu hören. Ein Beispiel dafür ist das Bild des blauen Schlüssels in Dianas Appartement [01:59:05], das durch die Akzentuierung durch die Tonspur bedeutungsvoll wird, obwohl die Wichtigkeit des Schlüssels, als Symbol für Camillas Tod, noch nicht durch die Handlung erklärt wurde. Der Unterschied in Musik und im Sounddesign kann als Hinweis interpretiert werden, dass der zweite Teil des Films die Realität darstellt und der erste nur eine Illusion.

5.4 „2001: A Space Odyssey“ (Stanley Kubrick, 1968)

„2001: A Space Odyssey“ ist ein Sci-Fi Film von Regisseur, Autor und Produzent Stanley Kubrick. Das Drehbuch stammt von Kubrick selbst und dem britischen Autor Arthur C. Clarke, der nach Veröffentlichung des Films auch einen gleichnamigen Roman herausbrachte. „2001: A Space Odyssey“ ist ein Umriss der menschlichen Entwicklung und Evolution und wirft die Frage auf, wohin diese führt. Es werden Themen wie technischer bzw. wissenschaftlicher Fortschritt, künstliche Intelligenz, die Suche nach außerirdischem Leben und nicht zuletzt auch Religion bzw. die Frage nach einer höheren Instanz behandelt. Hinsichtlich der Narration und Dramaturgie ignoriert Kubrick in „2001: A Space Odyssey“ viele filmische bzw. genrespezifische Konventionen, bedient sich bei der Erzählung der Handlung nicht hauptsächlich des Dialoges, sondern überlässt der Bild- und Tonsprache viele Ausdrucksmöglichkeiten. Kubrick eröffnet den FilmzuseherInnen nicht nur einen großen Interpretationsspielraum, er lässt auch das Ende offen. Der Stil der Erzählung wirkt objektiv und distanziert, es werden nur wenige Spannungsbögen im Sinne der klassischen Hollywood-Erzählweise aufgebaut und aufgelöst.

5.4.1 Narration und Dramaturgie

Grundsätzlich ist der Film in fünf Abschnitte bzw. Kapitel gegliedert. Der erste [00:04:35-00:19:50] beschreibt die evolutionären Schritte vom Primaten zum Homo erectus. Eine Gruppe von Menschenaffen in einer vorsteinzeitlichen Landschaft steht in Konflikt mit anderen Tieren um die verfügbare Nahrung sowie mit einer anderen Gruppe ihrer Artgenossen um eine Wasserquelle. Eines Morgens findet die Gruppe zu ihrer Verwunderung einen scheinbar funktionslosen Monolithen vor ihrer Höhle. Nach dem Erscheinen des Monolithen lernt einer der Primaten, einen Knochen als Werkzeug bzw. als Waffe zu verwenden, was es der Gruppe ermöglicht, sowohl andere Tiere zu jagen als auch ihre konkurrierenden Artgenossen in die Flucht zu schlagen.

Der zweite Abschnitt [00:19:50-00:54:40] handelt von Dr. Haywood Floyd (William Sylvester), der im Jahr 1999 zum Mond reist, um den rätselhaften Fund eines Monolithen zu untersuchen. Die Bildmontage beim Übergang der Kapitel vermittelt den Eindruck, dass die zu Beginn von Kapitel 2 sichtbare Raumfahrt die direkte Folge der erlernten Nutzung von Gegenständen als Waffe bzw. Werkzeug ist. Sie startet mit Bildern von Raumstationen im Weltall und dem Spaceshuttle, das Dr. Haywood Floyd zu einer Raumstation bringt. Dort angekommen, muss dieser den Fund vor WissenschaftlerInnen anderer Nationen verbergen, da die Informationen darüber offensichtlich nicht nach außen dringen dürfen. Von der Raumstation reist Dr. Haywood Floyd weiter zum Mond, wo er in einer Konferenz den mit dem Fund befassten Personen die weitere Vorgehensweise erläutert. Schließlich wird ihm die Ausgrabungsstätte gezeigt, wo er den Monolithen berührt. Daraufhin ist ein lautes Signal zu hören, das offenbar vom Monolithen ausgesendet wird.

Der dritte Abschnitt [00:54:40-01:27:22] ist mit „Jupiter Mission - 18 Months later“ untertitelt und handelt von den Astronauten Dave Bowman (Keir Dullea) und Dr. Frank Poole (Gary Lockwood), die an Bord des Raumschiffes „Discovery“ zusammen mit dem sprechenden Supercomputer HAL 9000 und einem dreiköpfigen, in einen künstlichen Dauerschlaf versetzten Forschungsteam zum Jupiter reisen, der als Ziel des vom Monolithen ausgesendeten Signals identifiziert wurde. Der eigentliche Sinn der Mission ist den beiden Astronauten im Gegensatz zu dem auf den Namen HAL hörenden, sehr

menschlich wirkenden Computer nicht bekannt. Die Rahmenbedingungen werden den ZuseherInnen durch ein Fernsehinterview, das die beiden Astronauten ansehen, erläutert. Nach einem Gespräch zwischen Bowman und HAL, in dem der Computer Zweifel an der Mission aufgrund der geheimnisvollen Umständen in der Vorbereitung der Expedition äußert, gibt der eigentlich vor Fehlern gefeierte Computer an, dass ein Bauteil des Raumschiffes ausgetauscht werden müsse, da dies in absehbarer Zeit ausfallen werde. Nach Rücksprache mit der Mission Control auf der Erde, die den Astronauten das OK gibt, das Teil austauschen zu dürfen, aber auch anmerkt, dass ihren Berechnungen zufolge die Information HALs über den Ausfall nicht korrekt zu sein scheint, führt Bowman den Ausbau durch, wofür er sich aus dem Schiff begeben muss. Da bei näherer Begutachtung alles in Ordnung scheint, beschließen die beiden Astronauten, auf den Vorschlag HALs, das Bauteil vorerst wieder einzubauen, einzugehen und auf den Ausfall zu warten. Die beiden Astronauten haben kein gutes Gefühl und ziehen sich in eine Raumkapsel zurück, um von HAL ungehört die bei Ihnen aufgetauchten Zweifel hinsichtlich der Intentionen des Computers zu besprechen. Sie beschließen, HAL abzuschalten, falls sich seine Information als falsch herausstellen sollte.

Kapitel 4 trägt den Namen „Intermission“ [01:27:22-01:57:00]. Poole begibt sich nach draußen, um die funktionierende Einheit wieder einzubauen. Als er aus der Raumkapsel steigt, die ihn zur Stelle des Umbaus am Raumschiff bringt, wendet sich die, wie es scheint, ein Eigenleben entwickelnde Raumkapsel gegen ihn und schleudert ihn ins All. Bowman beobachtet die Ereignisse vom Cockpit der Discovery aus, sieht jedoch nicht, was genau passiert. Auch HAL gibt vor, nichts zu wissen. Bowman steigt ebenfalls in eine Raumkapsel und eilt seinem vermutlich bereits toten Kollegen hinterher. Als er nach der Bergung zur Discovery zurückkehrt, verweigert ihm HAL den Zugang zum Schiff. Er eröffnet Bowman, dass er das Gespräch zwischen ihm und seinem Kollegen zwar nicht gehört, aber ihre Lippen gelesen hat. In Folge dessen hat er Maßnahmen ergriffen, um seine Abschaltung zu verhindern. Als Grund gibt er an, die Mission nicht zu gefährden zu dürfen. Bowman schafft es schließlich doch mit Gewalt an Board der Discovery zu kommen und entfernt anschließend die Module, die für HALs Funktionen verantwortlich sind. Während der Abschaltung artikuliert HAL sehr emotional und menschlich wirkend seine Angst vor dem „Tod“. Danach startet auf einem Bildschirm

eine Videobotschaft, die Bowman die wahre Intention ihrer Mission erläutert, nämlich den Ursprung des Monolithen zu finden.

Kapitel 5, „Jupiter and beyond the infinite“ [01:57:00-02:20:30] beschreibt die Ankunft Bowmans, der die Mission fortsetzt, beim Planeten Jupiter, wo er auf einen, im All schwebenden Monolithen stößt. Die folgende, durch visuelle Effekte verfremdete, psychodelisch anmutende Sequenz, in der bunte Formen und verfremdete Landschaften an Bowman vorbeiziehen, könnte die Reise durch ein Wurmloch darstellen, nach der er schließlich an einem unbekannten Ort in Raum und Zeit landet. In einem surreal wirkenden Schlafzimmer mit spätbarocker Einrichtung sieht der Astronaut, aus seiner Raumkapsel blickend, sein offensichtlich gealtertes Ego. Danach ist aus der Sicht dieses gealterten Bowman ein noch älterer Mann im Nebenraum zu sehen, der einen Bademantel trägt und eine Mahlzeit einnimmt. Der junge Bowman inklusive der Raumkapsel ist verschwunden. Die Person im Nebenraum hat ebenfalls die Gesichtszüge und Frisur des Astronauten, jedoch wiederum um einige Jahre älter. Als dieser aufsteht und in das Schlafzimmer geht, ist der gealterte Bowman im Raumanzug verschwunden, stattdessen liegt ein uralter Greis im Bett. Zuletzt verschwindet auch der alte Bowman im Bademantel, der alte Greis sieht stattdessen einen Monolithen vor seinem Bett und versucht, diesen zu berühren. Die nächste Einstellung zeigt das Bett, in dem in einer Fruchtblase ein übergroßes Baby liegt. In der letzten Szene befindet sich das riesige Baby samt Fruchtblase zwischen Erde und Mond im All und beobachtet die Erde.

5.4.2 Tongestaltung

Die Gestaltung der Tonspur unterscheidet sich in „2001: A Space Odyssey“ durch den sparsamen Einsatz der Dialoge von vielen anderen Produktionen. Weite Teile des Films und ganzer Szenen beinhalten keine Sprache, sondern Musik, Geräusche bzw. sehr reduzierte Atmos oder auch Stille. Die Dialoge sind auf das Notwendigste reduziert, der Filmmusik und den Geräuschen kommt so eine tragende Rolle zu.

Die Musik stellt ein wesentliches Element der Narration im Film dar, da sie nicht nur einen größeren zeitlichen Anteil als die Sprache einnimmt, sondern eine wesentliche Bedeutung transportiert, die durch die Dialoge oder das Bild nicht kommuniziert wird. Die Funktion der Filmmusik ist somit nicht, wie in vielen anderen Filmen, Emotionen der Charaktere auszudrücken oder Szenen zu dramatisieren. Es kommen bereits existente Werke aus klassischer und neuer Musik zum Einsatz.

„2001: A Space Odyssey“ beginnt mit einer etwa drei minütigen Intro-Sequenz, in der das Stück „Atmosphères“ vom ungarischen Avantgarde Komponisten György Ligeti zu hören ist. Das Bild bleibt währenddessen schwarz. Die Komposition weist keine klaren Strukturen auf, keinen klaren Anfang und kein klares Ende im musikalischen Sinne.

In der darauffolgenden Titel-Sequenz [00:02:58] ist zu den Bildern des in einer Linie über der Erde aufgehenden Mondes und der Sonne, der Beginn des ersten Abschnittes aus „Also sprach Zarathustra“ von Richard Strauß zu hören. Dieser Abschnitt ist auch die Vertonung des ersten, als „Sonnenaufgang“ betitelten Kapitels, in Friedrich Nietzsches gleichnamigem Werk (welches in der Komposition von Richard Strauß vertont wurde) und thematisch deutliche Parallelen zu „2001: A Space Odyssey“ aufweist. „Also sprach Zarathustra“ ist im Verlauf des Filmes noch zwei weitere Male zu hören. Bei [00:15:20] untermalt dieses Werk, nachdem wiederum kurz eine aufgehende Sonne im Bild zu sehen ist, den evolutionären Schritt vom Primaten zum Homo Erectus. Zum dritten Mal zu hören ist es am Ende des Filmes [02:19:00], als Bowman nach dem Erscheinen des Monolithen in Form eines übergroßen Babys wiedergeboren wird. Der musikalische Verweis auf Nietzsches Werk ist ein wesentlicher Hinweis auf eine mögliche Interpretation des Films. So kann die Wiedergeburt Bowmans den in „Also sprach Zarathustra“ thematisierten „Übermenschen“ darstellen, welcher dieser Interpretation folgend den nächsten Schritt in der Evolution der Menschheit darstellt.

Im Gegensatz dazu sind die Bilder des Weltalls (die Reise von Dr. Heywood Floyd zur Raumstation [00:19:55] und von der Raumstation zum Mond [00:33:43] sowie im Abspann [02:20:00]) mit „An der schönen blauen Donau“ von Johann Strauß unterlegt. Dies steht in starkem Kontrast zu Richard Strauß' Komposition. Der Donauwalzer kann in diesem Kontext für die Harmonie und die Schönheit der Natur bzw. der Schöpfung stehen. Am Beginn von Kapitel 3 (der Reise zum Jupiter [00:54:40]) ist anstelle von „An

der schönen blauen Donau“ das Adagio aus der „Gayane Ballet Suite“ von Aram Chatschaturjan zu hören. Das von einer Solo-Geige gespielte, melancholisch klingende Stück kann als Metapher für die Einsamkeit und Abgeschiedenheit im All interpretiert werden.

Ein weiteres, im Film wiederholt verwendetes Stück ist das „Kyrie“ aus György Ligetis Requiem, welches immer dann zu hören ist, wenn ein Monolith sichtbar wird ([00:12:15], [00:05:53], [02:00:10]) und welches diesem einen mystischen, überwirklichen Charakter verleiht. Damit nimmt es eine leitmotivische Funktion ein.

In Kapitel 3 und 4 ist der Einsatz der Musik dagegen den Dialogen und den Geräuschen untergeordnet. Dabei liegen die Sprache und auch Geräusche nicht zwingend auf dem Centerkanal und sind von vorne zu hören, wie es in den meisten Filmen üblich ist. Die räumliche Verteilung bleibt überdies nicht über eine Szene hinweg gleich, sondern springt oft simultan mit der Bildebene. Dies bestärkt die visuellen Eindrücke der oft nicht geraden, also vertikalen räumlichen Ausrichtung der Ereignisse im Bild durch die aufgehobene Gravitation.

Der Stimme von HAL (gesprochen von Douglas Rain) klingt, entgegen der nach filmischen Konventionen üblichen Darstellung von Computerstimmen, vergleichsweise menschlich, aber übertrieben freundlich und einfühlsam. Dies wirkt besonders gegen Ende des vierten Kapitels (nachdem der Computer absichtlich vier Menschen getötet hat und Bowman eröffnet, diesen nicht mehr an Bord des Schiffs lassen zu wollen [01:41:24]) sehr grotesk. Man kann von einer inkongruenten verbalen Kommunikation sprechen, der semantische Gehalt der Sprache stimmt nicht mit der Prosodie und der Klangfarbe der Stimme überein. Später im Film, als HAL seine Angst vor dem Ableben äußert, wirkt seine Stimme noch menschlicher, emotionaler als davor. Dies bedeutet eine größere Variation in der Sprachmelodie, Klangfarbe und Lautstärke, die von der sonst monotonen Wiedergabe der Worte abweicht. Es wirkt so, als würde er echte Angst empfinden [01:52:00]. Die Verdeutlichung seiner schwindenden Funktion, während Bowman ihn abschaltet, ist auf der Tonebene sehr effektiv umgesetzt. Zuerst wird seine Stimme immer langsamer, seine Wortwahl immer einfacher und der Klang der Stimme artifizieller [01:53:00]. Zum Schluss hin, mit nun in Klangfarbe und Tonhöhe stark verfremdeter Stimme (die eher den Konventionen einer Computerstimme entspricht),

gibt HAL ein einfaches Kinderlied wieder, währenddessen die schon anfangs tiefe Tonhöhe immer weiter abnimmt, bis seine Stimme schließlich gänzlich verstummt.

Die Umgebungsgeräusche auf der Discovery sind reduziert auf Geräusche mit mechanischem oder elektrischem Ursprung im Raumschiff und können der filmischen Wirklichkeit zugeordnet werden.

Bemerkenswert ist der Einsatz von Stille im Film. So sind in den meisten Sci-Fi Filmen im Weltall Atmos oder Geräusche der Raumschiffe zu hören. In 2001: A Space Odyssey wird das All dagegen physikalisch korrekt als lebens- und bewegungsfremder Raum dargestellt, in dem keine Schallausbreitung möglich ist. Das Resultat ist absolute Stille. Auf die FilmzuseherInnen wirkt die oft über lange Szenen hinweg ausgehaltene Lautlosigkeit eher unangenehm. Sie vermittelt eine Ungewissheit und Orientierungslosigkeit. Die Stille stellt die Weite des Universums, ebenso wie die Abwesenheit jeglichen Lebens, also den Tod, metaphorisch dar.

Sehr effektiv wirkt die indirekte Subjektivierung, die bei allen Szenen eingesetzt wird, in denen sich einer der Astronauten in seinem Raumanzug befindet. So ist beispielsweise in der über sechs Minuten dauernden Szene, in der Bowman das defekte Teil ausbaut [01:12:10], außer dem konstanten Rauschen der Sauerstoffversorgung in seinem Helm und der langsamen, gleichmäßigen, aber durch die verhältnismäßig hohe Lautstärke im Vordergrund stehenden Atmung des Astronauten nichts zu hören. Das Filmpublikum befindet sich daher in der Hörperspektive des Astronauten selbst oder seines Kollegen, der die Geräusche durch die Intercom hört. Die vermittelte Bedeutung ist unmissverständlich: Der Astronaut ist allein in der Weite des Weltalls, außer ihm gibt es dort draußen kein Leben. Das Ableben von Frank Pooles, nachdem dieser von der Raumkapsel ins All geschleudert wird, wird ebenfalls durch den Einsatz dieser Geräusche bzw. deren Ausbleiben unterstrichen. Zu Beginn der Szene [01:30:20] ist nur das Rauschen der Sauerstoffversorgung und das Atmen aus der Sicht Pooles zu hören, auch wenn im Bild die Ereignisse aus der Sicht seines Kollegen Bowman im Inneren der Discovery gezeigt wird. Nach der im Bild nicht sichtbaren Attacke der Raumkapsel ist das Atmen verstummt [01:32:16]. Es handelt sich also um eine Parallelisierung, die Tonebene liefert dem Filmpublikum Informationen, die vom Bild nicht kommuniziert werden.

5.5 „Der Dunkle Reigen“ (Dominik Bücheler, 2017)

„Der Dunkle Reigen“ ist ein mittellanger Spielfilm von Dominik Bücheler aus dem Jahr 2017 und handelt von der verbotenen Liebe zwischen zwei Frauen, die im Jahr 1919 in einem österreichischen Bergdorf eine Hexenjagd auslösen. Das Historiendrama behandelt die Themen Vertreibung und Flucht, sowie die homosexuelle Beziehung zweier Frauen, die von der Gesellschaft nicht toleriert wird. Der Aufbau der Handlung verläuft chronologisch und wird über zwei Handlungsstränge, die parallel zueinander verlaufen und sich immer wieder kreuzen erzählt. Im Lauf der Handlung kommt es zu einer Reihe von Konflikten, die sich immer weiter zuspitzen, bis die Situation eskaliert.

5.5.1 Narration und Dramaturgie

In den Wirren des ersten Weltkriegs gelangen die beiden heilkundigen Frauen Agnes (Anne Gleich) und ihre Mentorin Jadwiga (Elisabeth Rass) in ein österreichisches Bergdorf, wo die beiden um Hilfe für die kranke und geschwächte Jadwiga bitten. Nach der anfänglichen Zurückweisung durch Heinrich (Wolfgang Hundegger), dem Mesner des Dorfes, erbarmt sich schließlich Heinrichs Tochter Frieda (Isabella Jeksche) und versorgt die beiden mit einer warmen Mahlzeit. Anschließend bittet Frieda die beiden Fremden wieder zu gehen. Jadwiga jedoch eröffnet Frieda, die an einem grauen Star leidet, dass sie ihr helfen könne, ihr Augenlicht wiederzuerlangen, worauf sich Frieda, nach einigen Überredungsversuchen einlässt. Agnes führt den Eingriff durch und kümmert sich danach um Frieda. Am nächsten Morgen klopft Heinrich an der Tür, da seine Tochter nicht zum Gottesdienst erschienen ist. Wutentbrannt zerrt er seine Tochter vor das Haus, wo einige Dorfbewohner das Geschehen misstrauisch beobachten. Er reißt Frieda den Verband vom Gesicht, woraufhin diese bemerkt, dass ihr Augenlicht langsam zurückkehrt. Frieda widersetzt sich den Aufforderungen ihres Vaters, die beiden Heilkundigen hinauszuerwerfen. Der Vorfall spricht sich herum und fortan hilft Agnes mit ihrem medizinischen Wissen, die Krankheiten und Verletzungen

der Dorfbewohner zu versorgen. Mesner Heinrich beobachtet die Ereignisse misstrauisch und äußert beim Pfarrer Bedenken bezüglich der Fremden im Haus seiner Tochter. Der Zustand von Jadwiga verschlechtert sich zunehmend und wenig später stirbt diese an ihrer Krankheit. In den darauffolgenden Tagen helfen Agnes und Frieda einer Frau bei der Geburt ihres Kindes, das aber tot geboren wird. Verzweifelt über die Ereignisse der vergangenen Tage wird Agnes von Frieda getröstet, wobei sich die beiden Frauen näherkommen. Heinrich unterdessen versucht wiederum, den Pfarrer gegen Agnes aufzubringen, da er durch seinen Aberglauben die Feuerbestattung Jadwigas und Agnes' Wiederbelebungsversuch bei der Geburt des toten Babys als einen heidnischen Ritual interpretiert. Der Pfarrer weist Heinrich jedoch wiederum ab, woraufhin dieser zur Hütte seiner Tochter geht und schockiert durchs Fenster beobachtet, wie Agnes und Frieda miteinander schlafen. In seinem Stolz verletzt, ruft der Mesner die Bewohner des Dorfes in der Kirche zusammen, wo er die die Ereignisse schildert und die Bewohner aufhetzt, dem „gottlosen Treiben [...] ein Ende zu machen“. Heinrich und seine Handlanger stürmen darauf das Haus seiner Tochter und Heinrich peitscht Frieda, nach einer hitzigen Diskussion vor Agnes' Augen mit seinem Gürtel aus, die von den Schergen des Mesners festgehalten wird. In ihrer Verzweiflung entwickelt Agnes übernatürliche Kräfte und schließlich gelingt es nur Agnes und der verwundeten Frieda, lebend aus dem Haus zu gelangen.

5.5.2 Tongestaltung

Die Tongestaltung ist quasi-naturalistisch. Die minimalistische, naturalistische Vertonung der Umgebungsgeräusche und der Dialoge unterstreicht die Kargheit des Lebens in einem winterlichen Bergdorf. Die Musik ist bis auf den Beginn des Filmes extradiegetisch, und kommentiert und dramatisiert die Ereignisse auf der Bildebene.

Die erste Szene des Films [00:00:15] zeigt die Reise von Agnes und Jadwiga ins Dorf, zu der „Großer Gott wir loben dich“ zu hören ist, wobei eine Betrachtung dieser Szene in Bezug auf die Diegese aufschlussreich ist. Die Musik stellt sich als diegetisch heraus, als der von Heinrich dirigierte Kirchenchor im Bild zu sehen ist [00:01:00]. Die Bildebene

springt von diesem Zeitpunkt an zwischen dem Chor, Agnes und Jadwiga sowie den misstrauischen Dorfbewohnern beim Anblick der beiden hin und her. Durch das bekannte Kirchenlied wird der kulturelle Rahmen, in dem sich die Handlung abspielt sehr effektiv etabliert. Im Bild wird anfangs nur die Jahreszahl eingeblendet, das kulturelle Umfeld wird durch die Assoziationen der Zuseher mit dem Kirchenlied erklärt.

Die Musik im Rest des Filmes ist extradiegetisch und entspricht der klassischen Funktion der Filmmusik dramaturgisch wichtige Momente der Handlung zu untermalen, zu verstärken bzw. zu dramatisieren. Heinrichs Auftritte sind zumeist von langsamen, tiefen und bedrohlichen Streichern untermalt. Zum ersten Mal deutlich wird dies, als er seiner Tochter eintrichert, den Bitten der Fremden nicht nachzugeben [00:04:30]. Später, als er Frieda vor dem Gottesdienst nicht in der Kirche antrifft [00:11:00] und zur Hütte seiner Tochter stürmt, sind die Streicher dynamischer und es kommen perkussive Elemente hinzu. Bei der Hütte angelangt, lässt die Lautstärke kurz nach und wird langsam wieder lauter, wodurch sich eine Spannung aufbaut, die aufgelöst wird, als Agnes anstelle seiner Tochter die Türe öffnet. Ebenfalls zum Einsatz kommt die extradiegetische Filmmusik bei Jadwigas Tod sowie bei ihrer anschließenden Verbrennung [00:17:00] und in der Szene, in sich Agnes und Frieda zum ersten Mal küssen. Dabei ist die Tonhöhe aber höher und die enthaltenen Intervalle harmonischer. Der Kontrast wird insbesondere deutlich, als Frieda bei der Verbrennung von Jadwiga Agnes' Hand nimmt [00:17:25]. Die musikalische Harmonie drückt ihr Mitgefühl und ihre Zuneigung zu Agnes aus. Nach dem drauffolgenden Schnitt der Bildebene auf Heinrichs Gehstock, hinter dem in einiger Distanz der Scheiterhaufen zu sehen ist, mischt sich eine Dissonanz in die Harmoniefolge [00:17:35]. Der Kontrast polarisiert das Bild und lässt vermuten, dass Heinrich das Geschehen aus der Distanz misstrauisch beobachtet.

[00:12:15] In der Szene, in der Friedas Augenlicht zurückkehrt, ist eine sehr hohe, langsame Tonfolge aus zwei Akkorden zu hören. Die Intervalle des ersten Akkordes klingen Dissonant, eine Spannung baut sich auf. Diese Dissonanz löst sich im zweiten Akkord, gerade als die Schärfe des Bildes wieder zunimmt. Das Bild wird also durch die Musik paraphrasiert, die Auflösung der Dissonanz unterstreicht die Wiederkehr des Augenlichtes. Die Paraphrasierung wird jedoch durch die abstrakte Klangfarbe relativiert, die zusammen mit der hohen Frequenz der Töne dem Ereignis etwas

Geheimnisvolles, Mystisches verleihen, was als Hinweis auf eine mögliche Übernatürlichkeit des Vorganges gesehen werden kann. Ein kurzer Einsatz desselben Klanges ist bereits bei der Operation, bevor Frieda dem Eingriff zustimmt [00:08:30], zu hören.

Wichtig in Bezug auf die Narration des Films sind stimmliche Laute, die allerdings keine Sprache darstellen. Die Information über Jadwigas Krankheit wird zu einem wesentlichen Teil auch über die Tonebene transportiert. So ist im Bild oft ein blutiges Taschentuch zu sehen, es ist aber auch das Geräusch des Hustens, das den Charakter immer begleitet. Das Husten wird zum Ende hin heftiger und tritt immer häufiger auf womit es den Gesundheitszustand von Jadwiga metaphorisch darstellt.

In der Szene der fehlgeschlagenen Geburt [00:20:00] ist es aber vor allem die Abwesenheit der Geräusche, also die Stille, die vom ersten Moment an das Gefühl vermittelt, dass mit dem Neugeborenen etwas nicht stimmt – selbst als dieses im Bild noch nicht zu sehen ist.

6 Fazit

Die Analyse der Filme bestätigt die Annahme, dass mithilfe der Tongestaltung vielfältige Bedeutungen sowohl alleine als auch im Zusammenspiel mit der Bildebene und der Narration kommuniziert werden können. Die anfängliche Kritik am Tonfilm ist insofern auch aus heutiger Sicht noch relevant, sofern sie als Forderung interpretiert wird, die Tonspur als eigenständiges Element im Gesamtkontext des Spielfilmes einzusetzen, um das bloße Rezitieren der Bildebene zu vermeiden. Ansonsten kann das Potential, das durch die intermodale Assoziation von visuellen und auditiven Sinnesreizen gegeben ist, nicht genutzt werden. Bei einem sinnvollen Einsatz kann durch die Elemente der Tonspur im Zusammenspiel mit der Bildebene dem narrativen Gesamtkontext des Films ein deutlicher Mehrwert entstehen. Dieser vermittelt den FilmzuseherInnen zum einen Informationen über die räumliche und zeitliche Verortung sowie über die Vorgänge und Ereignisse in der filmischen Realität. Zum anderen können durch die zielgerichtete Gestaltung der Tonspur unterbewusst Emotionen ausgelöst und Stimmungen transportiert werden. Die Aufmerksamkeit des Filmpublikums kann auf wichtige Details gelenkt und bestimmte Ereignisse der Handlung verdeutlicht, hervorgehoben und kommentiert werden, wodurch diese bedeutungsvoll werden. Wichtig ist, dass die Bedeutungszuweisung immer erst durch die individuelle kognitive Leistung der FilmzuseherInnen entsteht. Kulturelle, historische, universelle und filmische Hörkonventionen müssen berücksichtigt bzw. die Rahmenbedingungen etabliert oder in Form einer Metakommunikation kommuniziert werden, um die intendierte Bedeutung auch erzielen zu können.

Die in der Fachliteratur dargestellten Konzepte liefern eine gute Ausgangsbasis in Bezug auf die dramaturgischen Funktionen der Tongestaltung. Insbesondere direkte und indirekte Subjektivierungen sind ein vielgenutztes Mittel in den der Analyse unterzogenen Filmen, um gezielt die Aufmerksamkeit des Filmpublikums auf die essentiellen Elemente und Ereignisse der Handlung zu richten, wohingegen unwichtigere Elemente ausgeblendet werden. Auch die Perspektive, aus der die akustischen Ereignisse zu hören sind, stellt somit einen wesentlichen Faktor in Bezug

auf die Unterstützung der Narration durch die Tonspur dar. Der Grad der Diegese kann bewusst eingesetzt werden, um eine Bedeutung zu transportieren. Musik, Sprache oder Geräusche sind, wenn diese nicht der filmischen Realität entspringen, in allen Filmen bedeutungsvoll. Jedoch ist der Einfluss des Grades der Diegese auf die Narration aufgrund der konzeptionellen Unterschiede in den analysierten Filmen schwieriger zu objektivieren. In Lars von Triers „Melancholia“ sind Geräusche und Sprache beispielweise fast ausnahmslos der Diegese zuzuordnen. Die Handlung wird von den Dialogen getragen, die Ereignisse werden von der extradiegetischen Filmmusik kommentiert. In Lynchs „Mullholand Drive“ sind die Dialoge zumeist diegetisch. Die Geräusche entspringen manchmal der filmischen Realität: Oft ist aber keine eindeutige Zuordnung möglich und die Geräusche verschwimmen in einem Klanggemisch mit der Musik. Die Geräusche besitzen so eine ähnliche Funktion wie die Filmmusik in „Melancholia“, nämlich die Ereignisse zu akzentuieren, zu kommentieren und so als bedeutungsvoll zu erklären. Die Beziehungen zwischen Bild, Ton und Narration sind vielfältig. Paraphrasierungen, Polarisierungen sind in den analysierten Filmen häufig zu finden, seltener dagegen Parallelisierungen oder Inkongruenzen, die jedoch, insbesondere auch durch den sparsamen Einsatz, große Wirkung in Bezug auf die erzählten Ereignisse entfalten.

Die empirische Analyse zeigt somit, dass die in der Theorie erarbeiteten Konzepte anwendbar sind, um die Tongestaltung in Bezug auf die Narration einordnen und deren Bedeutung deuten zu können. Der Einsatz der Bild-Ton-Beziehungen, der Hörperspektive, der Grad der Subjektivierung und Diegese sind in den untersuchten Filmen aber sehr unterschiedlich und auf den individuellen Kontext im jeweiligen Film zugeschnitten. Die unterschiedlichen Bedeutungszusammenhänge besitzen dabei eine große Vielfalt und die Beziehung zwischen Bild, Ton und Narration oft eine gewisse Komplexität. Eine völlige Objektivierung der Tongestaltung in Bezug auf die Narration im Spielfilm kann in der Praxis also nur schwerlich erfolgen. Dadurch wird aber umso deutlicher, dass immer der Gesamtkontext (das heißt das komplexe Netzwerk der Beziehungen zwischen den verschiedenen Kommunikationskanälen) entscheidend ist, welche Wirkung durch den Einsatz eines Gestaltungsmittels erzielt wird. Konzepte und Techniken zur Gestaltung der Tonspur, die in einem Film erfolgreich sind, werden dadurch in anderen Produktion mit großer Wahrscheinlichkeit nicht zielführend sein.

Die Gestaltung der Tonspur muss daher immer im Zuge einer aktiven Auseinandersetzung mit der Erzählstruktur, den Figuren, sowie den Höhen, den Tiefen und den Wendepunkten im Handlungsverlauf erfolgen, um hinsichtlich der Handlung des Films sowie der Wirkung in Kombination mit der Bildebene, sinnvolle kreative Entscheidungen treffen zu können, damit beim Filmpublikum auch die gewünschte Wirkung erzielt wird. Die Unterschiede der internationalen Produktionen im Vergleich zu „Der Dunkle Reigen“ sind merkbar und manifestieren sich vor allem in den Beziehungen zwischen Bild und Ton. Das Netzwerk an Beziehungen ist in den internationalen Produktionen ein sehr komplexes, ausgeklügeltes System. Die Bedeutung wird selten vorweggenommen, oft beruht sie auf der kognitiven Interpretationsleistung des Filmzusehers und schafft es so die Beziehung immer interessant zu gestalten. Bei „Der Dunkle Reigen“ sind Polarisierungen, Parallelisierungen vorhanden, jedoch handelt es sich bei der Musik und dem Geräuscheinsatz meist um eine Paraphrasierung, das heißt, die Tonebene liefert die gleichen Informationen bzw. löst die gleichen Emotionen aus wie Bild und Handlung. Des Weiteren ist die direkte Subjektivierung weniger stark ausgeprägt, die Geräusche sind weniger selektiv ausgewählt und die Aufmerksamkeitslenkung des Publikums ist nicht so stark auf die für die Narration wesentlichen Elemente hin optimiert. Die Unterschiede sind in vermutlich den größeren personellen, zeitlichen und technischen Ressourcen geschuldet, die bei internationalen Produktionen zum Einsatz kommen, sowie vorhandener Erfahrung und größerem Know-How der Audiodepartments. Ebenso sind die Gründe für die Unterschiede in der Planung und der Vorproduktion zu suchen, da bei „Der Dunkle Reigen“ die Tongestaltung in Bezug auf die Narration nur eine untergeordnete Rolle spielte.

Die Analyse hat auch eine Möglichkeit der Bedeutungsgebung gezeigt, die in der Literatur nur bedingt behandelt wurde. Der Einsatz der Filmmusik vermittelt in „2001: A Space Odyssey“ und „Melancholia“ eine wichtige Information über die Narration bzw. ihrer Interpretation. Die Funktion der Filmmusik unterscheidet sich von ihrer üblichen aber insofern, als mit ihrer Auswahl auf andere, außerfilmische Werke referenziert wird. Die Musik wird somit nicht (nur) zum Erzeugen von Stimmungen oder zum Verdeutlichen von Emotionen verwendet, sondern es ist die Auswahl der Stücke und deren thematischer Hintergrund, der bedeutungsvoll ist. In diesem Fall ist es

notwendig, diese Werke zu kennen und dadurch deren Bedeutung mit der Handlung assoziieren zu können, damit ein Mehrwert in Bezug auf die Narration entsteht. Es handelt sich daher um eine Form der Metakommunikation. Durch die bedeutungsvolle Musikauswahl werden über einen zusätzlichen Kommunikationskanal wichtige Informationen zur Interpretation der Narration geliefert.

Eine weitere wichtige Erkenntnis aus der Analyse besteht in der Wichtigkeit von Kontrasten bzw. Veränderungen, welche in der Fachliteratur zwar erwähnt, deren Bedeutung aber nur untergeordnet Beachtung geschenkt wird. Analog zu der in der Dramaturgie notwendigen Abwechslung von Konflikten bzw. Spannungen und deren Auflösungen sind auch in der Tongestaltung Kontraste essentiell. Zum einen können Unterschiede bzw. Veränderungen der Lautstärke, der Klangfarbe oder der räumlichen Verteilung von Klängen oder Geräuschen Spannungsbögen unterstreichen. Zum anderen entfalten Kontraste auch in Bezug auf die behandelten Konzepte große Wirkung, zum Beispiel Veränderungen im Grad der Subjektivierung oder in der Beziehung zwischen Bild und Ton.

Eine genauere Erarbeitung der narrativen Funktion von Kontrasten in der Tongestaltung sowie die Bedeutungsgebung durch akustische Verweise auf außerfilmische Referenzen zählen daher zu den wenig behandelten Gebieten der akustischen Gestaltung im Spielfilm und könnten die Basis für eine weitere Erschließung der Thematik dienen.

Literaturverzeichnis

Altmann, Hans; Ziegenhain, Ute (2007): Phonetik, Phonologie und Graphemik fürs Examen. 2., überarb. und erg. Aufl. Göttingen: Vandenhoeck & Ruprecht (Linguistik fürs Examen, 3).

Atmanspacher, Harald; Primas, Hans; Wertenschlag-Birkhäuser, Eva (Hg.) (1995): Der Pauli-Jung-Dialog und seine Bedeutung für die moderne Wissenschaft. Berlin, Heidelberg: Springer Berlin Heidelberg.

Bordwell, David (1985): Narration in the fiction film. Madison: Univ. of Wisconsin Press.

Chion, Michel; Gorbman, Claudia (Hg.) (1994): Audio-vision. Sound on screen. New York: Columbia Univ. Press.

Duden | Dis-so-nanz | Rechtschreibung, Bedeutung, Definition, Synonyme, Herkunft (2018). Online verfügbar unter <https://www.duden.de/rechtschreibung/Dissonanz#Bedeutung1>, zuletzt aktualisiert am 21.03.2018, zuletzt geprüft am 21.03.2018.

Duden | Pho-nem, Fo-nem | Rechtschreibung, Bedeutung, Definition (2018). Online verfügbar unter <https://www.duden.de/rechtschreibung/Phonem>, zuletzt aktualisiert am 09.03.2018, zuletzt geprüft am 09.03.2018.

Duden | Vo-kal | Rechtschreibung, Bedeutung, Definition, Herkunft (2018). Online verfügbar unter <https://www.duden.de/rechtschreibung/Vokal>, zuletzt aktualisiert am 27.03.2018, zuletzt geprüft am 27.03.2018.

Flückiger, Barbara (2007): Sound design. Die virtuelle Klangwelt des Films. Zugl.: Zürich, Univ., Diss., 1999. 3. Aufl. Marburg: Schüren (Zürcher Filmstudien, 6).

Friesecke, Andreas (2007): Die Audio-Enzyklopädie. Ein Nachschlagewerk für Tontechniker. Berlin: De Gruyter.

Görne, Thomas (2014): Tontechnik: Hören, Schallwandler, Impulsantwort und Faltung, digitale Signale, Mehrkanaltechnik, tontechnische Praxis: Carl Hanser Verlag GmbH & Company KG.

Görne, Thomas (2017): Sounddesign. Klang Wahrnehmung Emotion. München: Carl Hanser Verlag GmbH & Co. KG.

Hayward, Susan (1996): Key concepts in cinema studies. 1. publ. London u.a.: Routledge (Key concepts series).

Hellbrück, Jürgen; Ellermeier, Wolfgang (2004): Hören. Physiologie, Psychologie und Pathologie. 2., aktualisierte und erw. Aufl. Göttingen, Bern, Toronto, Seattle: Hogrefe, Verl. für Psychologie.

Jung, C. G. (2001): Archetypen. Hg. v. Lorenz Jung. München: Deutscher Taschenbuch Verlag (Dtv, 35175).

Karl Heinz Wagner: Phonetik und Phonologie: Kapitel 1: Einleitung und allgemeine Grundbegriffe. Universität Bremen. Online verfügbar unter <http://www.fb10.uni-bremen.de/khwagner/phonetik/kapitel1.aspx>, zuletzt geprüft am 02.04.2018.

Keutzer, Oliver; Lauritz, Sebastian; Mehlinger, Claudia; Moormann, Peter (2014): Filmanalyse. Wiesbaden: Springer VS (Film, Fernsehen, Neue Medien).

Lärmfibel. Online verfügbar unter <http://www.staedtebauliche-laermfibel.de/?p=5&p2=2.2>, zuletzt geprüft am 26.03.2018.

Lensing, Jörg U. (2006): Sound-Design, Sound-Montage, Soundtrack-Komposition. [über die Gestaltung von Filmtönen]. 1. Aufl. Stein-Bockenheim: Mediabook Verl.

Leonhard, J. F. (1999): Medienwissenschaft: ein Handbuch zur Entwicklung der Medien und Kommunikationsformen: W. de Gruyter (Teil 2).

Mag. Dr. Anton Fuxjäger (2007): Diegese, Diegesis, diegetisch. Versuch einer Begriffsentwirrung 16,2.

Magee, Bryan (2002): The Tristan chord. Wagner and philosophy. New York: Metropolitan Books/Henry Holt and Company.

Raffaseder, Hannes (2010): Audiodesign. 2., aktualisierte und erw. Aufl. München: Fachbuchverl. Leipzig im Carl-Hanser-Verl.

Sengpiel, Eberhard (2017a): Frequenz Periodendauer umrechnen Frequenzformel
Formel Hz Hertz amplitude umrechnen nach ms Periode Dauer Schwingungsdauer
T Schwingung $t=1/f$ Formel Wellenlaenge Formelsammlung Akustik Rechner
berechnen Amplitude Kreisfrequenz - sengpielaudio Sengpiel Berlin.
sengpielaudio. Online verfügbar unter <http://www.sengpielaudio.com/Rechner-periodendauer.htm>, zuletzt aktualisiert am 12.02.2017, zuletzt geprüft am 19.03.2018.

Sengpiel, Eberhard (2017b): Sone phon Lautheit Lautstaerke 0,3 - 0,4 - 0,5 - 0,6
Definition sone phone dB Pegel umrechnen dBA dB Dezibel Rechner berechnen
Lautstaerkeskala Kenngroesse Luefter Ventilator Laerm Psychoakustik
Lautheitspegel Lautstaerkepegel SPL Lautheit sone in dB Lautstaerke Bewertung
Bewertungsfilter Zimmerlautstaerke - sengpielaudio Sengpiel Berlin. sengpielaudio.
Online verfügbar unter <http://www.sengpielaudio.com/RechnerSonephon.htm>,
zuletzt aktualisiert am 12.02.2017, zuletzt geprüft am 15.03.2018.

Sonnenschein, David (2001): Sound design. The expressive power of music, voice, and
sound effects in cinema. Studio City, Calif.: Michael Wiese Productions.

Stephan, Achim; Walter, Henrik (Hg.) (2003): Natur und Theorie der Emotion.
Interdisziplinäres Kolloquium. Paderborn: Mentis-Verl.

Tillmann, Schiel (1995): Akustische Phonetik. Institut für Phonetik und Sprachliche
Kommunikation, Universität München. Online verfügbar unter
<https://www.phonetik.uni-muenchen.de/studium/skripten/AP/APKap3.html>,
zuletzt aktualisiert am 21.04.2008, zuletzt geprüft am 02.04.2018.

Tristan and Isolde. Online verfügbar unter
<https://www.laits.utexas.edu/tristan/motive1-3.php>, zuletzt geprüft am 07.05.2018.

Umweltbundesamt: Schalldruckpegel, energieäquivalenter Dauerschallpegel und
Lärmindizes. Online verfügbar unter
<http://www.umweltbundesamt.at/umweltschutz/laerm/schalldruckpegel>, zuletzt
geprüft am 15.03.2018.

Watzlawick, Paul (2004): Wie wirklich ist die Wirklichkeit? Wahn, Täuschung, Verstehen.
Sonderausg., 3. Aufl. München, Piper.

Wuss, Peter (1993): Filmanalyse und Psychologie. Strukturen des Films im Wahrnehmungsprozeß. Berlin: Ed. Sigma Bohn (Sigma Medienwissenschaft, 15).

Filmverzeichnis

Bücheler (2017): Der Dunkle Reigen. Mit Isabella Jeschke, Anne Gleich, Wolfgang Hundegger, Elisabeth Rass. Servus TV (Salzburg), 24.11.2017, 31Minuten.

David Lynch (2001): Mullholand Drive. Straße der Finsternis. Mit Naomi Watts, Laura Elena Harring, Justin Theroux. Neal Edelstein, 141 Minuten.

Kubrick, Stanley; Clarke, Arthur C. (1968): Stanley Kubricks 2001, Odyssee im Weltraum. Mit Keir Dullea und Gary Lockwood. Stanley Kubrick. 1 DVD-Video (ca.143 Min)

Trier, Lars von (2011): Melancholia. Mit Kirsten Dunst und Charlotte Gainsbourg. Concorde Video, 135 Minuten.

Abbildungsverzeichnis

Abbildung 1: (Sengpiel 2017a) y ... Schalldruck (p) a ... Amplitude T ... Periodendauer der Schwingung λ ... Wellenlänge	15
Abbildung 2: Darstellung des Signalverlaufs verschiedener Klänge und Geräusche (bzw. deren Amplitude) in Abhängigkeit von der Zeit auf mikroskopischer Ebene. t ... Zeit $A(t)$... Amplitude T ... Periodendauer (Raffaseder 2010, S. 58).....	16
Abbildung 3: Darstellung eines Schallsignals als ADSR- Hüllkurve $A(t)$... Amplitude t ... Zeit $t(A)$... Attack- Time $t(D)$... Decay- Time $t(S)$... Sustain- Time $t(R)$... Release- Time (Raffaseder 2010, S. 65).....	17
Abbildung 4: Zur Veranschaulichung von im Alltag vorkommenden Schallereignissen und deren Schalldruckpegel in dB SPL können diese Beispiele dienen. (Umweltbundesamt)	25
Abbildung 5: Kurven gleicher Lautstärke (auch Fletcher- Munson Curves genannt) (Friecke 2007, S. 125).....	26
Abbildung 6: Zusammenhang zwischen den Einheiten Phon und Sone (Sengpiel 2017b)	27
Abbildung 7: Zeitfunktionen von einem Sinuston, einem Klang und einem Geräusch Die Periodendauer $T = 1/f$ (Hellbrück und Ellermeier 2004, S. 55).....	31
Abbildung 8: Musikalische Intervalle und deren Verhältnisse (Friecke 2007, S. 122).	33
Abbildung 9: Darstellung von einem harmonischen Tones, eines Klanges, eines Geräusches sowie von weißem Rauschen im Frequenzbereich (Lärmfibel)	36
Abbildung 10: Vokalviereck (Raffaseder 2010, S. 84).....	41
Abbildung 11: Die Kulisse von Dogville (2003).....	60
Abbildung 12: Takt 1-3 der Overtüre von Richard Wagners „Tristan und Isolde“ (Tristan and Isolde).....	90