

Die Wirkung eines Films, die entsteht ja in einer imaginären Mitte zwischen dem Absender und dem Empfänger, zwischen der Leinwand und den Bedürfnissen und den Emotionen sehr vieler Menschen, die man nicht kennt und an die man im Einzelnen nicht denken kann. Eine gute Geschichte muss das hergeben, und allzu triviale oder banale Geschichten sind vorbei, wenn im Kino das Licht angeht, und die guten Geschichten nimmt man mit und bewegt sie, wie es in den alten Märchen heißt, in seinem Herzen.

Wolfgang Kohlhaase, 2006¹

Diegetic Sound. Zur Konstituierung figureninterner und -externer Realitäten im Spielfilm

Didi Merlin (Poitiers)

Einleitung

Gilt das, was der Drehbuchautor Wolfgang Kohlhaase im oben wiedergegebenen Zitat für die Wirkung eines Films im Allgemeinen behauptet, eigentlich auch für die Filmmusik, den Ton, den Dialog? Wenn ja, was ist das dann, die „imaginäre[...] Mitte zwischen dem Absender und dem Empfänger, zwischen der Leinwand und den Bedürfnissen und den Emotionen“ (ebd.) der Zuschauer/innen? Wo liegt diese Mitte? Lässt sie sich in Begriffe fassen oder entzieht sie sich einer sprachlichen Darstellung? Im Folgenden soll ein Anlauf dazu unternommen werden, die Wirkung dessen, was bei der Rezeption eines Spielfilms *gehört* wird, gleichermaßen in Abhängigkeit vom audiovisuellen Material und von den intersubjektiven und subjektiven Faktoren zu bestimmen, die die Rezipient/innen in diejenigen Prozesse mit einbringen, in deren Verlauf figureninterne und -externe Realitäten im Spielfilm konstituiert werden. Dies soll zunächst im Rückgriff auf etablierte Begrifflichkeiten geschehen, in Form einer kurzen Bestandsaufnahme, einer Skizzierung von Forschungsansätzen sowohl zum Sound-, als auch zum Diegesebegriff. Im Anschluss geht es um Kreuzungspunkte, Schnittmengen, Wechselwirkungen zwischen Bild, Sound und Publikum. In einem dritten Teil werden die terminologischen Überlegungen um die Analyse dreier Beispielsequenzen ergänzt, sie stammen aus *LA BATTIGLIA DI ALGERI* (*SCHLACHT UM ALGIER*, Algerien/Italien 1966, Gillo Pontecorvo, 117 min, Musik: Ennio Morricone / Gillo Pontecorvo), *L'OURS* (*DER BÄR*, Frankreich/USA 2006, Jean-Jacques Annaud, 94 min, Musik: Philippe Sarde / Pjotr Illjitsch Tschaikowski) und *LENZ* (Schweiz 2006, Thomas Imbach, 95 min, Musik: Peter Bräker / Balz Bachmann).

¹ Wolfgang Kohlhaase in: *LEBEN IN GESCHICHTEN – WOLFGANG KOHLHAASE*. Regie: Lutz Pehnert, Produktion: cine plus im Auftrag des RBB und in Zusammenarbeit mit ARTE. RBB 2006, 43:20 min. Gesendet im RBB am 21.02.2010, 23h50. Zitierter Ausschnitt: 00:38:05-00:39:34.

1. Terminologiegeschichten: Sound und Diegetisierung

Frank Schätzlein, Medienwissenschaftler an der Universität Hamburg, schreibt in einem von ihm und Harro Segeberg 2005 herausgegebenen Sammelband der Schriftenreihe der Gesellschaft für Medienwissenschaft zum Begriff *Sound*:

Insgesamt besteht die Tendenz zu einer weiten Sounddefinition, die längst über das rein Musikalische hinausreicht. Der Sound kann dabei als ‚neue Ära‘ oder ‚neues Paradigma‘ (von der Musik zum Sound bzw. vom Ton zum Sound) der musikwissenschaftlichen Forschung verstanden werden. Und: Die zentrale Bedeutung des Sounds für die Produktion und Rezeption von Rock- und Popmusik zeigt die Notwendigkeit einer Erweiterung des traditionellen Instrumentariums musikalischer Analyse um die Verfahren einer Soundforschung (Schätzlein 2005, 27).

Die Mehrdeutigkeit des Soundbegriffs, die einem einheitlichen Gebrauch entgegensteht, ist zugleich Grund seiner zunehmend populären Verwendung sowohl in der Alltags- als auch der Wissenschaftssprache. Schätzlein zufolge lassen sich dem Begriff neben den ursprünglichen, aus der Fachsprache des Jazz stammenden Signifikaten *Klangbildung*, *Klangfarbe*, *Klangqualität* und *typischer oder charakteristischer Klang*² mindestens acht weitere Bedeutungsebenen zuordnen:

1. Schall allgemein (alles auditiv Wahrnehmbare),
 2. Tonebene der elektronischen Medien (beispielsweise die Gesamtheit aller Schallereignisse auf der Film-Tonspur oder die auditive Ebene eines Fernsehprogramms oder Computerspiels),
 3. Geräusch oder Klangeffekt (Sounds als zumeist unidentifizierbare und nicht konkret benennbare Geräusche oder Schallereignisse: ‚Sound Effects‘ bzw. ‚SFX‘),
 4. charakteristischer Klang einer Ware / eines Industrieprodukts (Sounddesign als akustisches Produktdesign, z.B. für Fahrzeuge, Haushaltsgeräte, Telefone oder sogar Lebensmittel [...]),
 5. charakteristischer Klang einer Marke oder eines Unternehmens (Corporate Sound als akustischer Teil des Corporate Design-Gesamtkonzepts, z.B. bei Großunternehmen und Fernsehprogrammen wie ‚Das Erste‘ oder ‚ProSieben‘),
 6. „Erläuterung technischer Verfahrensweisen“ der Musikproduktion (Röhren-Sound oder Moog-Sound),
 7. „Umschreibung musikalischer Grundstimmungen (softer, knackiger Sound)“,
 8. „Mittel der qualitativen Bewertung (origineller, abgestandener Sound)“.
- (Schätzlein 2005, 26f.)³

² Schätzlein nennt als frühe Quelle, in der sich diese Verwendungsweise des Begriffs nachweisen lässt: Berendt, Joachim Ernst (1953) *Das Jazzbuch. Entwicklung und Bedeutung der Jazzmusik*. Frankfurt a.M.: Fischer.

³ Zitate in 6-8 nach Schätzlein aus: Rösing, Helmut (1996) „Klangfarbe und Sound in der ‚westlichen‘ Musik“. In: Finscher, Ludwig (Hg.) (1996) *MGG. Die Musik in Geschichte und Gegenwart*. Sachteil. Bd. 5: Kas-Mein. 2., neubearb. Ausg. Kassel/Stuttgart. S.156-159, hier: S.158.

Eine genaue Deskription der spezifischen Wirkung der hier unter dem Soundbegriff subsumierten Einzelphänomene kann an dieser Stelle nicht geleistet werden, dies wäre nur im Rahmen detaillierter interdisziplinärer Einzeluntersuchungen möglich, in denen sowohl produktionstechnische als auch rezeptionstheoretische Ansätze eine Rolle spielen müssten. Hier soll lediglich darauf hingewiesen werden, dass sich die folgenden Ausführungen primär auf das von Schätzlein unter Punkt 2 genannte Verständnis von *Sound* beziehen, also auf die Gesamtheit aller akustischen Ereignisse auf der Film-Tonspur, unabhängig davon, ob es sich um Musik, Sprache, Geräusche oder Spezialeffekte handelt.

Der auf diese Art sehr weit gefasste Soundbegriff sei nun mit einem Adjektiv in Verbindung gebracht, das auf eine äußerst interessante wissenschaftshistorische Karriere zurückblicken kann:⁴ wie bereits am Titel dieses Artikels erkennbar, handelt es sich um den Terminus *diegetisch*. Um sich der Frage zu nähern, was sich unter *diegetic sound* verstehen lässt, soll zunächst ein kurzer Blick auf die Begriffe *diegetisch*, *Diegese* und *Diegetisierung* geworfen werden. Das Adjektiv *diegetisch* sowie das Substantiv *Diegese* wurden Anfang der 50er Jahre von dem französischen Philosophen und Kunsttheoretiker Etienne Souriau (1892-1979) bzw. von dessen Tochter Anne Souriau aus der altgriechischen Philosophie entnommen und auf einen filmwissenschaftlichen Kontext übertragen.⁵ Etienne Souriau gehörte der sogenannten ‚filmologischen Schule‘ (‚Ecole de filmologie‘) an, deren Mitglieder sich „der wissenschaftlichen Untersuchung ‚filmischer Tatsachen‘ in ihrer psychologischen und physiologischen Wirkung auf den menschlichen Organismus“ (Kessler 1997, 132) widmeten. Neben Etienne Souriau zählten zu den Filmologen u.a. der Journalist, Filmregisseur, -produzent und -theoretiker Gilbert Cohen-Séat (1907-1980), der Literaturhistoriker Mario Roques (1875-1961), der Philosoph, Sozialpsychologe, Neuropsychiater, Pädagoge und Politiker Henri Wallon (1879-1962) und der Wissenschaftstheoretiker und Philosoph Gaston Bachelard (1884-1962). Die Schule gründete an der Sorbonne ein eigenes Institut für Lehre und Forschung und rief darüber hinaus eine internationale Zeitschrift ins Leben, die *Revue Internationale de Filmologie*, die von 1947 bis 1961 publiziert wurde. Etienne Souriau definiert den Begriff *diegetisch* 1951 in einem in dieser Zeitschrift veröffentlichten Aufsatz wie folgt:

Diegetisch ist alles, was man als durch den Film repräsentiert in Betracht zieht und als in der Art von Realität enthalten, die durch die Bedeutung des Films *als wahr behauptet wird*. Man könnte versucht sein, all das als die Realität der Tatsachen zu bezeichnen; und es spricht auch nichts gegen die Verwendung dieses Begriffs, wenn man sich in Erinnerung ruft, dass es sich um eine fiktionale Realität handelt (Souriau 1951, 237; Hv.i.O., ÜB M)⁶.

⁴ Vgl. dazu beispielsweise die Beiträge in *montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation*. 16, 2. Marburg, 2007: Schüren. Website: http://www.montage-av.de/a_2007_2_16.html (Stand: 26.03.2010).

⁵ Zu den genauen historischen Umständen der Begriffsprägung vgl. Merlin (2010).

⁶ « Est diégétique tout ce qu'on prend en considération comme représenté par le film, et dans le genre de réalité *supposé* par la signification du film : ce qu'on peut être tenté d'appeler la réalité des faits ; et ce terme même n'a pas d'inconvénient si on se rappelle que c'est une réalité de fiction. » (Souriau 1951, 237; Hv.i.O.) Zur Übersetzungsproblematik, insbes. zur Abgrenzung gegenüber der Übersetzung durch Frank Kessler (vgl. Souriau 1997, 151f., ÜB Kessler): vgl. Merlin (2010). Übersetzungen durch Merlin seien hier und im Folgenden gekennzeichnet als ‚ÜB M‘.

Der Literaturtheoretiker Gérard Genette semiologisiert – höchstwahrscheinlich in Anlehnung an den Filmtheoretiker Christian Metz⁷ – den Diegesebegriff Souriaus dahingehend, dass er ihn im Sinne der de Saussureschen Unterscheidung von *signifié* vs. *signifiant* mit der Seite des Signifikats gleichsetzt.⁸ Genette weist darüber hinaus darauf hin, dass man unter *Diegese* nicht die erzählte Geschichte (im Sinne von Fabula bzw. Handlung) verstehen könne, „sondern das Universum, in dem sie spielt“ (ebd.) – womit er einen früher durch ihn selbst verwendeten Sprachgebrauch korrigiert.⁹ Andere Autor/innen schließen sich Genette an. Sie definieren Diegese nicht als die jeweils erzählte Geschichte, sondern als das raumzeitliche Gefüge, in dem die Handlung stattfindet, so beispielsweise David Bordwell und Kristin Thompson in ihrem Buch *Film Art. An Introduction* (2008 [1997]), einem internationalen Standardwerk der wissenschaftlichen Filmanalyse:

The total world of the story action is sometimes called the film's *diegesis* (the Greek word for ‚recounted story‘). In the opening of *North by Northwest*, the traffic, streets, skyscrapers, and people we see, as well as the traffic, streets, skyscrapers, and people we assume to be offscreen, are all diegetic because they are assumed to exist in the world that the film depicts (Bordwell / Thompson 2008 [1997], 76; Hv.i.O.).

Anton Fuxjäger, Film- und Medienwissenschaftler an der Universität Wien, macht in seiner Interpretation des Souriau-Textes von 1951 allerdings darauf aufmerksam, dass sich der Diegesebegriff in der ursprünglichen, auf Souriau bzw. dessen Tochter zurückgehenden Definition „eindeutig und wiederholt auch auf die zeitlichen Aspekte der erzählten Welt und auf die in ihr vorkommenden Ereignisse oder Handlungen bezieht“ (Fuxjäger 2007, 21). Mit anderen Worten: von Souriau wird die erzählte Geschichte als eine Teilmenge der Diegese betrachtet. Könnte man also mit Christian Metz behaupten, dass der Diegesebegriff „im Grunde die Gesamtheit der filmischen Denotation“ (Metz 1972 [frz. 1968], 137f.) bezeichnet? Fuxjäger verneint dies:

Diese Formulierung entspricht insofern nicht Souriaus Auffassung, als so gut wie jeder Film auch Informationen denotiert, die nicht die erzählte Welt betreffen, etwa die im Vor- und Abspann eines typischen Films enthaltenen Informationen darüber, wer den Film produziert, inszeniert, geschnitten etc. hat. [Absatz, Anm. M] Zutreffender wäre es, die Diegese als die Gesamtheit aller Denotationen einer *Erzählung* zu bezeichnen (Fuxjäger 2007, 23).

Fuxjäger stellt diesen Vorschlag im Anschluss jedoch wieder in Frage, da er befürchtet, dass es dadurch nicht gelinge, ausreichend klarzumachen, dass es Souriau nicht um die Ebene des Diskurses gehe, sondern um „die Ebene des konkreten oder materiellen Daseins der Erzählung“ (ebd.). Diese Einschätzung Fuxjägers wird dem weiter oben wiedergegebenen Souriau-Zitat jedoch nicht gerecht, was daran liegen mag, dass sich Fuxjäger nicht auf den Originaltext, sondern auf die Übersetzung durch Kessler (vgl. Souriau 1997) bezieht.

⁷ Vgl. Genette (1998 [1994], 21), Metz (1972 [frz. 1968]; 137f., 144f., 175, 195f.), Kessler (2007, 11), Merlin (2010, 4-9).

⁸ Vgl. Genette (1998 [1994], 201; frz. Orig. 1983, 13).

⁹ Vgl. Genette (1998 [1994], 16; frz. Orig. 1972, 72).

Wie sich im Einzelnen nachweisen lässt, rückt in Kesslers Übersetzung – unter dem Einfluss einer simplifizierenden Umdeutung der Souriauschen Texte durch Christian Metz – die Zuschauerabhängigkeit des Diegetisierungsprozesses in den Hintergrund (vgl. Merlin 2010). Dadurch wird Souriau eine ähnlich starke Ontologisierung der diegetischen Realität unterstellt, wie diese in der Metzschen Konzeption von *Film als Text* zum Ausdruck kommt, in der sich Diegese *objektiv*, d.h. zuschauerunabhängig, als Funktion des Textes bestimmen lässt (vgl. Metz 1971, 14). In den Souriauschen Originaltexten (vgl. Souriau 1951, 1953 [1951]) findet sich jedoch die Beschreibung eines sensiblen Gleichgewichts zwischen einer in der Vorstellung des Zuschauers mehr oder weniger stark ontologisierten diegetischen Welt und dem zugleich bestehenden Wissen darüber, dass diese Ontologisierung durch einen *aktiven* Rezeptionsprozess auf Seiten der Zuschauer/innen zustande kommt, der eine fiktionale Realität schafft, welche einen grundlegend anderen Status besitzt als die Welt des Zuschauerraums.

Die Spannung zwischen diesen Polen, also einerseits dem von einer Vielzahl von Faktoren abhängigen Prozess der Konstituierung fiktionaler Welten und andererseits dem Resultat, dem Produkt dieses Prozesses (welches trotz der diesbezüglich bestehenden Analogien nicht mit den audiovisuellen Rohdaten identisch ist, die unsere Sinnesorgane im Kino aufnehmen), ist in den meisten weiteren Forschungsbeiträgen zum Diegesebegriff spürbar. So schlägt z.B. Hans Jürgen Wulff einerseits ein Schichtenmodell der Diegese vor, in dem die diegetische Realität *diegeseintern* in „vier miteinander koordinierte[...] Teilschichten“ (Wulff 2007, 40) zerlegt wird: in eine „*physikalische Welt*“, eine „*Wahrnehmungswelt*“, eine „*soziale Welt*“ und eine „*moralische Welt*“ (ebd., Hv.i.O.). Von diesem Modell, das sich primär auf das (fortwährend neu ausgebildete) Produkt des Konstituierungsprozesses bezieht, lässt sich der *diegeseextern* stattfindene Konstituierungsprozess selbst unterscheiden, den Wulff in Anlehnung an Odin (1983) als *Diegetisierung* bezeichnet. Bei Britta Hartmann (vgl. 2007, 61ff.) findet sich eine ähnliche Differenzierung, wenn sie die Diegese in ein *Diskursuniversum* einbettet. Der Diegetisierungsprozess wird in dieser Konzeption auf den audiovisuell vermittelten Diskurs verdichtet, d.h. er wird als primär durch den filmischen Text bedingt aufgefasst und weniger in Abhängigkeit von den Rezipient/innen gelesen: Laut Hartmann umfasst das Diskursuniversum außerhalb der diegetischen Welt angesiedelte „pragmatische[] Setzungen des Textes“ (ebd., 63), zu denen epistemologische, axiologische und moralische Postulate, Genreregeln, etc. gerechnet werden, die „„oberhalb“ der Welt der Figuren, jenseits ihrer Wahrnehmungsmöglichkeiten“ (ebd., 63) liegen.

Die sich im Verlauf der Wissenschaftsgeschichte etablierenden unterschiedlichen Verwendungsweisen des Diegesebegriffs kreisen um ein ähnliches Problemfeld, das auf der Basis unterschiedlicher erkenntnistheoretischer Verortungen jeweils aus unterschiedlichen Perspektiven betrachtet wird. Dieses Problemfeld lässt sich auf einer allgemeinen Ebene durch folgende Fragen charakterisieren:

1. Nach welchen subjektiven Vorgaben und nach welchen intersubjektiven Gesetzmäßigkeiten läuft der Prozess der *Diegetisierung* im Einzelfall ab?
2. Unter welchen Bedingungen verfügt eine diegetische Form der Realität über welchen ontologischen Status?
3. Welche am Prozess der Diegetisierung beteiligten Elemente sind selbst Teil der Diegese, welche der erstgenannten Elemente sind dies nicht?
4. Was bedeutet es genau, ‚Teil der Diegese‘ zu sein, d.h. wie funktionieren die Prozesse der Identifizierung und Analogiebildung, die diegetische Elemente mit nicht-diegetischen über Ähnlichkeitsstrukturen verbinden?
5. Welche diegetischen Elemente bzw. welche Formen dieser Elemente werden primär durch den filmischen Text gesetzt und welche werden eher durch die auf Seite der Rezipient/innen existierenden Wissensbestände und die dort stattfindenden perzeptiven, kognitiven, somatischen und emotionalen Vorgänge konstituiert, und wie funktionieren die wechselseitigen Abhängigkeiten zwischen diesen Faktoren?

2. Schnittmengen: Diegetic Sound

Eine breit angelegte, sowohl theoretisch fundierte als auch auf die Diskussion zahlreicher Beispiele aus der internationalen Filmgeschichte zurückgreifende Einführung zu den komplexen Beziehungen zwischen visuellen und akustischen Phänomenen im Spielfilm ist das bei den *Cahiers du Cinéma* publizierte Buch *Un art sonore, le cinéma. Histoire, esthétique, poétique* von Michel Chion (vgl. 2003, it. 2007, engl. 2009). Das Buch stellt eine Neubearbeitung und Erweiterung zweier anderer Bücher von Chion dar: *Le son au cinéma* (1985) und *La Toile trouée, ou la parole au cinéma* (1988). Chion stellt in Kapitel 15 des genannten Buches («Les trois frontières» [„Die drei Grenzen“], vgl. S. 221ff.) zunächst ein zweiteiliges Modell möglicher Ton-Bild-Relationen vor. Ein Sound werde häufig als ‚in‘ (engl. ‚on‘ oder ‚onscreen‘) bezeichnet, wenn die Quelle, durch die der Sound produziert wird, im Bild sichtbar ist. Ist die Quelle nicht sichtbar, rede man undifferenziert entweder von ‚off‘ oder ‚hors-champ‘ (= außerhalb des Bildfeldes, engl. ‚off‘ oder ‚offscreen‘). Chion präzisiert dieses zweiteilige Modell, indem er eine klare Unterscheidung der Begriffe *off* und *hors-champ* einführt:

[...] [D]ie Filmmusik beispielsweise, oder die Off-Stimme eines kommentierenden Erzählers haben nicht denselben Status wie die Musik eines die meiste Zeit unsichtbaren Pianisten, der der Nachbar der Figuren ist (ABSCHIED, DIE HÜNDIN, DAS FENSTER ZUM HOF), oder wie die Stimme einer Figur, die eine Veränderung der Kamerachse aus dem Bild geworfen hat, nicht jedoch aus der Handlung, und die

weiterhin spricht. Im letzten Fall bleibt das, was nicht sichtbar ist, in der Handlung verankert, im ersten gehört die unsichtbare Quelle des Sounds nicht zum selben Raum-Zeit-Gefüge (Chion 2003, 223; ÜB Merlin)¹⁰.

Chion ordnet nun dem erstgenannten Fall, bei dem der Sound nicht Teil der erzählten Welt ist, den Begriff *off* zu, dem zweiten dagegen, bei dem der Sound aus demselben Handlungsraum stammt, nur eben vorübergehend nicht sichtbar ist, den Begriff *hors-champ* – außerhalb des Bildfeldes. So gelangt er zu einem dreiteiligen Modell, das er als *tri-cercle* bezeichnet, da es sich als dreiteiliges Kreisdiagramm visualisieren lässt (vgl. ebd., 225). Die bereits erwähnten Begriffe *on*, *off* und *hors-champ* sind darin präzise definiert durch jeweils zwei semantische Kategorien, die in Form folgender Gegensatzpaare gegeben sind:

- *visualisé* vs. *acousmatique* (*visualisiert* vs. *akusmatisch*)
- *diégétique* vs. *non diégétique* (*diegetisch* vs. *Non-diegetisch*)

Es fällt auf, dass Chion die Begriffe *diegetisch* und *non-diegetisch* nicht definiert, auch nicht im Glossar film- und musikwissenschaftlicher Begriffe am Ende des Buches (vgl. ebd., S. 411-438), er scheint vorauszusetzen, dass die Leser/innen verstehen, was damit gemeint ist. Für den Begriff *akusmatisch* findet sich hingegen ein Eintrag im Glossar. Chion weist dort darauf hin, dass er den Begriff von Pierre Schaeffer (1952) entlehnt habe und definiert ihn wie folgt: „Die akusmatische Hörsituation ist jene, in der wir einen Sound hören, ohne die Ursache zu *sehen*, von der er hervorgebracht wird“ (Chion 2003, 411, Hv.i.O., ÜB Merlin)¹¹.

Chion erwähnt im Zusammenhang mit der Erläuterung des Triple-Modells (ebd., 224), dass seine Übersetzerin Claudia Gorbman¹² in der englischen Übersetzung seines Buches *L'audio-version: Son et image au cinéma* (1990, engl. 1993) die drei zentralen Begriffe des Modells wie folgt übersetzt: „[...] ‚in‘ durch *onscreen*, ‚hors-champ‘ durch *offscreen*, und ‚off‘ durch *nondiegetic*.“ (Chion 2003, 224; Hv.i.O., ÜB M)¹³. Chions beiläufiger Kommentar zur englischen Übersetzung ist meiner Ansicht nach nur scheinbar nebensächlich. Die Diskrepanz zwischen der französischen Terminologie und ihrer englischen Übersetzung macht auf ein generelles Problem aufmerksam, dem Chion nicht Rechnung trägt: wenn man die zwei Gegensatzpaare, die er als Definitionskriterien verwendet, miteinander kombiniert, entsteht eigentlich ein Quadrupel- statt ein Tripel-Modell. Ersteres sei hier in tabellarischer Form dargestellt:

¹⁰ « [...] [L]a musique du film, par exemple, ou la voix off du commentateur-narrateur n'ont pas le même statut que la musique du pianiste la plupart du temps invisible qui est le voisin des personnages (*Abschied, La Chienne, Fenêtre sur cour*) ou que la voix du personnage que le changement d'axe de la caméra a rejeté en dehors du champ, mais non de l'action, et qui continue de parler. Dans ce cas-ci, ce qui n'est pas visible demeure situé dans l'action, et dans ce cas-là, la source invisible du son n'appartient pas au même espace-temps » (Chion 2003, 223).

¹¹ « La situation d'écoute acousmatique est celle où l'on entend le son sans *voir* la cause dont il provient [...] » (Chion 2003, 411, Hv.i.O.).

¹² Laut Christine Noll Brinckmann (2007, 74) ist Claudia Gorbman die erste Filmwissenschaftlerin, die die von Genette entwickelte Ausdifferenzierung des Souriauschen Diegesebegriffs zur Klassifizierung des Filmtons verwendet. Zu Gorbman (1976) vgl. S. 26ff.

¹³ « [...] ‚in‘ par *onscreen*, ‚hors-champ‘ par *offscreen*, et ‚off‘ par *nondiegetic* » (Chion 2003, 224; Hv.i.O.).

	visualisé	acousmatique
diégétique	in (engl. onscreen)	hors-champ (engl. offscreen)
non diégétique		off (engl. nondiegetic)

Gorbmann macht sich in der Übersetzung zunutze, dass ein Ton, der sich nicht im raumzeitlichen Gefüge der Figurenhandlung verorten lässt und daher als *nondiegetic* klassifiziert wird, nach Chions Modell automatisch auch nicht sichtbar, d.h. *akusmatisch* ist – dies setzt Chion als selbstverständlich voraus. Da es also nach Chions Modell für einen *non-diegetischen* Ton nur die Möglichkeit gibt, gleichzeitig auch *akusmatisch* zu sein, verwendet Gorbmann für diesen Fall den Begriff *non-diegetisch*. In diesem Punkt ist daher nicht, wie man zunächst denken könnte, Gorbmanns Übersetzung ungenau, sondern das Modell von Chion. Der Fall, dass der Sound eines Films zugleich *non-diegetisch* und *visualisiert* ist, ist bei Chion von vornherein ausgeschlossen, er existiert aber durchaus, nämlich dann, wenn sich die Quelle des Sounds nicht dem Handlungsraum der erzählten Geschichte zuordnen lässt, sondern einem anderen Handlungsraum, dessen Darstellung die eigentliche Figurenhandlung kurzfristig unterbricht. Letzteres kann mit dem Ziel geschehen, die Primärhandlung in verschiedene Abschnitte zu untergliedern oder auch, um einen *non-diegetischen* Erzählerkommentar zu illustrieren. Beides lässt sich z.B. in Fatih Akins GEGEN DIE WAND (Deutschland/Türkei 2004, 116 min) beobachten. Die Primärhandlung des genannten Films ist wie in klassischen Dramen in fünf Akte unterteilt. Zwischen den einzelnen Akten tritt der türkische Klarinettist Selim Sesler mit einem Orchester am Ufer des Bosphorus auf. Sie begleiten die Berliner Schauspielerin, Sängerin und Regisseurin Idil Üner, die die Volkslieder *Saniye'm (Meine Saniye)*, *Penceresi Yola Karsi (Ihr Fenster geht zur Straße)* und *Su karsiki dagda bir fener yanari (Dort drüben auf dem Berg brennt ein Leuchtfeuer)* singt. Die Lieder thematisieren das Glück und die Schmerzen, die mit der Liebe verbunden sein können.¹⁴ Diese Musik ist eher Teil des narrativen Diskurses als der Diegese. Sie verfügt über eine *visualisierte* Quelle, und dennoch ist sie relativ zur Primärhandlung *non-diegetisch*, da sie sich nicht in deren raumzeitliches Koordinatensystem einfügen lässt: es kommt während des Films niemals zu einer Überschneidung beider Handlungsräume.

David Bordwell und Kristin Thompson verwenden in ihrem bereits erwähnten Buch *Film Art. An Introduction* dieselben inhaltlichen Unterscheidungskriterien wie Chion, allerdings ordnen sie den vier Möglichkeiten, die sich durch die wechselseitige Kombination dieser Kriterien ergeben, keine neuen Begriffe zu (vgl. 2008 [1997], 278f.). Tabellarisch ergibt sich hier die folgende Übersicht (in diese seien die von Chion verwendeten Begriffe in eckigen Klammern eingefügt, um die Unterschiede zwischen beiden Modellen deutlich zu machen):

¹⁴ Vgl. die Liedtexte in Wienen / Twele (2004, 21): <http://www.bpb.de/files/CDVFQZ.pdf> (Stand: 28.03.10).

	onscreen [visualisé]	offscreen [acousmatique]
diegetic [diégétique]	[in (engl. onscreen)]	[hors-champ (engl. offscreen)]
non diegetic [non diégétique]		[off (engl. nondiegetic)]

Bordwells und Thompsons System besitzt den Vorteil, dass sich nun eine größere Anzahl von Beispielen einbeziehen lässt, eben auch der bei Chion vernachlässigte Fall eines Sounds, der sich im Rahmen der Bordwell-Thompsonschen Terminologie als *nondiegetic onscreen* kategorisieren ließe (wie das genannte Beispiel aus GEGEN DIE WAND). Das Klassifizierungsverfahren ist nun relativ einfach zu handhaben: erstens muss man prüfen, ob die Quelle eines Sounds im raumzeitlichen Gefüge der erzählten Geschichte verankert ist oder nicht, und zweitens, ob diese Quelle visualisiert ist oder nicht. In den Worten von Bordwell und Thompson:

Sound has a spatial dimension because it comes from a *source*. Our beliefs about that source have a powerful effect on how we understand the sound. [...] For purposes of analysing narrative form, we described events taking place in the story world as *diegetic* (p. 76). For this reason, **diegetic sound** is sound that has a source in the story world. The words spoken by the characters, sounds made by objects in the story, and music represented as coming from instruments in the story space are all diegetic sound (Bordwell / Thompson 2008 [1997], 278; Hv.i.O.).

Alternatively, there is **nondiegetic sound**, which is represented as coming from a source outside the story world. Music added to enhance the film's action is the most common type of nondiegetic sound. When Roger Thornhill is climbing Mount Rushmore in *North by Northwest* and tense music comes up, we don't expect to see an orchestra perched on the side of the mountain. Viewers understand that movie music is a convention and does not issue from the world of the story. The same holds true for the so-called omniscient narrator, the disembodied voice that gives us information but doesn't belong to any of the characters in the film. An example is *The Magnificent Ambersons*, in which the director, Orson Welles, speaks the nondiegetic narration (Bordwell / Thompson 2008 [1997], 279; Hv.i.O.).

[...] [D]iegetic sound can be either onscreen or offscreen, depending on whether the source is inside the frame or outside the frame (Bordwell / Thompson 2008 [1997], 279; Hv.i.O.).

Die beiden von Bordwell und Thompson hier genannten Unterscheidungen *diegetic* vs. *nondiegetic* und *onscreen* vs. *offscreen* seien im Folgenden einer kritischen Betrachtung unterzogen. Beginnen wir mit dem ersten Oppositionspaar. Bordwell und Thompson betonen, dass die diegetische Realität grundsätzlich eine andere sei, als die außerfilmische, in der der Film produziert wurde:

[...] [T]he distinction between diegetic and nondiegetic sound doesn't depend on the real source of the sound in the filmmaking process. Rather, it depends on our understanding of the conventions of film viewing. We know that certain sounds are represented as coming from the story world, while others are represented as coming from outside the space of the story events. Such viewing conventions are so common that we usually do not have to think about which type of sound we are hearing at any moment (Bordwell / Thompson 2008 [1997], 279).

Es geht also bei der Unterscheidung zwischen diegetischem und non-diegetischem Sound nicht darum, die im außerfilmischen Sinn realen Quellen zu lokalisieren, die diese Sounds im Verlauf des Produktionsprozesses hervorgebracht haben, sondern darum, ob sich in der Vorstellung der Zuschauer/innen die Quelle eines Sounds dem Raum-Zeitgefüge der erzählten Geschichte *zuordnen* lässt oder nicht, wobei dieser Zuordnungsprozess im Fall der Verwendung bestimmter Konventionalisierungen quasi automatisiert sein kann.¹⁵ Wie weiter oben erwähnt, geht diese Definition, wissenschaftshistorisch betrachtet, auf Etienne Souriau (bzw. auf dessen Tochter Anne) zurück.¹⁶ Entgegen der Interpretationen des Souriauschen Diegesebegriffs durch Metz (vgl. 1972 [frz. 1968], 137f.), Genette (vgl. 1998 [1994], 201) oder Kessler (vgl. Souriau 1997) stellen sowohl die Rezipientenabhängigkeit als auch die Prozesshaftigkeit der Diegesekonstituierung bereits wesentliche Merkmale der Souriauschen Definition dar – auch wenn Souriau die Hypothese vertritt, dass sich die intersubjektive Schnittmenge der Resultate der einzelnen Rezeptionsvorgänge experimentell bestimmen lässt.

Wie lässt sich nun entscheiden, ob eine Zuordnung der Sound-Quelle zum raumzeitlichen Gefüge der Figurenhandlung möglich ist oder nicht, wenn man im Rekurs auf die in dieser Hinsicht übereinstimmenden Souriauschen und Bordwell-Thompsonschen Definitionen beachtet, dass die *Diegese* nicht als eigenständige Realität auf der Leinwand vorliegt, sondern erst in einer – einleitend auch von Kohlhaase konstatierten – *Interaktion* zwischen audiovisuellem Input und Rezipient/innen konstituiert werden muss? Dazu schreibt Anton Fuxjäger:

[...] [J]ene immer wieder referierte Faustregel zur Unterscheidung von diegetischen und nicht-diegetischen Elementen des Plots/Diskurses funktionieren nicht, sofern man bei ihrer Anwendung exakt genug vorgeht: „Nicht-diegetische Elemente des Films sind definiert als solche, die von einer Figur des Films nicht wahrgenommen werden können“ (Rother 1997, 58). Ich würde meinen: Üblicherweise gehen wir davon aus, dass die Figuren in einem Film eben diesen Film prinzipiell nicht wahrnehmen können, dementsprechend auch keinen seiner Teile oder Elemente [...]. Zu sagen, dass irgendetwas von

¹⁵ In LISBON STORY (D 1994, Wim Wenders, 100 min, Musik: Jürgen Knieper / Madredeus) findet sich eine exzellente Veranschaulichung dieses Zusammenhangs. Dort gibt es eine Szene, in der ein Toningenieur (Rüdiger Vogler), der für eine Spielfilmproduktion O-Töne in Lissabon sammeln soll, dabei beobachtet wird, wie er eine ganze Reihe von ‚natürlich‘ wirkenden Sounds mit Hilfsmitteln erzeugt, die man, wenn der Produktionsvorgang nicht visualisiert wäre, niemals als Quellen der entsprechenden Sounds akzeptieren würde (vgl. LISBON STORY, 00:20:34-00:24:18). Die Szene ist eine Hommage an den Beruf des ‚Geräuschemachers‘ (engl. ‚Foley artist‘), dessen Tätigkeit bei vielen Spielfilmproduktionen unverzichtbar ist, vgl. z.B. *The Story of Jack Foley* auf der Website des Sound-Editors Philip Rodrigues Singer: <http://www.marblehead.net/foley/index.html> (Stand: 26.03.2010).

¹⁶ Vgl. die Souriausche Diegese-Definition auf S. 3.

dem, was wir im Kino zu sehen oder hören bekommen, ‚diegetisch‘ sei, ist insofern falsch im Sinne der Urheber des Konzepts, als die ‚Diegese‘ (bei fiktionalen Erzählungen) ja nur in unserer *Vorstellung* existiert. – Im *Kino* haben wir es lediglich mit Zeichen zu tun, die auf die Diegese verweisen oder genauer gesagt: uns dazu veranlassen, eine Vorstellung der Diegese zu konstruieren (Fuxjäger 2007, 26f.; Hv.i.O.).

Im Folgenden wirft Fuxjäger die Frage auf, was genau die Phänomene, die Hartmann (2007) auf der Ebene eines *Diskursuniversums* bzw. Wulff auf einer „Diegese ‚zweiter Ordnung‘“ (Wulff 2001, zit. nach Hartmann 2007, 63, Anm. 19) verorten würde, von solchen Phänomenen unterscheidet, die in der Vorstellung der Zuschauer/innen von einer Wahrnehmungsrealität der Figuren (also einer „Diegese ‚erster [...] Ordnung‘“, ebd.) zugeordnet werden: „Inwiefern unterscheidet sich die Art und Weise, in der z.B. *superimposed titles* über die erzählte Welt informieren, von jener, in der dies die bewegten Bilder von Figurenhandlungen tun?“ (Fuxjäger 2007, 27; Hv.i.O.). Fuxjäger zögert nicht mit der Antwort:

Die üblicherweise als ‚diegetisch‘ klassifizierten Elemente unterscheiden sich von jenen, die zwar ebenfalls über die Diegese informieren, aber dennoch nicht als diegetisch gelten, durch den Umstand, dass sie auf nachahmende Weise über die erzählte Welt informieren. [...] Zwischen den bewegten Bildern von den Figurenhandlungen, die uns auf der Leinwand begegnen, und den Vorgängen in der erzählten Welt, auf die sie verweisen (sollen), besteht eine große Ähnlichkeit. ‚Nicht-diegetische Elemente‘ des Plots/Diskurses wie etwa *superimposed titles* oder ‚nicht-diegetische‘ Musik informieren zwar ebenfalls über die erzählte Welt, tun dies aber nicht, indem sie Aspekte der erzählten Welt nachahmen (Fuxjäger 2007, 27; Hv.i.O.).

Fuxjäger macht also die Einschätzung des Faktums, ob es sich eher um „mimetische[...]“ oder eher um „nicht-mimetische[...] Informationen“ (ebd., 31) über die im raumzeitliche Gefüge der erzählten Geschichte situierten Objekte, Figuren und Ereignisse handelt, zum Definitionskriterium dafür, ob letztere als *diegetisch* oder *nondiegetisch* aufzufassen sind. Auf den überwiegend mimetischen Charakter der Filmerzählung weist (im Kontext einer Untersuchung der Zusammenhänge zwischen Lichtgestaltung und Diegesebildung) auch die inzwischen emeritierte Züricher Filmwissenschaftlerin Christine Noll Brinckmann hin:

Während es im Roman nicht zur Verwechslung von Diegese und Gestaltung kommt, ist dies bei der visuellen filmischen Präsentationsweise leicht gegeben. Im Grunde beschreibt der Film seine Ereignisse bereits ‚im Gewand der Diegese‘, Signifikant und Referent, Darstellung und Dargestelltes sind – jedenfalls was die äußere Erscheinung der Dinge betrifft – weitgehend kongruent. Auch sind viele Gestaltungsweisen so transparent oder konventionalisiert, dass die Zuschauer gleichsam durch sie hindurch direkt auf die fiktionale Welt zu blicken glauben. Erst die genauere Analyse vermag einen Keil zwischen die verschiedenen Schichten zu treiben (Noll Brinckmann 2007, 73).

Für die „visuelle[...] filmische[...] Präsentationsweise“ (ebd.) mag diese Auffassung zutreffen – auch wenn zu berücksichtigen ist, dass sie sich in erster Linie auf den fiktionalen Mainstreamfilm bezieht, weniger auf experimentellere Filme. Gilt die für den Mainstreamfilm konstatierte „weitgehend[e] [K]ongruen[z]“ (ebd.) von Darstellung und Dargestelltem jedoch auch für akustische Phänomene, für den Sound? Diese Frage soll dann, wenn es um die Diskussion der Beispielsequenzen geht, ins Zentrum der Aufmerksamkeit gerückt werden.

Wie steht es nun um die Abgrenzungskriterien zwischen „mimetische[n] und nicht-mimetische[n] Informationen über die Diegese“ (Fuxjäger 2007, 31)? Die Frage, die in Fuxjägers Überlegungen eine entscheidende Rolle spielt, nämlich ob das Gegensatzpaar *mimetisch* vs. *nicht-mimetisch* eine präzisere Beschreibung der Phänomene erlaubt, die zur Diegesekonstitution durch den Zuschauer beitragen, als das Begriffspaar *diegetisch* vs. *nicht-diegetisch*, wird von Fuxjäger auf widersprüchliche Weise beantwortet. Anfangs zieht Fuxjäger eine klare Grenze zwischen *mimetischen* und *nicht-mimetischen* Informationen über die erzählte Welt:

Die schriftlichen Prologe am Beginn der STAR-WARS-Filme haben keine Ähnlichkeit mit den diegetischen Ereignissen, von denen sie berichten; die bedrohliche Musik, die vor einer Mordszene ertönt und uns darüber informiert, dass Gefahr lauert, hat keinerlei Ähnlichkeit mit dieser Gefahr oder sonst einem akustischen Ereignis in der erzählten Welt, während ‚diegetische Musik‘ große Ähnlichkeit mit musikalischen Ereignissen in der erzählten Welt hat. [Absatz, Anm. M] Man könnte hier mit Peirce von ikonischen im Unterschied zu symbolischen Informationen über die Diegese sprechen. Die ‚diegetischen‘ Elemente des Plots/Diskurses sind – in der hier zur Debatte stehenden Hinsicht – *icons*: Sie repräsentieren die entsprechenden Dinge und Vorgänge in der erzählten Welt „by [...] similarity“ (Peirce 1998, 276). Die ‚nicht-diegetischen‘ Elemente sind dagegen als Symbole zu klassifizieren: Die Verbindung zwischen *superimposed titles*, Stimmen von *noncharacter narrators* und nicht-diegetischer Musik mit den entsprechenden diegetischen Zusammenhängen besteht aufgrund von Konventionen der Verbalsprache, des typischen Einsatzes bestimmter Musik etc (Fuxjäger 2007, 27f.; Hv.i.O.).

Im weiteren Verlauf seiner Argumentation weicht Fuxjäger die hier dargestellten Grenzen zwischen *mimetischen* und *nicht-mimetischen* Informationen allerdings wieder auf, indem er feststellt, dass es auch Zwischenstufen gibt, beispielsweise Zeitungsschlagzeilen, die die Rezipient/innen einerseits auf *nicht-mimetische* Art über die Figurenhandlung informieren, andererseits aber bezüglich ihrer materiellen Beschaffenheit eine starke Ähnlichkeitsbeziehung zu einem Element der erzählten Welt aufweisen, eben in ihrer Eigenschaft als Zeitungsschlagzeilen. Fuxjäger gelangt daher zu einer Definition des Begriffs *diegetisch*, die zwar unter dem Aspekt, dass sie nicht a priori *filmische Diegese* und *filmischen Diskurs* zusammenfallen lässt, relativ präzise ist, die jedoch im Gegenzug keine klaren Abgrenzungskriterien zwischen *mimetischen* und *nicht-mimetischen* Informationen über die Diegese angibt:

Als diegetisch wird nicht nur alles bezeichnet, was zur erzählten Welt gehört, sondern auch all jene Anteile des Plots/Diskurses, die den Rezipienten [...] auf nachahmende Weise über die erzählte Welt informieren, wobei diese Nachahmung mehr oder weniger treu und ein Element des Plots/Diskurses daher auch mehr oder weniger diegetisch sein kann (Fuxjäger 2007, 32f).

Hier stellt sich die Frage, worauf sich der Akt der Nachahmung eigentlich bezieht, d.h. was genau durch den audiovisuellen Text nachgeahmt bzw. durch die Rezipient/innen als Nachahmung eingeschätzt wird. Bei fiktionalen Welten muss dies nicht zwangsläufig die äußere Wahrnehmungsrealität der Zuschauer/innen sein:

Das Diegetische muss nicht mit der normalen Wahrnehmungserfahrung des Zuschauers übereinstimmen. Wenn etwa der Held in *DIE HARD* (*STIRB LANGSAM*, 1988, John McTiernan) barfuß über ein Feld zersplitterten Glases laufen muss, kann der Zuschauer dessen Empfindungen nur empathisch simulieren – und er darf erstaunt darüber sein, wie schmerzempfindlich der Held damit umgeht. Eine derartige Differenz von ‚wahrscheinlichem Wahrnehmungseindruck‘ einer Figur und dem Horizont der ‚zuschauereigenen Körpererfahrung‘ wird gerade in komischen Genres zum Vergnügen des Publikums lustvoll inszeniert. Wahrnehmungswelt der Zuschauer und Wahrnehmungswelt der Figuren sind dann nicht deckungsgleich, sondern Teil der Inszenierung wie Teil einer spezifischen Distanz, die zwischen Diegese und Zuschauer aufgerichtet wird (Wulff 2007, 42; H.v.i.O.).

Auf ähnliche Weise äußert sich auch Hartmann:

Die Konstruktion einer erzählten Welt erfordert im Übrigen keineswegs, dass die Bilder und Töne in einem Abbildungs- oder auch nur Ähnlichkeitsverhältnis zu Erscheinungen unserer Realität stehen (vgl. ähnlich Koch 2003, 167; 179): Auch Animationsfilme bilden Diegesen aus. [...] Diegesen müssen nicht der Realität des Zuschauers ähneln, darauf beruht oft sogar ihr Reiz, erst von daher können sie eine besondere Erfahrung vermitteln. Gertrud Koch spricht von einer „autonome[n] filmische[n] Welt“ (ibid., 163), der wir gegenübergestellt sind, und Wulff betont mit Michotte van den Berg (2003) die „Eigen- und Selbständigkeit der Diegese gegenüber der primären Realität des Zuschauers“ [...]. [...] Für die Imagination einer solchen fiktionalen Realität ist es unerheblich, ob diese ‚realistisch‘ ist oder nicht [...]. Von daher können wir selbst Unwahrscheinliches oder nach unseren Maßstäben Unmögliches als konstituierend für die erzählte Welt akzeptieren, z.B. das unerklärliche Auftauchen des Monolithen in Stanley Kubricks *2001: A SPACE ODYSSEY* (USA/GB 1968), die Existenz von Geistern, Monstern und Untoten im Horror-Film, von sprechenden Tieren oder Teekesseln im Märchen, die schiere Unzerstörbarkeit des menschlichen Körpers im Action-Kino oder auch das In-Geltung-Stehen anderer gesellschaftlicher Werte und Moralvorstellungen (Hartmann 2007, 58).

Bezieht man diese Aussagen auf Fuxjägers oben zitierte Diegese-Definition, so wird deutlich, worin das grundlegende Problem bei der Suche nach einer möglichst exakten Prädikatierung durch den Begriff

diegetisch besteht, ein Problem, das ich als *diegesetheoretisches Paradoxon* bezeichnen möchte: Um zu entscheiden, ob ein Phänomen mehr oder weniger *mimetisch* und daher mehr oder weniger *diegetisch* im Sinne Fuxjägers ist, muss ein Vergleich zwischen diesem Element und anderen Elementen der Diegese vorgenommen werden; um diesen Vergleich aber vornehmen zu können, muss zumindest für die anderen Elemente, mit denen dieser Vergleich vorgenommen wird, die Diegese bereits konstituiert sein. Mit anderen Worten: Man kann nicht entscheiden, ob ein Element eher zur erzählten Welt gehört oder eher zum filmischen Diskurs, der dabei hilft, diese Welt zu konstituieren, wenn man dieses Element nicht mit etwas vergleichen kann, was bereits klar der erzählten Welt zuzuordnen ist.

Bei Filmen, die versuchen, bezogen auf audiovisuelle Oberflächenstrukturen möglichst *naturalistisch* die äußere Wahrnehmungsrealität der Zuschauer/innen abzubilden, ist die Diegesekonstituierung aufgrund starker Ähnlichkeitsbeziehungen zwischen *außerfilmischer* (in der Souriauschen Terminologie: *afilmischer*) Realität, Diskursebene (nach Souriau: *filmophanischer* Realität) und *diegetischer* Realität¹⁷ weniger ein Problem als dann, wenn die Beschaffenheit der diegetischen Welt stark von der außerfilmischen Realität abweicht oder wenn Diskursformen eingesetzt werden, die weniger konventionalisiert sind, die noch nicht oder nicht mehr den Seh- bzw. Hörgewohnheiten des Zuschauers entsprechen oder die aus anderen kulturellen Kontexten stammen. Um im Verlauf der Filmrezeption bzw. der Konstituierung einer narrativen Wirklichkeit trotz des geschilderten diegesetheoretischen Problems dennoch relativ stabile diegetische Welten ausbilden zu können, finden Diegetisierungsprozesse fortwährend sowohl in einem *horizontalen*, als auch – im Rahmen der durch die audiovisuellen Wahrnehmungsdaten gegebenen Anhaltspunkte – in einem *vertikalen* Abgleich von Ähnlichkeitsstrukturen statt: *horizontal* insofern, als es um fikcionalitätsspezifische Regelmäßigkeiten bzw. logische Kohärenzen innerhalb der diegetischen Realität geht, *vertikal* im Hinblick auf Analogien zwischen außerfilmischen, diskursiven und diegetischen Phänomenen. Die genannte Stabilisierung der erzeugten narrativen Welten kann dabei durch manche Filme bewusst unterstützt, durch manche jedoch auch gezielt konterkariert werden. Als aktuelles Beispiel für den letztgenannten Fall vgl. z.B. SHUTTER ISLAND (USA 2010, Martin Scorsese, 138 min, Musik: Robbie Robertson), einen Film, in dem über weite Strecken des Handlungsverlaufs keine Gewissheit darüber besteht, ob der den Zuschauer/innen präsentierte audiovisuelle Input als intersubjektiv wahrnehmbare äußere Realität der Figuren zu lesen ist oder als Halluzination des Protagonisten.

Nun zum zweiten Oppositionspaar, das Bordwell und Thompson zur Klassifizierung von filmischen Soundphänomenen relativ zum Geschehen auf visueller Ebene einsetzen: der Differenzierung *onscreen* vs. *offscreen*. Hier noch einmal das entscheidende Zitat: „[...] [D]iegetic sound can be either onscreen or

¹⁷ Souriau unterteilt ein *allgemeines filmisches Universum* in sieben unterschiedliche Realitätsebenen, wobei der Begriff des *Universums* hier nicht, wie häufig missverstanden wird, in einem *raumzeitlichen*, sondern in Anlehnung an den englischen Mathematiker und Logiker Augustus De Morgan (1806-1871) in einem *logischen* Sinn zu verstehen ist. Neben den erwähnten *afilmischen*, *filmophanischen* und *diegetischen* Realitäten lassen sich laut Souriau als Teilmengen des genannten Gesamtphänomens eine *profilmische*, *filmographische*, *spektatorielle* sowie eine *kreatorielle* Wirklichkeitsebene unterscheiden, vgl. Souriau (1951, 1953 [1951]).

offscreen, depending on whether the source is inside the frame or outside the frame” (Bordwell / Thompson 2008, 279).

Die Züricher Medienwissenschaftlerin und ehemalige Filmtonemeisterin Barbara Flückiger (Autorin des Buches *Sound Design. Die virtuelle Klangwelt des Films*, vgl. 2007 [2001]) kritisiert diese von Bordwell und Thompson vertretene Unterscheidung unter Verweis auf Untersuchungen von Jörg Tüschmann (1994) und Laurent Jullier (1995) als zu stark auf ein Primat visueller Phänomene bezogen:

Die Problematik rührt von einer Grundeigenschaft akustischer Ereignisse her: Sie sind invasiv und ubiquitär, sie durchdringen Mauern und gehen um Ecken. In vielen Fällen lässt sich nicht entscheiden, ob die Klangobjekte ihre Quelle im Bild haben oder nicht. Mehr noch, diegetische Klangobjekte befinden sich auch dann im abgebildeten Raum, wenn ihre Quellen außerhalb des Bildes situiert sind, insofern sie die Figuren ebenfalls hören können. Mit anderen Worten: der Off-Ton ist ein weites Feld (Flückiger 2005, 142).

Flückiger schlägt im Folgenden eine operationale Terminologie vor, die nicht mehr rein zu Definitionszwecken daran gebunden ist, ob sich ein Sound innerhalb oder außerhalb des Bildrahmens befindet. Sie stellt Musik und Figurenäußerungen sowie Geräuschen, die sich direkt auf die Figurenhandlung beziehen, Geräusche zur Seite, die „zur Charakterisierung eines Ortes dien[en], zum Beispiel [...] Hundegebell“ (Flückiger 2005, 144). Es ist für diese Art von Geräuschen, die sie in Anlehnung an den von R. Murray Schafer (1994 [1977]) geprägten Begriff *sound mark* als *Orientierungslaute* bezeichnet, unerheblich, wo genau sich die Soundquelle befindet:

Der Begriff *Orientierungslaut* bezeichnet nicht ein spezifisches Klangobjekt an sich, sondern dessen Funktion, einen Ort geographisch, zeitlich, kulturell, ethnisch oder sozial zu definieren. Ein Hundegebell – um beim einfachsten Beispiel zu bleiben – kann auch die Quelle, einen bestimmten Hund nämlich, bezeichnen und hat dann eine indexalische Funktion. Die Unterschiede zwischen den beiden Funktionen werden klanglich markiert. Das Hundegebell mit indexalischer Funktion befindet sich im Vordergrund, während das Hundegebell mit Orientierungsfunktion via verminderte Lautstärke und zusätzliche Raumparameter – insbesondere Hall – in den Hintergrund verlegt wird (Flückiger 2005, 144; H.v.i.O.).

Systemisch organisierte Orientierungslaute bilden nach Flückiger größere Einheiten aus, die sie als *Atmosphären* bezeichnet:

Die einzelnen Orientierungslaute sind auf der Tonspur in spezifischen, übergeordneten Strukturen organisiert. Als Begriff für diese Substrukturen schlage ich den bereits bestehenden Terminus *Atmosphäre* vor. Im Technikerjargon und für die Organisation von Geräuscharchiven hat sich der

Begriff etabliert, um ganze Lautsphären sprachlich zu fassen, wie zum Beispiel *Hafen, Bahnhof, Bergwiese*. [...] [Atmosphären] sind das akustische Setting, das sowohl die raumzeitliche Orientierung ermöglicht als auch den emotionalen Rahmen nicht nur für die Zuschauer, sondern auch für die Filmfiguren abgibt (Flückiger 2005, 145; Hv.i.O.).

Flückigers pragmatischer Ansatz zeichnet sich dadurch aus, dass er auf eine relativ anschauliche Weise zugleich produktionstechnischen, als auch rezeptionsästhetischen Überlegungen Rechnung trägt. Es würde u.U. zu interessanten Ergebnissen führen, die hohe Operationalität, die die Begriffe *Sound mit indexalischer Funktion, Orientierungslaut* und *Atmosphäre* in Bezug auf die Konstituierung diegetischer Realitäten aufweisen, mit einer Ausdifferenzierung der Fuxjägerschen Idee, Diegetisierungsprozesse als Hypothesenbildungen zu beschreiben, die sich auf die fortwährende Analyse von Analogiestrukturen stützt, zu kombinieren. Dabei wären Fragen zu klären, die mit dem oben skizzierten diegesetheoretischen Dilemma zusammenhängen, d.h. es müsste berücksichtigt werden, dass die genannten Analogiestrukturen sowohl in vertikaler, als auch in horizontaler Hinsicht auftreten können.

3. Internal diegetic Sound

Aus der Perspektive von Bordwell und Thompson ließe sich auf die insbesondere von Fuxjäger und Flückiger geäußerte Kritik an den bisher vorgestellten traditionellen Unterscheidungskategorien entgegen, dass diese Kategorien dennoch auf eine Vielzahl von Filmen anwendbar seien. Phänomene, die mit der von Chion (2003) bzw. der von Bordwell / Thompson (2008 [1997]) verwendeten Terminologie (vgl. die entsprechenden tabellarischen Übersichten) nicht eindeutig zu beschreiben sind, werden in dieser Sichtweise als Resultate eines Spiels mit sonst relativ stark etablierten Konventionen betrachtet:

At many times, however, a film's narration deliberately blurs boundaries between different spatial categories. Such a play with convention can be used to puzzle or surprise the audience, to create humor or ambiguity, or to achieve other purposes (Bordwell / Thompson 2008 [1997], 279).

Chion argumentiert hier ähnlich. Er gibt zu, dass es Fälle gibt, die sich mit seinem Triple-Modell nicht klassifizieren lassen. Während Bordwell und Thompson allerdings von *häufigen* Ausnahmen reden (s.o.), sind Chions Ansicht nach sind diese Fälle marginal:

Wir ignorieren nicht, dass man in den Unterscheidungen weiter gehen kann. Sie decken nicht alle Fälle ab (z.B. nicht denjenigen des internen Sounds), aber wir schätzen, dass diese drei Kategorien die Analyse der am häufigsten auftretenden Fälle ermöglichen (Chion 2003, 224; ÜB M).¹⁸

¹⁸ « Nous n'ignorons pas qu'on peut aller plus loin dans les distinctions, qui ne couvrent pas tous les cas (par exemple celui du son interne), mais nous estimons que ces trois catégories permettent d'analyser les cas les plus fréquents » (Chion 2003, 224).

Allerdings weist Chion im gleichen Atemzug darauf hin, dass es Fragen gibt, die von zu großer Relevanz sind, als dass man sie komplett aus der Betrachtung ausschließen könnte, nur weil sie sich nicht so leicht mit den beschriebenen Kategorien klassifizieren lassen:

Diese dreifache Differenzierung ermöglicht die Etablierung einer ersten, groben und provisorischen Typologie dieser räumlichen und zeitlichen Beziehungen. Man sollte allerdings später nicht vergessen, dem inneren Monolog einer Figur (der inneren Stimme) einen Platz zu geben, sowie der Stimme eines Schauspielers, der sich innerhalb des Bildfeldes befindet, uns aber den Rücken zudreht, oder auch den Stimmen unsichtbarer Gespenster [...] Und gilt dasselbe nicht auch für das Geräusch des Windes, wenn sich nichts bewegt, was ihn visualisieren würde, wir ihn aber trotzdem hören? (Chion 2003, 224; ÜB M)¹⁹

Um Soundphänome klassifizieren zu können, die sich auf innere Vorgänge von Figuren beziehen, schlagen Bordwell und Thompson eine weitere Binnendifferenzierung vor: sie unterscheiden *internal* von *external diegetic sound* bzw. *subjective* von *objective diegetic sound*. Mit Hilfe dieser Differenzierung lassen sich nun Fälle wie der von Chion erwähnte innere Monolog einordnen:

Often a filmmaker uses sound to represent what a character is thinking. We hear the character's voice speaking his or her thoughts even though that character's lips do not move; presumably, other characters cannot hear these thoughts. Here the narration uses sound to achieve subjectivity, giving us information about the mental state of the character. Such spoken thoughts are comparable to mental images on the visual track. A character may also remember words, snatches of music, or events as represented by sound effects. In this case, the technique is comparable to a visual flashback. [Absatz, Anm. M] The use of sound to enter a character's mind is so common that we need to distinguish between internal and external diegetic sound. **External diegetic sound** is that which we as spectators take to have a physical source in the scene. **Internal diegetic sound** is that which comes from inside the mind of a character; it's subjective. Nondiegetic and internal diegetic sounds are often called **sound over** because they do not come from the real space of the scene. Internal diegetic sound can't be heard by other characters (Bordwell / Thompson 2008 [1997], 284; Hv.i.O.).

Diese auf den ersten Blick sinnvoll erscheinende Unterscheidung zieht ein größeres Problem nach sich: wenn für einen *internal diegetic sound* die Quelle des Sounds nicht in der auf diegetischer Ebene als *real* eingeschätzten äußeren Wahrnehmungswelt der Figuren verortet ist, dann fragt sich, ob man ihn nicht als *non-diegetisch* bezeichnen müsste. Dies wäre die logische Konsequenz aus der Formulierung, dass dieser Sound „nicht aus dem *realen* Raum der Szene komm[t]“ (ebd., ÜB u. Hv. M), was einer der Hauptkriterien für *non-diegetischen* Sound darstellt. Wenn man diesen Sound jedoch nun weiterhin als *diegetisch* einstuft

¹⁹ « Cette triple distinction va nous permettre d'établir une première typologie, grossière et provisoire, de ces rapports de temps et de lieu. Il ne faudra pas oublier plus tard de donner une place au monologue intérieur d'un personnage (voix interne), à la voix de l'acteur lorsqu'il est visible dans le champ mais nous tourne le dos, ou aux voix de fantômes invisibles [...] Quid même du bruit du vent, quand rien ne bouge qui viendrait le visualiser dans l'image mais qu'on l'entend ? » (Chion 2003, 224).

(wie Bordwell und Thompson dies durch die Begriffsbildung *internal diegetic sound* tun), so wird damit eine zweite Wirklichkeitsebene postuliert, die einen Spezialfall der Diegese darstellt: die einer *figureninternen* Wahrnehmungsrealität.

Akzeptiert man Sounds, deren Quelle weder sichtbar noch im raumzeitlichen Gefüge der Figurenhandlung verortet ist, die also in der bis hierhin etablierten Terminologie als *offscreen* und *non-diegetisch* angesehen werden müssten, als *diegetisch*, dann geraten die (durch Fuxjäger und Flückiger aus ihren Verankerungen gelösten) klassifikatorischen Kategorien *onscreen* vs. *offscreen* und *diegetic* vs. *nondiegetic* weiter ins Schwanken: dann wären viele andere bisher als *non-diegetisch* eingestufte Sounds plötzlich auch als *diegetisch* einzuschätzen, nämlich dann, wenn sie eine mehr oder weniger starke Ähnlichkeitsbeziehung zu Vorgängen der inneren Wahrnehmungsrealität der Figuren aufweisen – und die Konzeption letzterer hängt, da es sich um das fortlaufend erneuerte Resultat von Diegetisierungsprozessen handelt, sowohl vom audiovisuellen Input als auch von den Zuschauer/innen ab. Wendet man diese Überlegungen auf konkrete Beispiele an, z.B. auf die von Chion erwähnte Sequenz, in der ein einsamer Cowboy durch die Wüste reitet und ein im non-diegetischen Raum verortetes Orchester seine emotionale Situation in der Musik zu spiegeln versucht (vgl. Chion 2003, 233), dann müsste diese Musik als *non diegetic offscreen* klassifiziert werden, wenn ein/e Zuschauer/in auf die raumzeitliche Situierung des Orchesters als wesentliches Definitionskriterium zurückgreift. Ein anderer Zuschauer oder eine andere Zuschauerin müsste dieselbe Musik jedoch als *internal diegetic* auffassen, wenn, nach subjektivem Maßstab geurteilt, eine hohe Ähnlichkeitsbeziehung zwischen den Eigenschaften der Musik und denjenigen der Innenwelt des Protagonisten gegeben ist (u.U. wäre für manche Zuschauer/innen auch der paradoxe Fall einer Doppelklassifikation möglich). Kurz: Das Problem des Bordwell-Thompsonschen Definitionskriteriums für *internal diegetic sounds* besteht darin, dass, sobald eine Zuordnung eines Sounds zur Innenwelt einer Figur erfolgt, die Lokalisierung der Soundquelle im raumzeitlichen Universum der Außenwelt, die bisher das Kriterium für die Unterscheidung *diegetisch* vs. *non-diegetisch* darstellte, nicht mehr relevant ist.

Im Folgenden seien die konkreten Auswirkungen dieser allgemeinen Behauptung anhand von drei kurzen Filmbeispielen erläutert. Das erste ist die Eingangsszene aus *LA BATTIGLIA DI ALGERI* (SCHLACHT UM ALGIER, Algerien / Italien 1966, Gillo Pontecorvo, 117 min, Musik: Ennio Morricone / Gillo Pontecorvo). Am Anfang des Filmes, noch vor dem Titelvorspann (00:00:35-00:02:33), sieht man eine Gruppe französischer Soldaten, die einen Algerier durch Folter zur Kooperation gezwungen haben. Der Algerier zittert am ganzen Körper und kann sich kaum selbst auf den Beinen halten. Dann betritt der Kommandant der Soldaten den Raum: Colonel Mathieu. Die Figur des Colonel ist in Anlehnung an die historisch reale Person des General Jacques Massu konzipiert. Massu war Chef der 10. Fallschirmjägerdivision, die während des Algerienkrieges 1957 den Auftrag hatte, mit brutalsten ‚Anti-Terror-Methoden‘ einen Aufstand in der Altstadt von Algier niederzuschlagen. Der Algerier, der in der genannten Eingangsszene gezeigt wird, soll den Soldaten den Aufenthaltsort von ‚Ali la Pointe‘ (alias Ali Ammar) verraten, eines historisch verbürgten

Untergrundkämpfers der algerischen Widerstandsbewegung FLN. Als der Kommandant befiehlt, dem Gefolterten eine französische Uniform anzuziehen, damit er auf dem Weg zu dem Versteck nicht identifiziert werden könne, rinnen dem Algerier Tränen über die Wangen. Man sieht sein Gesicht in Großaufnahme, er dreht es langsam dem Zuschauer zu. Mit der Bewegung des Kopfes beginnt die Filmmusik, zunächst sehr leise, dann mit zunehmender Intensität: man hört den Anfang der *Matthäuspasion* von Johann Sebastian Bach (BWV 244). Das Gesicht des Algeriers verharrt einen Moment, seine Augen starren ins Leere.



Abb. 1: LA BATTIGLIA DI ALGERI (SCHLACHT UM ALGIER, 00:02:06)



Abb. 2: LA BATTIGLIA DI ALGERI (SCHLACHT UM ALGIER, 00:02:13)

Der Kommandant gibt den Befehl zum Aufbruch und macht sich auf den Weg zur Tür. Der Algerier wirft ihm einen kurzen Blick hinterher, dann dann schüttelt er den Kopf und stürzt sich mit einem lauten Schrei („Non!!!!“), aus dem Bildrahmen, Schnitt in die Nah, er reißt das Fenster auf, ein Soldat hält ihn zurück, ohrfeigt ihn, droht ihm an, wieder mit der Folter zu beginnen, führt ihn ab. Schnitt: Titel des Films unter dem Beginn der nächsten Szene: Soldaten springen von einem LKW in der Altstadt von Alger. Abrupter Wechsel der Musik: Trommelwirbel.



Abb. 3: LA BATTIGLIA DI ALGERI (SCHLACHT UM ALGIER, 00:02:21)

Die Musik, die mit der langsamen Kopfbewegung des Algeriers einsetzt, das erste Stück Musik in dem Film, ist sie *diegetisch* oder *non-diegetisch*, *offscreen* oder *onscreen*? Für einen distanzierten Beobachter, der zudem über das Wissen verfügt, dass es sich um die *Matthäuspassion* handelt, stellt die Klassifizierung kein Problem dar: *offscreen*, die Quelle der Musik, sprich: das Orchester, das die *Matthäuspassion* spielt, ist nicht zu sehen; *non-diegetisch*, die Quelle der Musik befindet sich auch nicht im gleichen Raum-Zeit-Gefüge wie die handelnden Figuren, letztere können die Musik nicht hören.

Ist dies die einzige Deutungsmöglichkeit? Wer sagt uns, dass der Algerier die Musik nicht hört, die *Matthäuspassion*, die Musik des gefolterten Christus? Weil er kein Christ ist? Weil er Bach nicht kennt? Warum verwenden Pontecorvo und Morricone an dieser Stelle die *Matthäuspassion*? Sicher nicht deshalb, um eine unüberbrückbare Differenz zwischen dem Kulturraum, dem der Algerier angehört, und demjenigen, aus dem die französischen Soldaten stammen, deutlich zu machen. Im Gegenteil: Es geht um das gemeinsame Dritte, um das, was beide Kulturräume miteinander verbinden *müsste*: das Mitleid mit dem Gefolterten, mit den Gefolterten beider Seiten. Die unversöhnliche Opposition verläuft nicht zwischen den unterschiedlichen Kulturräumen, sondern zwischen Tätern (im Plural) und einem Opfer (im Singular), zwischen einer Gruppe von Folterern und einem Gefolterten, der, zumindest dem äußeren Anschein nach, mit sich allein ist. Die Musik lässt sich in dieser Interpretation als ein materialisiertes Analogon lesen, als eine *innere Stimme*, die die Gefolterten, die ihrer Sprache beraubt worden sind, sprechen lässt – egal, ob es sich um den gekreuzigten Christus handelt oder um einen namenlosen Algerier.

Wie Bordwell und Thompson richtig anmerkten (vgl. 2008 [1997], 279; hier zit. auf S. 10), handelt es sich bei der Entscheidung, ob ein Sound *diegetisch* ist oder nicht, um einen Zuordnungsprozess, der durch die Zuschauer/innen im Einzelfall vorgenommen wird. Die im *außerfilmischen* Sinn reale, in den Worten Souriaus: *profilmische* (d.h. im *afilmischen* Raum-Zeit-Gefüge zur Produktion eines Films bereit gestellte) Quelle der Musik, im vorliegenden Fall die des Orchesters, das die *Matthäuspassion* einspielt, ist irrelevant. Es geht um die Frage, ob zur genannten Musik eine Quelle *vorstellbar* ist, die sich im Handlungsraum der Figuren verorten lässt. Die Antwort lautet dann, wenn man in der diskutierten Szene nur die *äußere*

Sinnesrealität der Figuren im Auge hat: Das Orchester ist nicht anwesend, weder im sichtbaren noch im nicht-sichtbaren Teil des Raum-Zeit-Gefüges, in dem die Figuren agieren. Wenn man im Rahmen der Bordwell-Thompsonschen Konzeption von *internal diegetic sounds* eine innere Wahrnehmungsrealität zulässt, welche anderen als der im jeweils gegebenen Fall relevanten Figur nicht zugänglich sein muss, dann handelt es sich in der Eingangsszene von *LA BATTIGLIA DI ALGERI* um eine Musik, die in hohem Maße *diegetisch* ist: Sie verleiht der emotionalen und kognitiven Entwicklung, die sich in der inneren Realität des Gefolterten abspielt, eine Stimme. Diese Stimme ist für den Algerier selbst, aber auch für diejenigen Zuschauer/innen hörbar, die über einen *empathisierenden Prozess* mit dem Algerier verbunden sind. Sie ist hörbar, bevor sie für die französischen Soldaten wahrnehmbar ist. Letzteres ändert sich erst, als die innere Bewegung mit der äußeren zusammenfällt und sich schließlich in dem verzweifelten Schrei artikuliert, der dann auf einer – sowohl im Hinblick auf die französischen Soldaten, als auch auf die Zuschauer/innen – intersubjektiven Ebene zugänglich ist. Es würde die moralische, und davon abhängig, die politische Provokation dieser Szene um einen bedeutenden Anteil schmälern, wenn man den Versuch der Soldaten, von außen, d.h. über das Zufügen von physischen Schmerzen, Kontrolle über die innere Realität des Gefolterten gewinnen zu wollen, *nicht* im Einsatz der Musik mit künstlerischen Mitteln gespiegelt sähe. Der Musik wird hier die Funktion *zugeordnet*, die Innenwelt des Gefolterten *hörbar zu machen*; sie erzählt das langsame Wachsen seines Entschlusses, schließlich doch noch die Kooperation mit den Folterern zu verweigern und sich dem Befehl des Kommandanten, der den Verrat öffentlich machen würde, durch Selbstmord zu entziehen (der Selbstmord scheitert, und der Verrat wird am Ende des Films für die Bewohner/innen der Kasbah sichtbar, trotz der Kostümierung durch die französische Uniform).

Die Filmmusik verfügt in dieser Deutungsvariante als Teil eines audiovisuellen Gesamtphänomens über eine gleichermaßen darstellende und mitteilende Funktion: Sie macht das Reifen eines Entschlusses einer Figur auf akustischer Ebene greifbar. Dies ist unabhängig davon erkennbar, ob Rezipient/innen wissen, dass Pontecorvo und Morricone an dieser Stelle die *Matthäuspassion* einsetzen. Die Musik übernimmt hier, ähnlich wie im Fall eines inneren Monologes, die Aufgabe, den Zuschauer/innen einen Vorgang zu vermitteln, der in der inneren Realität einer Figur stattfindet und insofern subjektiv ist, dass er von anderen Figuren nicht oder in reduzierter Intensität wahrgenommen werden kann, was jedoch nicht bedeutet, dass er weniger *diegetische* Relevanz besitzt. Die Rolle der Musik erschöpft sich nicht in ihrer deiktischen Funktion, sondern sie schwimmt, in Abhängigkeit von den einzelnen Rezipient/innen, mit dem zu repräsentierenden Vorgang selbst. Nimmt man darüber hinaus die außenweltlich sichtbaren Auswirkungen dieses Vorgangs ernst – die Langsamkeit der Bewegung, die Tränen, den starren Blick – so könnte man davon reden, dass sich *onscreen* Indikatoren auf die Quelle dieser Musik finden: Sie befindet sich im Bewusstsein der dargestellten Figur, unabhängig von der erkenntnistheoretischen Frage, wo und wie sich dieses, in raumzeitlichen Kategorien der Außenwelt gesprochen, im Einzelnen manifestiert.

Das Beispiel macht die Perspektivenabhängigkeit deutlich, mit der die Klassifizierung eines Elementes des audiovisuellen Diskurses als *diegetisch* oder *non-diegetisch* vorgenommen wird. Für beide Varianten lassen sich Argumente finden, je nachdem, ob und wie stark die Quelle der Musik in die innere Realität einer Figur verlagert wird.

Ein ähnliches Problem taucht dann auf, wenn es um Träume von Figuren geht. Für die Rolle des Sounds am Übergang vom Wachzustand zum Traum und auch für den umgekehrten Fall lassen sich eine Menge Beispiele finden, mit deren Hilfe es sich lohnen würde, eine Systematik der Verwendung von Sound an diesen Stellen auszuarbeiten. Gleiches gilt für den Sound innerhalb des Traums. Hier soll nur ein Beispiel genauer diskutiert werden, mit dem Ziel, bereits einen Teil der Fragen aufzuwerfen, denen in umfassenderen Studien mit größerer Akribie nachgegangen werden müsste. Eine geeignete Traumsequenz findet sich in L'OURS (DER BÄR, FR / USA 2006, Jean-Jacques Annaud, 94 min, Musik: Philippe Sarde / P.I. Tschaikowski). Die Hauptfigur ist ein kleiner Bär, dessen Mutter gleich zu Beginn des Films durch einen herunterfallenden Felsbrocken getötet wird. Der Bär ist gezwungen, sich allein in den Bergen von British Columbia durchzuschlagen. In einer Nacht schläft er auf einem Felsvorsprung (00:10:53-00:12:47). Er träumt von den Fröschen, die er tagsüber gejagt hat. Die Szene beginnt mit einer Totale auf eine Bergschlucht und einem langsamen Aufschwenk, der der Silhouette des rechten Bergabhangs folgt. Der Schwenk wird von einem ebenfalls relativ langsamen Zoom auf den Felsvorsprung abgelöst, auf dem der kleine Bär schläft. Nach einer Überblendung sieht man den Bären in der Nahaufnahme.

Die visuelle Verdichtung wird begleitet durch Streicherklänge von Piotr Illjitsch Tschaikowski, eingespielt vom *London Symphonic Orchestra* unter der Leitung von Carlo Savina. Darüber hinaus ist vereinzelt ein Käuzchen zu hören, das die Atmosphäre der nächtlichen Berglandschaft evoziert.²⁰ Dann findet eine Überblendung auf den Körper des Bären statt, der nun den ganzen Bildraum einnimmt, die langsame Zufahrt wird fortgesetzt, die Streichermusik ebenfalls. Dazu kommen nun die regelmäßigen Atemgeräusche des Bären. Danach folgt eine Abblende ins Schwarz, begleitet von einem Fade Out der Musik und einem Fade In vom Quaken der Frösche, die schließlich im Traum zu sehen sind. In der ganzen Traumsequenz findet sich keine Musik. Diese setzt erst am Ende der Sequenz wieder ein, beginnend mit zwei lang anhaltenden Tönen in zwei unterschiedlichen Frequenzen, die dann in die Streichermusik übergehen. Auf visueller Ebene ist das Ende des Traums durch eine erneute Schwarzblende markiert, danach sieht man den Bären in einer statischen Kameraeinstellung, im Ton ein kurzes, erleichtertes Seufzen, der Traum ist zu Ende, Schnitt, Abschwenk den Berghang hinunter, im Tal ein Lichtpunkt und eine Rauchwolke, ein Lagerfeuer; dort lauert, ohne dass der Bär dies zum gegenwärtigen Zeitpunkt ahnt, die im *figurenextern-diegetischen* Sinn *reale* Gefahr: zwei Trapper, die auf Bärenjagd sind.

²⁰ Das Käuzchengeräusch wäre in der Terminologie Flückigers als *Orientierungslaut* einzustufen, d.h. es wird in der Regel zwar dem diegetischen Raum zugeordnet, eine exakte Positionsbestimmung der Quelle ist jedoch nicht möglich, vgl. Flückiger (2005, 144f.; hier zit. auf S. 16.). Bordwell und Thompson müssten diesen Sound als *external diegetic* klassifizieren, ohne die Frage beantworten zu können, ob er *onscreen* oder *offscreen* ist.

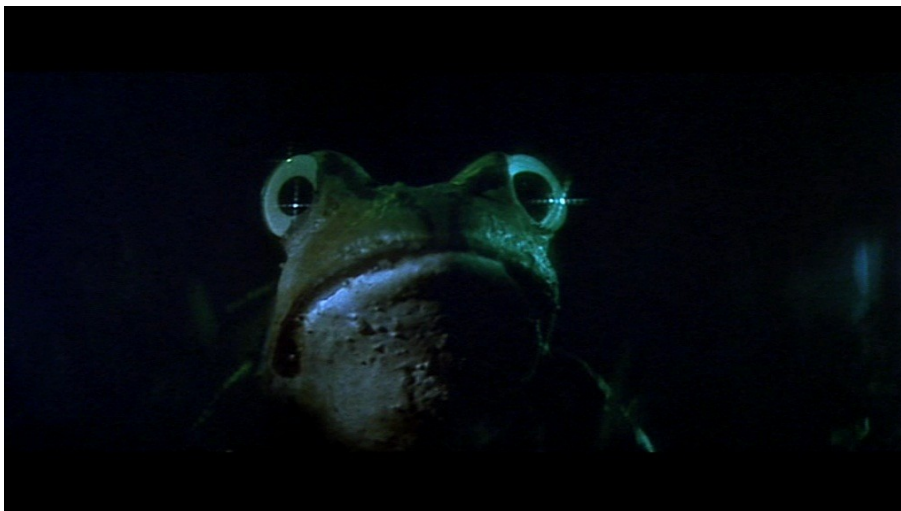


Abb. 4: L'OURS (DER BÄR, 00:11:37)

Die zwischen den beiden Schwarzblenden eingefügten Traumanimationen, die von Bretislav Pojar unter Beteiligung des Prager *Studio Kratky Film* und des *Filmexports Tschechoslowakei* konzipiert wurden, zeigen eine zunehmende Verfremdung von Fröschen. Die Frösche weisen zu Beginn der Sequenz sowohl in visueller, als auch akustischer Hinsicht eine starke Ähnlichkeit mit denjenigen Vertretern ihrer Art auf, die man auf *außerfilmisch-externer* (vertikal analoger) sowie auf *figurenextern-diegetischer* (horizontal analoger) Ebene als real einstufen müsste, doch die Ähnlichkeit nimmt im Verlauf der Sequenz ab.



Abb. 5: L'OURS (DER BÄR, 00:11:53)

Der Traum entwickelt sich zum Albtraum, was auch auf der Tonebene gespiegelt wird: Sound-Ingenieur Laurent Quadglio und Filmmusik-Komponist Philippe Sarde lassen das anfangs noch relativ *naturalistisch* anmutende Quaken der Frösche zunehmend dumpfer und bedrohlicher klingen, und sie spielen nach und nach ein windähnliches Pfeifen dazu, das am Ende, als einer der Frösche in einem bereits durch eine irreal

Farbgebung entstellten Tümpel landet, in ein undefinierbares Geräusch mündet, das sich sowohl durch seine Ähnlichkeit mit einem realen Aufprall auf der Wasseroberfläche auszeichnet, als auch durch einen stark artifiziellen Charakter, den man bereits als *prä-musikalisch* einschätzen könnte.

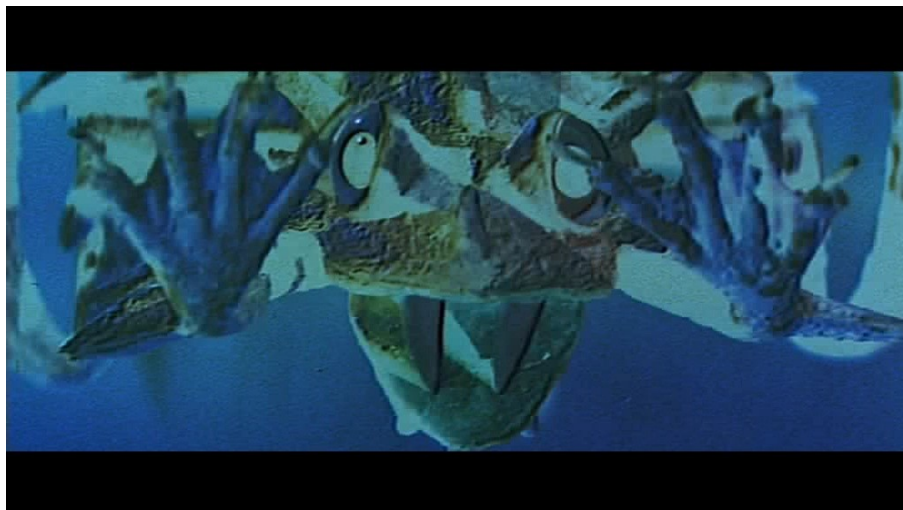


Abb. 6: L'OURS (DER BÄR, 00:12:01)

Lässt sich dieser zunehmende Verfremdungsprozess, der sowohl in visueller als auch akustischer Hinsicht innerhalb der Traumsequenz zu beobachten ist, in das terminologische Raster der oben vorgestellten etablierten Klassifikationskriterien einordnen? Ist die Quelle des im Traum des Bären zu hörenden Sounds *onscreen* oder *offscreen*? Zu welcher Wahrnehmungswelt gehört dieser Sound? Zu derjenigen der Frösche, derjenigen des Bären oder ausschließlich derjenigen der Zuschauer/innen? Ist dieser Sound *diegetisch* oder *non-diegetisch*, ist er, falls er diegetisch sein sollte, im *figureninternen* oder *figurenexternen* Sinn diegetisch?

Im Fall der Überleitungen zum Traum ist der Fall eindeutig: die Tschaikowski-Musik ist *offscreen* und *non-diegetisch*. Zu dem Zeitpunkt, zu dem die Musik einsetzt, sehen wir eine Totale auf die Berge von British Columbia, wir wissen noch nicht, dass es um den schlafenden Bären geht. Weder ist die Quelle der Musik im visualisierten Handlungsraum zu erkennen, noch findet sich eine Figur, deren innere Entwicklung durch die Musik gespiegelt sein könnte, sprich: In der Berglandschaft ist kein Orchester zu sehen, und die Berge werden normalerweise, d.h. unter gewöhnlichen Bedingungen und von der überwiegenden Mehrheit der Zuschauer/innen, nicht als Lebewesen eingeschätzt – was nicht heißen soll, dass es nicht Zuschauer/innen gibt, für die letzteres eine Tatsache darstellt; dies ist dann der Fall, wenn bestimmte animistische Weltanschauungen vertreten werden bzw. wenn Situationen, in denen animistische Deutungsmuster über eine rein simulative Ebene hinaus an Relevanz gewinnen, zum persönlichen Erfahrungshorizont von Zuschauer/innen gehören. Es gibt darüber hinaus eine Vielzahl von Filmen, in denen das Motiv der Belebtheit des eigentlich Unbelebten in die dramaturgischen und ästhetisch-formalen Strukturen eingearbeitet ist, u.a. dann, wenn figureninterne Perspektiven dargestellt werden, die figurenextern als ein

Resultat psychischer Krankheit gewertet werden können.²¹ Der letztgenannte Fall soll in der Analyse des dritten Filmbeispiels, in dem es um die Wahrnehmungsrealität einer schizophrenen Hauptfigur geht, genauer untersucht werden.

Zurück zum Traum des kleinen Bären: Der Einsatz der beiden Schwarzblenden und der Tschaikowski-Musik vor dem Beginn und nach dem Ende der Darstellung des Traums lassen sich vergleichsweise eindeutig dem narrativen Diskurs, nicht jedoch den figurenintern oder figurenextern verorteten diegetischen Realitäten zuordnen. Wie steht es nun um denjenigen Teil der Szene, der sich innerhalb dieser explizit gekennzeichneten visuellen und akustischen Markierungen befindet?

Würde man den während der Darstellung des Traums wahrnehmbaren Sound im Rekurs auf die (an Gérard Genettes literaturtheoretischen Überlegungen orientierte) Terminologie Claudia Gorbmans als *metadiegetisch* bezeichnen (vgl. Gorbman 1976, 450), würde man sich der Frage, welche dramaturgische Bedeutung den innerhalb des Traums auftretenden audiovisuellen Phänomenen im Einzelnen zukommt, durch ein stark formalisierendes Vorgehen entziehen. *Metadiegetisch* hieße im gegebenen Kontext nur, dass die visuellen und akustischen Inhalte des Traums einen anderen Realitätsstatus aufweisen als die Elemente der als *figurenextern* konstituierten diegetischen Wirklichkeit, konkret: Die im Traum hör- und sichtbaren Frösche sind, ontologisch betrachtet, nicht dieselben Frösche, denen der kleine Bär in der der Nachtszene direkt vorausgehenden Szene in seiner Außenwelt begegnet. Hier werden sehr schnell die Grenzen des Gorbmanschen bzw. Genetteschen Modells deutlich: Genette bezeichnet eine Geschichte als *metadiegetisch*, wenn sie *innerhalb* einer Rahmenhandlung erzählt wird. Die für das Verständnis der psychischen und sozialen Vorgänge, die das Verhalten der erzählten Figuren prägen, aufschlussreiche Frage, inwiefern die diegetische und die metadiegetische Ebene miteinander zusammenhängen und inwiefern sich beide Ebenen voneinander unterscheiden, wird ausgeblendet. Inwiefern stellt die metadiegetische Ebene ein Abbild, eine Erinnerung, eine Vorahnung der figurenexternen Realität dar? Inwiefern spiegelt die metadiegetische Ebene Prozesse der figureninternen Realität? Wie real im Sinne der figurenexternen Realität ist die figureninterne Realität? Wie real im Sinne der außerfilmischen Wirklichkeit der Zuschauer/innen sind beide diegetischen Realitäten, d.h. sowohl die figureninterne, als auch -externe? Inwiefern weist die metadiegetische Ebene non-diegetische Elemente auf, d.h. Elemente, die sich nicht als Teil des auf der metadiegetischen Ebene etablierten Raum-Zeit-Gefüges verstehen lassen?

Die Beantwortung dieser Fragen ist von den einzelnen Zuschauer/innen abhängig. Ähneln die Frösche im Traum des kleinen Bären denjenigen, mit denen er tagsüber gespielt hat? Welche Vorgänge in der inneren Realität des Bären lösen die Verfremdungseffekte aus, die im Verlauf der Traumsequenz auftreten? Die Angst vor den unbekanntem Lebewesen und neuen Situationen, mit denen der kleine Bär täglich konfrontiert wird?

²¹ Vgl. z.B. den oben erwähnten Film SHUTTER ISLAND (USA 2010, Martin Scorsese), DAS WEISSE RAUSCHEN (2001, Hans Weingartner) oder auch LENZ (CH 2006, Thomas Imbach), aus dem die die dritte hier analysierte Beispielszene entnommen ist.

Die Angst, ohne seine Mutter nicht überleben zu können? Oder lassen sich die Verzerrungen der figurenexternen Realität innerhalb der Traumsequenz einfach auf die überschwängliche Phantasie des Bären zurückführen, auf seine Experimentierlust, seinen Spieltrieb? Inwieweit handelt es sich bei den Traumhalten tatsächlich um *Inhalte*, d.h. um das, was sich der Bär im Traum vorstellt, und inwiefern ging es den an der Gestaltung des Films Beteiligten einfach nur darum, in der audiovisuellen Darstellung eine formale Differenz zwischen figureninterner und -externer Realität deutlich zu machen, eine Differenz, die in ihrer faktisch vorliegenden Form primär der Kreativität derjenigen geschuldet ist, die die Animationssequenzen konzipiert haben? Um die Fragen zuzuspitzen: Inwiefern sind die in der Traumsequenz hörbaren Sounds als *internal diegetic* zu klassifizieren (d.h. inwiefern entspringen sie der inneren Realität des Bären), inwiefern sind sie als *external diegetic* einzustufen (d.h. wie stark sind die Analogien zur diegetischen Außenwelt) und inwiefern müsste man sie als *non diegetic* betrachten (d.h. inwiefern verweisen sie als Teil des narrativen Diskurses auf die außerfilmischen Urheber dieser Diskurses)? Die Zuschauerabhängigkeit, die hier ins Spiel kommt, besteht nicht nur im Hinblick auf die Einschätzung der einzelnen Soundphänomene innerhalb der Traumsequenz. Sie ist auch dann gegeben, wenn man mit Hilfe der Gorbmanschen Kategorien *metadiegetic*²² und *extradiegetic*²³ versucht, eine formale Klassifizierung der gesamten Sequenz vorzunehmen, da nicht intersubjektiv entscheidbar ist, ob sich die Geräusche während des Traums *tatsächlich* in der Vorstellungswelt des Bären befinden (vgl. Anm. 20: „sound [...] imagined by a character“) oder ob sie auf außerhalb der Diegese liegende Quellen verweisen und daher durch den Bären nicht wahrnehmbar sind. Auf dieses Problem macht auch Gorbman selbst aufmerksam. Da sie jedoch der Deskription des *filmischen Diskurses* (nach Souriau: der *filmophanischen* Ebene) ein Primat gegenüber einer phänomenologischen Analyse der (sich in Interaktion zwischen Diskurs und Zuschauer/innen fortwährend neu konstituierenden) *diegetischen Realitäten* einräumt, plädiert sie dafür, die genannten Kategorien beizubehalten:

Ultimately, these terms are arbitrary and unsuitable for careful film analysis, since what is metadiegetic for one auditor may be extradiegetic for another. But the distinctions can be very useful pedagogically for outlining the very problem of levels of narrativity in films, and they also lead to considerable enlightenment in describing cinematic styles (Gorbman 1976, 450).

Kommen wir zum dritten Beispiel. Es handelt sich um mehrere Sequenzen aus LENZ (CH 2006, Thomas Imbach, 95 min, Musik: Peter Bräker / Balz Bachmann), einer sehr freien Adaption von Georg Büchners Romanfragment *Lenz* (das 1839, zwei Jahre nach dem Tod des Autors, zum ersten Mal publiziert wurde). In Imbachs Film spielt Milan Peschel einen Berliner Filmregisseur, der in den Schweizer Bergen von

²² „Metadiegetic: sound apparently ‚narrated‘ or imagined by a character as secondary narrator, e.g., in *SCARLETT STREET*, the murdered wife’s voice that rings in Edward G. Robinson’s ears; likewise, in *BLACKMAIL*, the famous ‚knife‘ auditory hallucination; in *BLOW-UP*, the sounds of tennis as Hemmings walks away from mimed tennis match“ (Gorbman 1976, 450). Der von Bordwell und Thompson verwendete Begriff *internal diegetic* (vgl. 2008 [1997], 284) bezeichnet also einen Spezialfall der Phänomene, die Gorbman als *metadiegetic* klassifiziert.

²³ „Extradiegetic: ‚soundtrack‘ sound, not perceivable by diegetic characters.“ (Gorbman 1976, 450). Diese Definition entspricht in der logischen Konsequenz derjenigen, die Bordwell und Thompson für den Begriff *nondiegetic* angeben: „[Nondiegetic sound] is represented as coming from a source outside the story world.“ (2008 [1997], 279).

Schizophrenie-Anfällen geplagt wird. Der Protagonist hatte ursprünglich geplant, Büchners *Lenz* zu verfilmen, entscheidet sich dann jedoch, ein Selbstporträt zu drehen. Er besucht gegen den Willen seiner Ex-Frau Natalie seinen neunjährigen Sohn Noah, der sich in den Winterferien mit einem Kindermädchen in einer Hütte in den Bergen aufhält.

Bereits in den Einstellungen rund um die Titelsequenz des Films (00:06:20-00:07:21) wird deutlich, dass sich die innere und äußere Realität der Hauptfigur nur schwer voneinander abgrenzen lassen. Bevor der Vorspann eingeblendet wird, sind bereits 6 Minuten und 20 Sekunden vergangen, wir haben die Hauptfigur bereits eine Weile beobachtet: u.a. bei einem nächtlichen Streifzug durch eine Hügelandschaft, beim Ablecken von Tautropfen von Tannenzweigen, beim Träumen auf einem mit Laub bedeckten Waldboden (Rückblende in eine Urlaubserinnerung: sein Sohn Noah am Strand), beim Telefonieren mit seiner Ex-Frau, beim Einkaufen, bei der Besichtigung eines verlassenen Hauses. In den O-Ton der Gegenstände, die den Protagonisten umgeben, mischen sich hin und wieder einzelne Anschläge einer E-Gitarre, die sich an einer Stelle, als sich die Figur, mit einer blonden Lockenperücke bekleidet und die Gesichtsmuskulatur zu wilden Zuckungen verzerrt, in den Ästen einer Tanne verfangt, zu einer kurzen Improvisation verdichten. Die letzte Einstellung vor der Titelsequenz ist eine Großaufnahme des Gesichts der Figur, die in dem verlassenem Haus aufmerksam auf die Geräusche lauscht, die zu hören sind, darunter liegt ein gedämpftes elektronisches Fiepen.



Abb. 7: LENZ (00:06:14)

Dann sieht man eine Winterlandschaft, einen Feldweg, der von gefrorenen Gräsern überwuchert ist, zu beiden Seiten kleine Tannen. Der Himmel ist bedeckt, Wind- und Tiergeräusche. Schnitt: ein Betonblock, der über dem Boden zu schweben scheint, pendelt langsam hin und her. Die Stille wird unterbrochen durch das Ächzen seiner metallischen Verankerung, es setzt immer dann ein, wenn der Block zu schwingen beginnt. Simultan dazu ein gedämpfter elektronischer Ton.



Abb. 8: LENZ (00:06:37)

Direkt im Anschluss ein Sprung in die Subjektive: Handkamera, Bewegung durch das mit Raureif überzogene Feld, die Perspektive auf die Gräser gerichtet. Der Horizont liegt meist außerhalb der Kadrierung. Im Ton schwere Atemgeräusche, die Schritte auf dem gefrorenen Gras, gedämpftes Stimmengewirr, wie aus einem nicht richtig eingestellten Radiosender. Und erneut das elektronische Fiepen aus dem verlassenen Haus, mit sich steigernder Intensität. Über diese Einstellung läuft der Titelvorspann. Kurze Zeit später, nachdem die Hauptfigur mit einer Seilbahn einen Berg hochgefahren und auf seinem Koffer mit irrsinniger Geschwindigkeit einen Abhang hinunter gerutscht ist, sieht man, wie der Kondensstreifen eines Flugzeugs das Matterhorn zu ‚durchbohren‘ scheint (00:08:25-00:08:32), darunter liegt ein weiteres Mal ein anhaltender elektronischer Ton, dessen Frequenzspektrum sich leicht verändert hat, er klingt stimmhafter als in den vorausgehenden Einstellungen.



Abb. 9: LENZ (00:06:31)

Wie sind die Sounds, die in den genannten Sequenzen aus LENZ zu hören sind, im Hinblick auf ihre Verankerung im raumzeitlichen Gefüge der erzählten Geschichte zu bewerten? Bei manchen Geräuschen, wie dem Ächzen des schwingenden Betonblocks, dem lauten Atmen oder den Schritten über dem gefrorenen Gras, scheint auf den ersten Blick klar zu sein, dass die Soundquelle im Handlungsraum zu finden ist und dass die Geräusche potentiell auch von anderen Figuren als der Hauptfigur wahrnehmbar wären. Sicher ist dies jedoch nicht. All diese scheinbar als *external diegetic sounds* klassifizierbaren Geräusche klingen leicht verfremdet, sei es durch ihre überdimensioniert hohe Lautstärke, sei es durch die Überlagerungen mit Tönen, die sich eindeutig als elektronisch produziert identifizieren lassen. Man könnte diese Verfremdungen als Resultat einer subjektiven Wahrnehmung durch die Hauptfigur deuten. Sind die genannten Geräusche dann also *internal diegetic*? Ein ähnliches Problem stellt sich im Hinblick auf den Gitarrensound, der in den Einstellungen vor der Titelsequenz zu hören ist. Im Bild ist keine Gitarre zu sehen, der Sound könnte also *non-diegetic* sein, wenn man nicht die Tatsache mit in Betracht ziehen müsste, dass die Hauptfigur im weiteren Verlauf des Films häufiger beim Gitarrenspiel zu sehen und zu hören ist, allerdings handelt es sich ausnahmslos um eine akustische Gitarre. Lässt sich der elektronisch verstärkte Gitarrensound am Anfang des Films dann als *external diegetic offscreen* klassifizieren, oder ist er doch, wie zunächst vermutet, *non-diegetic*?

Wie steht es mit den eindeutig elektronisch produzierten Sounds? Dem leisen Fiepen im verlassenen Haus, der elektronischen Verstärkung der Pendelbewegung des Betonblocks oder dem sonoren Brummen unter der Matterhorn-Einstellung? Die Einstellung im verlassenen Haus zeigt die Hauptfigur in Großaufnahme, bei einem intensiven Hörvorgang (vgl. Abb. 7). Was hört die Figur? Nur die Stille im Haus und die Vogelgeräusche aus dem Garten, oder auch das elektronische Fiepen? Mit hoher Wahrscheinlichkeit handelt es sich bei dem elektronisch erzeugten Ton nicht um einen Sound, der in der äußeren Wahrnehmungsrealität der Figur existiert, d.h. er ist nicht von anderen Figuren hörbar und daher nicht *external diegetic*. Ist der genannte Sound nun der inneren Realität der Figur zuzuordnen (dann wäre er *internal diegetic*) oder wird er als durch die betreffende Figur nicht wahrnehmbar eingeschätzt (in diesem Fall wäre er *non-diegetic*)?

Das Hauptproblem bei *internal diegetic sounds* besteht darin, dass es auf visueller Ebene keine Garantie dafür gibt, dass eine Figur einen bestimmten Sound auch tatsächlich wahrnimmt; selbst dann, wenn sich dieser Sound auf visueller Ebene in Form eines bestimmten mimischen oder gestischen Verhaltens ausdrücken sollte, bleibt die Verkopplung des jeweils relevanten akustischen mit dem jeweiligen visuellen Phänomen eine Frage des individuellen Zuordnungsprozesses, der durch den einzelnen Zuschauer oder die einzelne Zuschauerin vorgenommen wird.

Im vorliegenden Fall werden vermutlich diejenigen Zuschauer/innen, die nach der Besichtigung der ersten Szenen des Films bereits zu der Einschätzung gelangt sind, dass das Verhalten der Hauptfigur Schizophrenie-ähnliche Symptome aufweist (und darüber hinaus auch diejenigen, denen Büchners *Lenz* oder auch die

Handlung des Films schon vorher bekannt war), eher dazu bereit sein, das elektronische Fiepen unter der letzten Einstellung im verlassenen Haus als *internal diegetic* zu interpretieren, als eine Art ‚inneres Pfeifen‘, eine akustische Halluzination, die mit der Erkrankung der Hauptfigur in Zusammenhang steht. Andere Zuschauer/innen werden dieser Deutung aus nicht minder nachvollziehbarem Grund widersprechen: Für sie klingt der betreffende Sound zu technisch-artifizell, als dass er als Bestandteil der inneren Realität einer Figur akzeptiert werden könnte; für sie drängt sich die außerfilmische Soundquelle in den Vordergrund, sie haben beim Hören des elektronischen Fiepens sofort den Sound-Ingenieur vor Augen, der die Tonspur bearbeitet. Und es gibt sicher auch Zuschauer/innen, von denen der genannte Sound mit beiden Realitätsebenen verknüpft wird, sowohl mit der *diegetischen*, als auch mit der *non-diegetischen*.

Für die elektronisch erzeugten Sounds in den anderen Einstellungen gilt Ähnliches. Bereits auf visueller Ebene stellt sich die Frage, ob der über dem gefrorenen Boden schwebende Betonblock (vgl. Abb. 8) oder das scheinbar von einem Flugzeug ‚durchbohrte‘ Matterhorn (vgl. Abb. 9) Teil des Wahrnehmungsprozesses der Hauptfigur sind oder nicht. Da die entsprechenden Einstellungen in der Montage nicht mit einem ihnen korrespondierenden Blick der Figur verbunden werden, ist die Antwort auf die Frage, ob es sich bei diesen Einstellungen um Subjektiven oder Objektiven handelt, Ermessenssache. In Abhängigkeit davon kann nicht eindeutig entschieden werden, ob der in diesen Einstellungen zu hörende Sound *internal diegetic* oder *non-diegetic* ist.

Die geschilderte Zuordnungsproblematik im Hinblick auf die audiovisuelle Gestaltung findet sich auch in den Details der dramaturgischen Konzeption. Die Hauptfigur des Imbach-Films heißt Lenz, wie der Protagonist des Büchner-Fragments. Nachdem Lenz sich dazu entschieden hat, ein Selbstporträt zu drehen, anstatt das Fragment zu verfilmen, werden die diegetische Realität des Imbach-Films und die Diegese des von Lenz ursprünglich geplanten Films (d.h. die Diegese des Film im Films, in der Genette-Gorbmanschen-Terminologie: die Metadiegeese) ununterscheidbar. Dieser Zusammenhang wird im Dialog thematisiert, als Lenz in einem fortgeschrittenen Stadium seiner Krankheit wahllos Passant/innen anspricht, die zufällig an der Berghütte vorbeikommen. Ein Schweizer und eine Dänin nehmen seine Einladung an, mit ihm in der Hütte eine Suppe zu essen. Hier ein Auszug aus dem Gespräch, das sich bei dieser Gelegenheit entwickelt:

Schweizer: Du spielst dich selber, oder?

Lenz: Ja, ich filme mich selber, einfach. Ich hab's aufgegeben, mit Schauspielern zu arbeiten. So Leute wie ihr, ist viel besser, ist viel interessanter, ist viel echter.

Dänin: Wirklich?

Lenz: Ja, ihr versteckt euch nicht, ihr sagt immer die Wahrheit.

Schweizer: Ja, hoffen wir mal.²⁴

²⁴ LENZ (00:39:21-00:39:33).

Die Zuschauer/innen werden im Unklaren darüber gelassen, ob die beiden Passant/innen Schauspieler/innen sind oder ob es sich im außerfilmischen Sinn um *reale* Passant/innen handelt, die bei den Dreharbeiten des Imbach-Films spontan dazu aufgefordert wurden, in dem Film mitzuwirken. Die verschiedenen Realitätsebenen schieben sich ineinander: Die *außerfilmische* Realität (zu der möglicherweise die Passant/innen gehören, die dann dokumentarisch gefilmt wurden, womit die Realität der Zuschauer/innen Eingang in den Film fände), die *profilmische* Realität (Milan Peschel, der sich relativ schutzlos seiner Rolle ausliefert), die *diegetische* Realität (in der Lenz als Filmregisseur agiert), die *metadiegetischen* Realitäten (diejenige des Films im Film sowie diejenige des Büchnerschen Romanfragments, das über die Analogien der jeweils erzählten Geschichten hinaus auch als Requisite in den anderen Realitätsebenen präsent ist: als Reclamheftchen).

Aufgrund der psychischen Krankheit der Hauptfigur, welche sich sowohl in einem spezifischen Einsatz der audiovisuellen Gestaltungsmittel, als auch in der narratologischen Struktur des Films spiegelt, handelt es sich bei den oben diskutierten Einstellungen aus LENZ um Extrembeispiele. Die Klassifikations-schwierigkeiten, die bei einer Verwendung der etablierten Terminologie zu Bild-Sound-Beziehungen in diesen Fällen entstehen, sind daher das Resultat einer kritischen Zuspitzung, einer gewissen Überzeichnung, die es ermöglicht, vergleichbare Probleme, die in anderen Filmen mit geringerer Signifikanz auftreten, besser zu verstehen.

Konklusion

Als Ergebnis der hier vorliegenden Analyse lässt sich festhalten, dass die in der Filmmusikforschung häufig verwendeten Begriffspaare *onscreen* vs. *offscreen*, *diegetisch* vs. *non-diegetisch* (bzw. entsprechend: *intradiegetisch* vs. *extradiegetisch*), *internal diegetic* vs. *external diegetic* zu einer präzisen Beschreibung der temporären Resultate der interaktiven Prozesse, die auf perceptiver, kognitiver und emotionaler Ebene zwischen audiovisuellem Input und Rezipient/innen stattfinden, nicht ausreichen. Klassifizierungen von Bild-Sound-Relationen mit Hilfe der genannten Terminologie produzieren in vielen Fällen logische Widersprüche, die nur mit Hilfe einer stärkeren Berücksichtigung individueller Variablen des Rezeptionsverhaltens von Zuschauer/innen bzw. Zuhörer/innen aufgelöst werden können. Das Problem einer nicht herstellbaren Intersubjektivität der jeweils vorgenommenen Klassifizierungen stellt sich mit besonderer Schärfe dann, wenn es um die Beurteilung von Sounds geht, bei denen die Möglichkeit gegeben ist, dass sie sich *inneren Realitäten* der handelnden Figuren zuordnen lassen. Die Nutzung dieser Möglichkeit hängt von den jeweils aktivierten kulturellen Wissensbeständen sowie den situationsspezifisch dominierenden psychosozialen Dispositionen auf Seiten der Rezipient/innen ab – ohne dass der Einfluss der intersubjektiv verifizierbaren audiovisuellen Wahrnehmungsdaten auf den Verlauf von Diegetisierungsprozessen unterschätzt werden soll.

Zu klären wäre die Frage, durch welche Komponenten die genannten *figureninternen Realitäten* im Einzelnen bestimmt sind – wobei evident ist, dass diese Innenwelten nicht als abgeschlossene Systeme aufgefasst werden können, sondern durch wechselseitige Abhängigkeiten mit externen Wahrnehmungsrealitäten geprägt sind. Die Suche nach Antworten auf diese Frage kann in der Kooperation von Film- und Musikwissenschaft allein nicht geleistet werden; eine genauere Deskription der Entstehung und Wirkungsweise von Phänomenen, die von Bordwell und Thompson als *internal diegetic sounds* bezeichnet werden, verlangt nach der Öffnung des Forschungsfeldes gegenüber einer ganzen Reihe von weiteren Einzeldisziplinen.

Literatur

- Bordwell, David / Thompson, Kristin (2008 [1997]) *Film art. An Introduction*. Eighth international edition. New York: McGraw-Hill.
- Chion, Michel (2003) *Un art sonore, le cinéma. Histoire, esthétique, poétique*. Cahiers du cinéma, Paris. [it. 2007 bei Kaplan Edizioni, Turin; engl. 2009 als *Film, a Sound Art* bei Columbia University Press, New York.]
- Flückiger, Barbara (2007 [2001]) *Sound Design. Die virtuelle Klangwelt des Films*. Marburg: Schüren. 3. Auflage 2007. Erstaufgabe 2001.
- Flückiger, Barbara (2005) Narrative Funktionen des Filmsounddesigns: Orientierung, Setting, Szenographie. In: Segeberg, Harro / Schätzlein, Frank (2005) *Sound. Zur Technologie und Ästhetik des Akustischen in den Medien*. Schriftenreihe der Gesellschaft für Medienwissenschaft (GFM). Marburg: Schüren. S. 140-156.
- Fuxjäger, Anton (2007) Diegese, Diegesis, diegetisch: Versuch einer Begriffsentwerrung. In: *montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation*. 16, 2. Marburg: Schüren. S. 17-37.
- Genette, Gérard (1972) *Discours du récit*. In: *Figures III*. Paris: Editions du Seuil. S. 65-282. [dt. vgl. Genette 1998 [1994].]
- Genette, Gérard (1983) *Nouveau discours du récit*. Paris: Editions du Seuil. [dt. vgl. Genette 1998 [1994].]
- Genette, Gérard (2007 [= 1972 + 1983]) *Discours du récit*. Paris: Editions du Seuil. [frz. Neuausgabe v. Genette 1972 und 1983 in einem Buch.]
- Genette, Gérard (1998 [1994], frz. 1972 + 1983) *Die Erzählung*. 2. Auflage. 1. Auflage: 1994. München: Fink. Übers. v. Andreas Knop. [frz. Orig. vgl. Genette 1972, 1983.]
- Gorbman, Claudia (1976) Teaching the Soundtrack. In: *Quarterly Review of Film Studies*. 1, 4. S. 446-452.
- Jullier, Laurent (1995) *Les sons au cinéma et à la télévision. Précis d'analyse de la bande-son*. Paris: Armand Colin.
- Kessler, Frank (1997) Etienne Souriau und das Vokabular der filmologischen Schule. In: *montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation*. 6, 2. Marburg: Schüren. S. 132-139.
- Kessler, Frank (2007) Von der Filmologie zur Narratologie. Anmerkungen zum Begriff der Diegese. In: *montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation*. 16, 2. Marburg: Schüren. S. 10-16.
- Merlin, Dieter (2010) *Der Diegesebegriff Souriaus*. Bislang unveröffentlichtes Manuskript. Universität Poitiers.
- Metz, Christian (1968) *Essais sur la signification au cinéma*. Bd. 1. Collection d'esthétique, 3. Paris: Klincksieck. [dt. vgl. Metz 1972.]

- Metz, Christian (1971) *Langage et cinéma*. Paris: Librairie Larousse. [dt. 1973: *Sprache und Film*. Frankfurt a.M.: Athenäum.]
- Metz, Christian (1972 [frz. 1968]) *Sémiologie des Films*. München: Wilhelm Fink. Übers. v. Renate Koch. [frz. Orig. vgl. Metz 1968.]
- Murray Schafer, Raymond (1994 [1977]) *The Soundscape. Our Sonic Environment and the Tuning of the World*. Rochester VT: Destiny Books.
- Noll Brinckmann, Christine (2007) Diegetisches und nondiegetisches Licht. In: *montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation*. 16, 2. Marburg: Schüren. S. 71-91.
- Odin, Roger (1983) Pour une sémio-pragmatique du cinéma. In: *Iris*, 1, 1, 1983. S. 67-82. [Engl. 1995: For a semio-pragmatics of film. In: Buckland, Warren (Hg.) (1995) *The film spectator from sign to mind*. Amsterdam: Amsterdam University Press. S. 213-226.]
- Schaeffer, Pierre (1952) *A la recherche d'une musique concrète*. Paris: Editions du Seuil.
- Schätzlein, Frank (2005) Sound und Sounddesign in Medien und Forschung. In: Segeberg, Harro / Schätzlein, Frank (2005) *Sound. Zur Technologie und Ästhetik des Akustischen in den Medien*. Schriftenreihe der Gesellschaft für Medienwissenschaft (GFM). Marburg: Schüren. S. 24-40.
- Souriau, Etienne (1951) La structure de l'univers filmique et le vocabulaire de la filmologie. In: *Revue internationale de Filmologie*. N° 7-8, 1951. Paris: Presses Universitaires de France. S. 231-240. [dt. vgl. Souriau 1997.]
- Souriau, Etienne (Hg.) (1953 [1951]) *L'Univers filmique*. Paris: Flammarion. [Fertiggestellt am 01.03.1951 (vgl. Vorwort), publiziert erst 1953.]
- Souriau, Etienne (1997) Die Struktur des filmischen Universums und das Vokabular der Filmologie. In: *montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation*. 6, 2. Marburg: Schüren. Übers. von Frank Kessler. [frz. Orig. vgl. Souriau 1951.]
- Türschmann, Jörg (1994) *Film – Musik – Filmbeschreibung. Zur Grundlage einer Filmsemiotik in der Wahrnehmung von Geräusch und Musik*. Münster: MakS-Publikationen.
- Wulff, Hans Jürgen (2007) Schichtenbau und Prozesshaftigkeit des Diegetischen. Zwei Anmerkungen. In: *montage AV. Zeitschrift für Theorie und Geschichte audiovisueller Kommunikation*. 16, 2. Marburg: Schüren. S. 39-51.

Internetquellen

- Wienen, Andrea / Twele, Holger (2004) *Gegen die Wand. Fatih Akin. Deutschland 2004*. Filmheft der Bundeszentrale für politische Bildung, Bonn. <http://www.bpb.de/files/CDVFQZ.pdf> (Stand: 28.06.10).
- Rodrigues Singer, Philip: *The Story of Jack Foley*. Quelle nach Rodriguez Singer: Artikel ‚Disk.151 Book.70.‘ <http://www.marblehead.net/foley/index.html> (Stand: 26.06.2010).

Filme

DAS WEISSE RAUSCHEN (D 2001, Hans Weingartner, 106 min, Marek Goldowski).

GEGEN DIE WAND (D/Türkei 2004, Fatih Akin, 116 min, Musik: diverse Komponist/innen und Interpret/innen).

LA BATTIGLIA DI ALGERI (SCHLACHT UM ALGIER, Algerien/Italien 1966, Gillo Pontecorvo, 117 min, Musik: Ennio Morricone / Gillo Pontecorvo).

LEBEN IN GESCHICHTEN – WOLFGANG KOHLHAASE. Dokumentation. Regie: Lutz Pehnert, Produktion: cine plus im Auftrag des RBB und in Zusammenarbeit mit ARTE. RBB 2006, 43:20 min. Gesendet im RBB am 21.02.2010, 23h50.

LENZ (Schweiz 2006, Thomas Imbach, 95 min, Musik: Peter Bräker / Balz Bachmann).

LISBON STORY (D 1994, Wim Wenders, 100 min, Musik: Jürgen Knieper / Madreus).

L'OURS (DER BÄR, Frankreich/USA 2006, Jean-Jacques Annaud, 94 min, Musik: Philippe Sarde / Pjotr Iljitsch Tschaikowski).

SHUTTER ISLAND (USA 2010, Martin Scorsese, 138 min, Musik: Robbie Robertson).

Empfohlene Zitierweise:

Merlin, Didi: Diegetic Sound. Zur Konstituierung figureninterner und -externer Realitäten im Spielfilm.

In: *Kieler Beiträge zur Filmmusikforschung* 6, 2010.

URL: <http://www.filmmusik.uni-kiel.de/beitraege.htm>

Datum des Zugriffs: 1.11.2010.

Kieler Beiträge für Filmmusikforschung (ISSN 1866-4768)

Copyright © by Didi Merlin. All rights reserved.

Copyright © für diese Ausgabe by Kieler Gesellschaft für Filmmusikforschung. All rights reserved.

This work may be copied for non-profit educational use if proper credit is given to the author and „Kieler Beiträge für Filmmusikforschung“.