

Bild-Text-Ton-Analysen intermedial – am Beispiel von Musik(video)clips

Abstract

Der Musikclip gehört seit den 1980er Jahren zum Forschungsbereich diverser Disziplinen und gilt Vielen als intermediales Phänomen schlechthin. Als problematisch erweist sich allerdings nach wie vor, dass das klangliche Material des Clips, populäre Musik, eine Herausforderung nicht nur für die Musikwissenschaften darstellt – greifbar wird dies mit Blick auf die anhaltenden Diskussionen um einen adäquaten Begriff der populären Musik. Darüber hinaus gilt Musik allgemein als ‚Sonderfall‘ für den Bereich der Medien-, Sprach- und Kulturwissenschaften, da an ihr weder rein medienästhetische noch kommunikations- und informationstheoretische Begriffe in ausreichender Weise greifen. Die Entwicklung eines transdisziplinär nachvollziehbaren Objektverständnisses des Musikclips bleibt daher desiderabel.

Der Beitrag zum Thema „Bild-Text-Ton-Analysen“ resultiert aus einer intensivierten Begegnung von Medienwissenschaft und Musikwissenschaft. Im Artikel wird die Konstitution von Bedeutung im intermedialen Zusammenspiel von Sprache/Text, Stimme und Musik fokussiert. Dies geschieht auf Grundlage einer näheren Bestimmung der Analysekriterien, die im Hinblick auf den speziellen Fall des popmusikalischen Umgangs mit Sprache erforderlich sind. Ziel ist es, die Bedeutungssedimente von vokaler Performanz im Kontext von populärer Musik offenzulegen. Für die Betrachtung des Musikclips ist dies ein wesentlicher Zwischenschritt. Anhand der Darstellung der klanglich-materiellen Vorprägungen gilt es, die Möglichkeitsbedingungen der (nachträglichen) intermedialen Transformation von Sprache auf die Bildebene auszuloten. In finaler Wendung ist es dann möglich, das inter- bzw. plurimediale Amalgam von Text-Stimme-Musik als Generator von Bedeutungsüberschüssen einzufassen.

1. Einleitung

Im Rahmen unseres Basler Forschungsschwerpunkts ‚Populärkulturanalysen‘ widmen wir uns bereits seit langer Zeit der Untersuchung des audiovisuellen *Klein*-Formats Musikvideoclip. Wir haben uns zum einen mit dem Aspekt der *Distribution* von Clips, also dem Musikfernsehen (MTV, VIVA u.a.) beschäftigt. Dieses hat sich in den vergangenen Jahren in gravierender Art und Weise weiter entwickelt. Das klassische Umfeld des Fernsehangebeters MTV wurde programmlich stark verändert: MTV ist inzwischen zu einem TV-Sender für verschiedenste Jugend-affine TV-Angebote wie Soaps, Reality-Serien, Comics usw. geworden (offline) mit einem geradezu radikal ausgebauten Netzangebot (online) (für eine aktuelle Marktsichtung siehe Schmidt/Neumann-Braun/Autenrieth 2009).

Zum anderen haben wir uns mit den Clips selbst beschäftigt und *Produktanalysen*¹ durchgeführt. Diese haben sich in der Hauptsache auf die Aspekte Bild und Text bezogen und die Rekonstruktion von Orientierungs- und Deutungsmustern² – beides soziologische Termini – zum Ziel gehabt. Methodisch wurden die Songtexte auf texthermeneutischer Methodengrundlage interpretiert, sodann die Bilder der Clips: Kurz gesagt wurde dabei ein Gesamtablaufprotokoll erstellt, es wurden einzelne Clip-Teilsequenzen identifiziert und über das Verfahren von Schlüsselbildern, also einzelnen Stills, bildhermeneutisch analysiert. In einem nächsten Schritt wurden die Ergebnisse der Text- bzw. Bildanalyse in Beziehung gesetzt mit der Intention, den *intermedialen Zusatznutzen* der Verschränkung von Text und Bild zu bestimmen. Die Kernfrage liegt auf der Hand: Welcher Mehrwert ergibt sich daraus, dass ein Song nicht nur zu hören sondern (dabei) auch zu sehen ist?

Recht stiefmütterlich behandelt wurde bei diesem Vorgehen leider der Ton, was – vorsichtig formuliert – recht unglücklich ist, da bekanntlich in Clips bereits vorliegende Songs bebildert werden. In der Regel kreiert ein Musiker einen Song, der in der Folge von einem Regisseur visuell ‚umgesetzt‘ wird. Ein Song vereint also die beiden Elemente Text und Ton, die ihrerseits im Clip durch ein drittes Element, das Bild/die Bilder, erweitert werden. Und mehr noch: Der Songtext wird nicht im Medium der Schriftlichkeit präsentiert, also nicht als zu lesender Text, sondern er wird vorgelesen, gesungen, gesprochen – jedenfalls ‚performed‘. Clips sind also recht komplexe Phänomene – wie komplex wird im Weiteren noch genauer vorzustellen sein.

Es dürfte mit diesen wenigen Ausführungen bereits deutlich geworden sein, dass die bisherigen Analysen (eigene wie fremde) alles andere als zufrieden stellend zu bewerten sind – im Gegenteil: Sie sind deutlich unterkomplex und dies in zweierlei Hinsicht: Erstens werden die einzelnen Elemente des Clips für sich allein genommen nicht hinreichend untersucht – als Desiderat lässt sich hier insbesondere der Ton benennen. Zweitens wird der spezifischen Komponente der Intermedialität von Bild – Text – Ton kaum hinreichend Beachtung geschenkt.

Unser aktuell durchgeführtes Forschungsprojekt,³ vom Schweizerischen Nationalfond finanziert, widmet sich genau diesem Ziel – nämlich der Entwicklung eines Analyseinstrumentariums zu einer hinreichend umfassen-

¹ Vgl. Neumann-Braun (Hg.) (1999).

² Vgl. grundlegend Oevermann (2001a und b), Lüders (1991), Lüders/Meuser (1997), Meuser/Sackmann (Hg.) (1991).

³ Das Forschungsprojekt trägt den Titel „Bild-Text-Ton-Analysen am Beispiel der Gattung Videoclip. Entwicklung eines integrierten Analyseinstrumentariums zur Interpretation von Musikvideos“ und wird von Prof. Dr. Klaus Neumann-Braun am ifm Basel geleitet (Laufzeit 07/2008 bis 07/2011; SNF-Aktenzeichen: 100012–119936/1). Mitwirkende Forschungsmitarbeiter sind Daniel Klug und Dr. Christofer Jost sowie Dr. Axel Schmidt. Wir danken Daniel Klug und Arnulf Deppermann für instruktive Diskussionshinweise sowie Vanessa Kleinschnittger für die Durchsicht des Manuskripts.

den Bild-Text-Ton-Analyse audiovisueller Formate, das insbesondere in der Lage ist, das intermediale Verschränkungsverhältnis und dessen Wirkungen zu erhellen.

Der Rahmen dieses Artikels ist eng gesteckt, eine umfassende Einlösung dieses Desiderats ist naheliegenderweise an dieser Stelle nicht möglich. Wir wollen uns deshalb beschränken und uns erstens ‚nur‘ mit dem *Verschränkungsverhältnis von Musik und Sprache* beschäftigen – ein Verhältnis, dem innerhalb des Clips eine *besondere Rolle* zukommt. Zweitens stellen wir an dieser Stelle noch nicht das soeben annoncierte Methodeninstrumentarium vor sondern vielmehr ein *Begriffsinventarium*, das insbesondere an der Schnittstelle von Musik- und Kulturwissenschaft die Relation von Musik und Sprache *kriterial* einzuholen bemüht ist. Diese gleichsam auf den Audio-Bereich des Clips fokussierten Überlegungen reflektieren dieses Thema drittens aus der Perspektive der Medien- und Kommunikationswissenschaft sowie der Musikwissenschaft, also relevanten Nachbardisziplinen der Sprachwissenschaft.

2. Zur Bestimmung des Verhältnisses von Musik und Sprache im Musikvideoclip

Zur hinführenden Bestimmung des Verhältnisses von Musik und Sprache im Musikclip sollen zunächst *drei zusammenhängende Fragekomplexe* behandelt werden, welche zentral für unsere fokussierte Thematik sind und diese zudem einleitend verorten sollen:

- Welchen Konstitutionsprinzipien folgen Musikclips?
- Inwiefern sind Musikclips intermedial?
- Und: Inwiefern kommt dem Verhältnis von Musik und Sprache innerhalb von Musikclips eine besondere Rolle zu?

Begonnen werden soll mit folgender *Ausgangsbeobachtung*: Wer sich mit Musikclips – zunächst verstanden als audiovisuelle Medienprodukte, welche einen Popsong zu Werbezwecken visualisierend begleiten und i.d.R. via TV-Ausstrahlung distribuiert werden⁴ – beschäftigt, stellt recht schnell dreierlei fest:

- Es handelt sich erstens um *semiotisch komplexe Phänomene* (mehrere Zeichensysteme und Gestaltungstraditionen⁵ sind in unterschiedlichen Bezugsverhältnissen involviert), welche

⁴ Weitere Definitionen des Musikclips finden sich bei Bergermann (2003), Keazor/Wübbena (2005), Weibel (1987) und Neumann-Braun/Schmidt (1999).

⁵ Lull (1987) betont, dass Clips als Synthesen aus verschiedenen (populären) Kunstformen zu verstehen sind: Sie beinhalten nicht nur die visuelle Präsentation eines Songs, sondern darüber hinaus Tanz, Choreographie, Schauspielerei, Narrationen/Stories, Mode, Kostümierung/Styling, Beleuchtung, Starinszenierungen und visuelle Techniken (etwa Effekte und Animation).

- zweitens gezwungen sind, ein künstlerisches Konzept auf kleinstem Raum zu entwickeln,⁶ was sie
- drittens häufig dadurch erreichen, dass sie *mit traditionellen Darstellungskonventionen brechen* (etwa Erzähltraditionen oder filmischen Konventionen).

Clips wirken daher häufig *dicht, überladen und zusammenhangslos* – letzteres insbesondere dann, wenn man auf die Bildebene fokussiert und diese an üblichen Film- bzw. Erzählcodes misst (also etwa minimale Anforderungen an eine Geschichte, Kausalität/*plot*, Aufbau einer fiktionalen, diegetischen Welt etc.). *Kohärenz* erhält die Bildebene des Clips u.a. durch den Rekurs auf den *Akt der Klangerzeugung bzw. die musikalische Aufführungspraxis*, also unter Rückgriff auf eine durch die Alltagserfahrung vermittelte (soziale) *Handlung*.⁷ Diese ist mit Blick auf das dem Clip zugrunde liegende *Primärprodukt*, nämlich den Popsong (als ‚Werk‘) bzw. den Tonträger (als Speichermedium) bzw. die Schallaufzeichnung (als technisch-apparativer Speicherprozess) allerdings *weder sichtbar noch rückholbar* und das aus drei Gründen: Erstens weil die realweltliche Klangerzeugung bzw. Aufführung als einmaliges Ereignis unwiederbringlich Vergangenheit geworden ist (*Flüchtigkeit*); zweitens bleibt die aufgezeichnete Musik immer dieselbe, ist und kann also nicht Resultat einer zukunfts offenen (sozialen) Handlung im Hier und Jetzt sein (*Reproduzierbarkeit*);⁸ und drittens wird Akustisches und Visuelles qua Speichermedium zeitlich und räumlich *entkoppelt* (Effekt der Phonographie).⁹ Dennoch ist die musikalische Aufführungspraxis sehr wohl *imaginierbar* oder – stärker formuliert – wird durch Musik und insbesondere die Popmusik, welche erst in und durch Aufführungspraxen lebt, *imaginativ auch forciert*.¹⁰

⁶ Die meisten Clips bewegen sich in einem Zeitrahmen von 3–5 Minuten und sind damit an die übliche Länge eines Popsongs gebunden. Diese für ein audiovisuelles Produkt relative Kürze wurde in der Literatur immer wieder als konstitutives Merkmal von Musikclips betont und zwingt die Clipregisseure – grob gesprochen – zu Strategien der Komprimierung (was im Einzelfall höchst Verschiedenes bedeuten kann, nämlich etwa Bilder nur assoziativ zu verketteten, ikonisch aufgeladene (Schlüssel-)Bilder zu bemühen, auf narrative oder gegenständliche Bilder ganz zu verzichten, mit Redundanzen zu arbeiten u.v.m.).

⁷ Vgl. Wulff (1999).

⁸ Sie bedarf daher als *phonographische Musik* zunächst keiner Interpretation (sondern nur das Drücken einer play-Taste), ganz im Gegensatz zur Aufführung von Kunstmusik, welche immer als Interpretation eines Werks begriffen wird (vgl. Helms 2003b).

⁹ Dies fundiert die sog. *acousmatische Erfahrung* (Großmann 1998, S. 110) bzw. *acousmatischen Sounds*. Chion fasst diese (unter Rekurs auf Pierre Schaeffer) als „sounds one hears without seeing their originating cause“ (1994, S. 71 ff.) und versteht die betreffenden Medien als „*acousmatic media*“, das sind „all which transmit sounds without showing their emitter“ (ebd.). Visualisierung respektive ‚Acousmatisierung‘ begreift Chion als (In-)Visibilisierung der Klangquelle. Dies geschieht im Musikclip, wenn die Darstellung zwischen Performance und Nicht-Performance hin- und herwechselt (allerdings auch im Falle der Performance auf simulativer Ebene; siehe unten).

¹⁰ Rösing (2003) unterscheidet systematisch verschiedene Möglichkeiten audiovisueller Musik-Wahrnehmung respektive verschiedene Formen der Verkopplung von Auditivem und Visu-

Grundlage hierfür sind in und durch Alltagserfahrungen (Welt- und Sozialisationsidealisierungen) fundierte ‚Normal- bzw. Ursprungsformen‘¹¹ realweltlicher Musikaufführungen, welche als rezeptionslenkendes Hintergrundwissen der audiovisuellen Wahrnehmung von Clips fungieren und im Falle von Musikclips die Form annehmen, dass a) gehörte Musik durch Musiker erzeugt wird (Musik wird bewirkt; Kausalität/Notwendigkeit) und b) diese Musik typischerweise in der Lage ist, musikrezeptionstypisches Anschlussverhalten (etwa Tanz) hervorzubringen (Musik bewirkt etwas; Möglichkeit/Wahrscheinlichkeit/Typikalität). Hieraus ergibt sich ein *für die Konstitution des Clips konstitutiver Effekt: Musik als Klangereignis lässt als Bewirktes Bilder ihrer Erzeugung respektive Aufführung (Performance) und als Bewirkendes Bilder typischen Rezeptionsverhaltens* (etwa Tanz, konzeptionell: Choreographie) ‚natürlich‘ bzw. ‚tonaffin‘ erscheinen (siehe Schaubild 1 und 2), so dass aufgrund der aus der natürlichen Wahrnehmung herrührenden Erfahrung visualisierte klangerzeugende Ereignisse kausal interpretiert werden (also gewissermaßen ‚naturalisiert‘) und daher die visuelle Komponente ‚Musik-Performance‘ im Musikclip *simulativ musikerzeugend* eingesetzt wird. Mit anderen Worten: Durch die Darstellung einer *Musik-Performance* im Clip entsteht der Eindruck, es handle sich um eine *Dokumentation der Aufführungspraxis bzw. gar um eine (zumindest das akustische) Medium erzeugende Handlungswiedergabe*.¹² Oder um mit Helms zu sprechen: „Das Video gibt vor [...] Musik-Machen beobachtbar zu machen.“¹³ All dies steht Musikclips als strukturelles Potenzial der Be-

ellem im Rahmen von Musikrezeption, welche bei der natürlichen Einheit musikerzeugender Praxis ihren Ausgang nimmt (Konzert), unterschiedliche Stufen der Imagination durchläuft, um schließlich bei ‚realen‘ (also nicht bloß imaginierten) Visualisierungen a) im Zuge der Entkopplung von Original-Ton und visueller Wahrnehmung (etwa Reisesmusik), b) im Rahmen von Primärmedien (Bühnenmusik wie im Falle von Ballet, Oper etc.) sowie c) im Rahmen audiovisueller Produkte (Film, Musikclip) anzukommen.

¹¹ Vgl. hierzu Großmann (1998, S. 109).

¹² Helms (2003a) folgend lässt sich zeigen, dass Clips wie *Handlungswiedergaben* wirken. Denn: Durch die Simulation einer Musik-Performance, die gerade aufgezeichnet wird (erste Simulationsebene), entsteht zudem der Eindruck, dass diese Performancehandlung zugleich auch das Medium selbst, also den Clip, erzeugt (zweite Simulationsebene). Obwohl beides – sowohl die den Clip erzeugenden Handlungen (also die Kamerahandlungen des ‚korporierten Regisseurs‘ i.S.v. Reichertz (1992, 2005)) als auch die Materialität des Mediums (Magnetband, Datei, Übertragung etc.) – visuell nicht zugänglich ist. „Das Video gibt vor – augenfälliger als die Schallaufzeichnung –, Musik-Machen beobachtbar zu machen. Dagegen entzieht sich dem Beobachter das Ding, das Material der Aufzeichnung“ (Helms 2003a, S. 102). Und weiter: „Die Unzulänglichkeit der dinglichen Gestalt des Mediums lässt den Eindruck entstehen, die am Bildschirm oder Lautsprecher beobachtete Handlung sei identisch mit der Handlung des Musikers, die das Medium erzeugte. Es entsteht die Vorstellung einer ‚Aufzeichnung‘“ (ebd., S. 103).

¹³ Dieser Eindruck wird durch einen weiteren, technischen (Zwischen-)Schritt der Medienentwicklung fundiert: Während Phonographie und Photographie das Akustische und Visuelle trennten, isolierten und in Daten unterschiedlicher Physikalität zerlegten (vgl. Kittler 1999), überwand der Tonfilm durch entsprechende optoelektrische Verfahren (Lichtton) diese Trennung wieder. Auf diese Weise wurden Aufnahmen möglich, die sowohl das Akustische

deutungsgenerierung prinzipiell zur Verfügung und kommt umso stärker zur Geltung, je mehr und ‚natürlicher‘ (respektive unartifizierter) ein Clip mit Performance- respektive Choreographie-Anteilen arbeitet.¹⁴

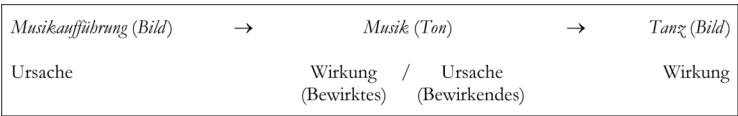


Schaubild 1: Musik als Bewirkendes und Bewirktes

Wählt der Clip allerdings als visuelle Darstellungsinhalte *weder Performance¹⁵ noch Tanz¹⁶* und hält zudem die *visuelle Binnenstruktur ungegenständlich und non-narrativ¹⁷* (also etwa im Falle von Graphiken), so entsteht der Eindruck eines *invertierten Bild-Ton-Verhältnisses*.¹⁸

Dieses kommt – grob gesprochen – dadurch zustande, dass die natürliche Wahrnehmungseinheit der Musikaufführung aufgelöst und die dadurch entstehende ‚visuelle Leerstelle‘¹⁹ mit *zunächst ‚tonfremd‘ erscheinenden Bildern aufgefüllt* wird, welche sich mehr oder weniger weit von der (imaginierbaren) ‚natürlichen‘ Wahrnehmungssituation Musik erzeugender Musiker entfernen (weit: Konzeptclip; nah: Performance-Clip) und aufgrund der simultanen Darbietung im Audiovisuellen zwangsläufig auf die Musik bezogen werden: Da das Visuelle nun nicht mehr den Ton zu erzeugen scheint (der musizierende Musiker ist verschwunden) und auch nicht mehr als typische Folge der Musik gelesen werden kann (es tanzt auch niemand), scheint nun in Ermangelung einer ‚natürlichen Erklärung‘ *in umgekehrter Weise der Ton das Bild hervorzubringen* bzw. das Bild erscheint als *Visualisierung* des Tons. Aufgrund dessen sind Musikclips in besonderer Weise in der Lage mit (populärkulturellen) Versatzstücken²⁰ zu arbeiten, also sinnhafte Zusammenhänge bloß fragmentarisch anzudeuten und assoziativ zu verketten, da solche als zunächst

als auch Visuelle einer ‚realen‘ Situation synchron fixierten, kurz: Aufzeichnungen im Sinne von Dokumentationen (*footage*). Diese neuerliche mediale ‚Normalität‘ (Film und Fernsehen) macht sich der Musikclip zunutze, um sie seinerseits wieder zu verfremden.

¹⁴ Dass Performance(-anteile) in Clips unterschiedlichen Graden der Verfremdung unterliegen und dieses Kriterium als systematischer Ausgangspunkt einer Typologisierung von Musikclips gelten kann, hebt insbesondere Altrogge (2000) hervor.

¹⁵ Siehe Beispiel 1 (Still aus Green Day, „American Idiot“).

¹⁶ Siehe Beispiel 2 (Still aus Gwen Stefani, „Hollaback Girl“).

¹⁷ Siehe Beispiel 3 (Still aus Gnarlz Barkley, „Crazy“).

¹⁸ In besonders radikaler Form gilt dies im Rahmen oszilloskopischer Anordnungen, da das Visuelle in diesem Fall eine technisch-physikalische Transformation des Tons darstellt, also im Peirce’schen Sinn als indexikalisches Zeichen zu verstehen wäre, wodurch das Bild nicht bloß so wirkt, als sei es vom Ton hervorgebracht, sondern es in einem physikalischen Sinn auch tatsächlich ist.

¹⁹ Vgl. Großmann (1998, S. 110).

²⁰ Etwa Zeitungsbilder, Portraits/Konterfeis, Bildikonen, konventionalisierte Bildfolgen, pop-musikalische Aufführungen als eingestreute Sinnfragmente etc.

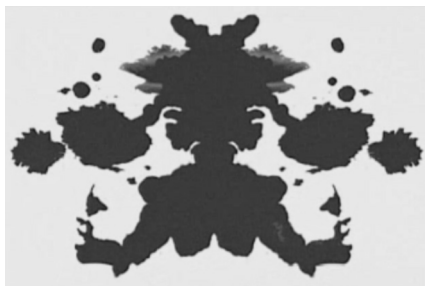
,tonfremd‘ wahrgenommenen Bilder eben unter der ästhetischen Prämisse der Inversion ‚gelesen‘ werden.²¹ Diese Merkmale der *Zitativität, Assoziativität und Inversion* mach(t)en den Musikclip zu einem Paradefall der sich etwa zur gleichen Zeit (1980er Jahre) etablierenden Intermedialitätsforschung.



Beispiel 1: Green Day – „American Idiot“



Beispiel 2: Gwen Stefani – „Hollaback Girl“



Beispiel 3: Gnarlz Barkley – „Crazy“

²¹ Die meisten Clips zeigen durch Performance- und Choreographieanteile auch, was näher läge bzw. erwartbarer/,natürlicher‘ wäre und etablieren damit einen archimedischen Punkt, wodurch ,tonfremde‘ Bilder als ,Abweichungen‘ von einem Idealtypus (im Weber’schen Sinn) markiert werden.

Versucht man nun vor diesem Hintergrund Musikclips mit Blick auf gängige *Intermedialitätssystematiken* zu verorten,²² so lässt sich zeigen, dass – neben vielen anderen möglichen Berührungspunkten – eine für systematische Zwecke *besonders bedeutsame, strukturelle Eigenheit des Musikclips* ins Auge sticht, nämlich: Die prototypisch als intermedial begriffenen Phänomene im Clip entstehen *vornehmlich im Verhältnis Bild-Ton (= AV)*,²³ welches jedoch seinerseits auf einem *basaleren, intermedialen Verhältnis der Tonspur selbst* beruht, nämlich dem *Verhältnis von Sprache und Musik in der musikalischen Performance*. Dass diese für Musikclips grundlegende Beobachtung plausibel und für das vorliegende Thema relevant ist, lässt sich mit Blick auf einschlägige Erkenntnisse der Intermedialitätsforschung wie folgt zeigen:

Erstens lässt sich unschwer erkennen, dass Musikclips zum einen *plurimediale* Phänomene sind, d.h. – der Systematik Rajewskys (2002) folgend – konventionell als distinkt wahrgenommene Medien sind, die mehrere Zeichensysteme involvieren; aber zum anderen aufgrund ihrer – historisch betrachtet – relativen Neuartigkeit ebenso gut als *intermediale* Phänomene im Sinne einer *Medienkombination* begriffen werden können, d.h. als Verschränkung mehrerer Medien und/oder Symbolsysteme (also etwa Musik und Filmbild oder Bildlichkeit und Oralität usw.), welche allesamt *im Endprodukt material präsent* sind.

Zweitens lässt sich engführend auf das *AV-Verhältnis* unschwer zeigen, dass oben angedeutete Strukturmerkmale des Musikclips Korrespondenzen mit gängigen Konzepten und Subdimensionen von Intermedialität aufweisen. Dies sei hier nur angedeutet:

- Das Merkmal der *Zitativität* korrespondiert mit dem Konzept der *intermedialen Bezugnahmen*,²⁴ verstanden als *Rekurse auf altermediale Produkte*

²² Verwendet wurden hier vor allem die Arbeiten von Rajewsky (2002), Schröter (1998) und Leschke (2007).

²³ Ausgeblendet bleibt hier die Materialität des jeweiligen Trägermediums (also: Filmstreifen vs. Magnetband vs. Datei) bzw. die technisch-apparative Ebene (also: Filmkamera/-projektor vs. Videokamera/-player/Bildschirm vs. digitale Kamera/Monitor), damit auch die (technische und perzeptive) Spezifik des generierten Bildes (Leinwandbild vs. Monitorbild) sowie die entsprechenden Mediensysteme (Kinematographie, Video, digitaler Film). Dies geschieht trotz des Umstands, dass gerade die Mediendifferenz ‚Film/Video‘ als typisch für die Gattung ‚Videoclip‘ erachtet wurde (vgl. etwa Kerscher/Richard 2003), da diese Differenz als Abgrenzungskriterium wohl für die Anfänge des Genres, nicht aber für heutige Produktionen geeignet ist. Entsprechend wurde die Bezeichnung ‚Videoclip‘ respektive ‚Musikvideo‘ – einem Vorschlag von Jacke (2003) folgend – durch die Bezeichnung ‚Musikclip‘ ersetzt. Die vorliegende Betrachtung beschränkt sich zudem auf die durch Apparate generierte Oberfläche (Bildschirm), d.h. auf das Produkt als kommunikativ-semiotisches Phänomen unter Ausblendung seines technisch-apparativen Zustandekommens.

²⁴ Die Idee der intermedialen Bezugnahme nimmt ihren Ausgang beim Konzept der Intertextualität (siehe zusammenfassend Fix 2001) bzw. Transtextualität (nach Genette 2004) und erweitert diese um den Aspekt der Mediendifferenz im Falle der *Intermedialität* im Vergleich zur *Intertextualität*. Während im Falle der *Intertextualität* keine Mediengrenzen überschritten

*und/oder semiotische Systeme, wobei letztere im kontaktnehmenden Medium material nicht präsent sind.*²⁵

- *Transmedialität*, verstanden als die Wanderung *medienunspezifischer* Phänomene, lässt sich am Merkmal der *Assoziativität* festmachen, da der Musikclip hier auf formalästhetische Gestaltungstraditionen und -prinzipien insbesondere des frühen Avantgardefilms (etwa Collage, Rhythmus als gemeinsame Grundlage von Auditivem und Visuellem, Sichtbarmachung von Sound etc.) zurückgreift.²⁶
- Das *invertierte Bild-Ton-Verhältnis* schließlich verweist auf einen intermedialen Produkten häufig zugeschriebenen Effekt, nämlich jenen der *Synästhesie*.²⁷ Hiermit ist das bereits oben angedeutete Phänomen gemeint, dass Clips den Eindruck eines Hörens von Bildern vermitteln bzw. von Bildern, die musikalisch-klangliche Qualitäten zu haben scheinen.²⁸

werden (Bezug von Literatur auf Literatur, von Text auf Text), kann sie aus der Perspektive der *Intermedialität*, für welche die Überschreitung von Mediengrenzen konstitutiv ist (etwa Bezug von Literatur auf Film), als ein (bedeutsamer) Fall von *Intramedialität* (Bezugnahmen innerhalb eines Mediums) gelten.

²⁵ Mit Spielmann (1998) lassen sich solche Bezugnahmen etwa am Beispiel des Films zeigen, welcher auf die Malerei dergestalt rekurren kann, dass Gemälde im Film zu sehen sind oder aber die Bildgestaltung der Malerei (bzw. des Tafelbildes) *formale Prinzipien* (etwa eine zentripetale Art der Kadrierung oder eine an zentralperspektivischen Idealen orientierte Kameraführung (dies zeigt Spielmann an den Filmen von Peter Greenaway)) entlehnt, ohne dass der Film in einem ontologischen bzw. substantiellen Sinn Malerei sein könnte, da er – was die medial-materiale Basis anbelangt – auf die Fixierung von Lichtverhältnissen auf Zelluloid und deren Projektion angewiesen ist und die Mittel der Malerei – Leinwand und Farbe – daher nur simulieren, nicht aber tatsächlich verwenden kann.

²⁶ Vgl. hierzu Bódy/Weibel (Hg.) (1987); Deutsches Filmmuseum Frankfurt (Hg.) (1993); Paech (1994).

²⁷ Clips sind konzeptionell darauf angelegt, Ton zu visualisieren (siehe auch Rösing 2003); daher auch die (ent-)sprechenden Bezeichnungen wie „Augenmusik“ (Barth/Neumann-Braun 1996), „Tönende Bilder“ (Altrogge 2000), „Werbende Klangaugen“ (Hausheer 1994), „Visueller Sound“ (Hausheer/Schönholzer (Hg.) 1994), „The Look of Sound“ (Aufderheide 1986) und „Visuelles Radio“ (Bechdolf 1996), durch welche Musikclips häufig charakterisiert wurden. Daher lässt sich auch – Schmidbauer/Löhr (1996) folgend – beim Clip weniger von einem *Soundtrack* (Tonspur eines Films), also einer Musik zum Bild, als vielmehr von einem *visual track* zur vorgängigen Musik sprechen.

²⁸ Zu verstehen ist dies im Sinne einer Farblichtmusik, d.h. einer Verkopplung auf der Ebene der technischen Signale und nicht der Symbole (vgl. Großmann 1998, S. 111), welche entweder auf subjektiver, synästhetischer Intuition (prominentestes Beispiel: das Farbenklavier des russischen Komponisten Alexander Skrjabin) oder auf objektiver, physikalischer Transformation (etwa im Falle des ‚Optophons‘ (Raoul Hausmann, 1922)) beruhen kann. Insbesondere die Arbeiten des Experimentalfilmers Oscar Fischinger – so Großmann (1998) – legen Zeugnis vom Versuch ab, die Grenzen von Ton- und Bildkanal zu überwinden (seine „Ton-Ornamente“ (1932) etwa versuchen dies gewissermaßen physisch, da sie eine Visualisierung der Tonspureinschreibungen im Tonfilmstreifen, also innerhalb einer mit Lichttonverfahren arbeitenden Filmkamera darstellen). Musikclips folgen *grosso modo* eher ersterem Prinzip, wobei der synästhetische Eindruck die Suche sowohl nach formalen (etwa Cut-Rhythmus-Korrespondenzen) und als auch semantischen (etwa Songtext-Bild-Relationen in Form von Illustrationen) Konnexen evoziert.

Der entscheidende Punkt ist nun *drittens*, dass Clips häufig auf obige, weil hervorstechende und auffällige, *pluri- bzw. intermediale Merkmale reduziert* werden und dabei das *primäre und tiefer greifende Verhältnis von Sprache und Musik übersehen* wird. Stärker formuliert: *Erst diese asymmetrische Verschachtelung von Symbolsystemen im Musikclip* vermag jene intermedialen Verhältnisse, jene Ästhetik hervorzubringen, für die der Clip sprichwörtlich geworden ist („Clip-ästhetik“) (siehe Schaubild 2).

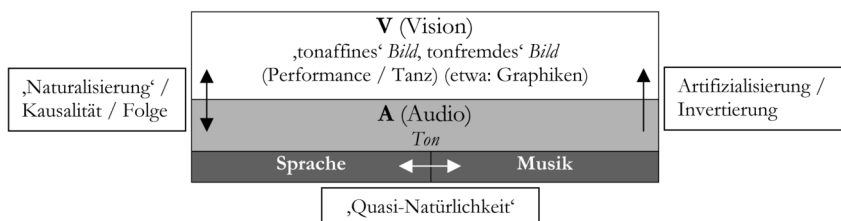


Schaubild 2: Asymmetrische Verschachtelung der Symbolebenen und Bild-Ton-Verhältnisse im Clip

Daher soll im Folgenden der *Audio-Bereich des Clips fokussiert* werden, also die Tonspur, da diese produktionsästhetisch primär ist und – mit Blick auf Semiose-Prozesse – aus oben genannten Gründen ‚natürlicher‘ wirkt. Umgekehrt ist es gerade deshalb auch jener Bereich, der *schwieriger ‚auseinanderzudevidieren‘* (also zu analysieren) ist, der in Musikclipanalysen *häufig stillschweigend vorausgesetzt* wird und dessen *intermediale Bedeutungspotenziale aufgrund dessen vernachlässigt* respektive gar nicht erst analytisch fruchtbar gemacht werden.

3. Text-Ton-Relationen und Populärmusikforschung

Wenn im Folgenden von der Relation von Musik und Sprache die Rede ist, so soll insbesondere auf die *analytischen Anforderungen* eingegangen werden, die sich mit diesem Beziehungsgeflecht in populärer Musik verbinden. Es sei an dieser Stelle hervorzuheben, dass in jüngster Zeit eine Reihe von Beiträgen zur *Sprache in populärer Musik* erschienen ist, die in unterschiedlichen Perspektivierungen die phänomenalen Eigenschaften von Sprache in Pop-Stücken ausloten und dadurch insgesamt zu einem differenzierten Erkenntnisstand bezüglich dieser Bedeutungsebene beitragen.²⁹ Ihre informative Dichte kann jedoch nicht darüber hinwegtäuschen, dass in der Populärmusikforschung eine allgemeine Unklarheit bezüglich des analytischen Umgangs mit den klanglichen Dimensionen und in der Folge mit der Bedeutungskonstitution durch musikalische Klanglichkeit herrscht.³⁰ Die

²⁹ Vgl. Bielefeldt (2008); vgl. Klein (2008); vgl. Bowman (2003).

³⁰ Vgl. Wicke (2003).

folgenden Ausführungen sind demnach an die globale Zielperspektive gekoppelt, den reflexiven Schärfegrad im Spannungsfeld von Material und Analyse zu erhöhen. In diesem Kontext werden auch grundlegende theoretische Fragestellungen zur Populärmusikforschung tangiert. Im Hinblick auf dieses Vorhaben kommt der Verschränkung der Symbolsysteme populäre Musik und Sprache durchaus eine Schlüsselposition zu. Gelingt es entlang der in diesem semiotischen ‚Fusionskern‘ eingelagerten intertextuellen Rekurspotentiale ein Beobachtungssystem zu etablieren, so ist zu erwarten, dass hiervon Impulse für die Beobachtung weiterer klangübergreifender Zusammenhänge in populärer Musik ausgehen. Nachfolgend sind es nun Überlegungen zur *Pragmatik von Gesang*, die in diesen Methodologie-Komplex einführen.

Im Rahmen von Populärmusik- und Clipforschung gilt es als gesicherte Erkenntnis, dass das Verhältnis von Ton und Text subsidiär in den globalen Zusammenhang von Musik und Sprache eingelagert ist.³¹ Hierdurch wird nicht grundlegend die Sinnhaftigkeit einer Text-Analyse in Frage gestellt. Entscheidend ist nur, welchen Stellenwert man ihr in einer integrierten (Bild-)Text-Ton-Analyse zuteil werden lässt. Angesichts der phänomenalen Zusammenhänge in einem sprachbasierten Musikstück erscheint diese Analyse als Nebenschauplatz. Denn der Text wird *gesungen/ vorgetragen*, materialisiert sich also erst im performativen Vollzug der Stimme.³² Jene Materialisierung durch die Stimme bleibt während eines Konzerts flüchtig, erscheint im Speichermedium Tonträger indes in verewigter Form. Die Sprachzeichenhaftigkeit von Gesang wird dadurch nicht grundlegend in Abrede gestellt. Für die Konstituierung und Sedimentierung von Bedeutung in vokalen Formaten sind jedoch nunmehr die *Relationen* zwischen Materialhaftigkeit einer Performance und Zeichenhaftigkeit des textuellen Basismaterials bedeutsam.³³ Den Text singen bedeutet gleichsam, Sinnzusammenhängen in expressiver Flüchtigkeit eine Gestalt zu geben, auf Bedeutungen *aus* zu sein. Gegebenenfalls kann dies auch zur Folge haben, dass Bedeutungen des Texts unterminiert werden, er also eine Umwertung über sich ergehen lassen muss. Sei es, er wird nach einer spezifisch musikalischen Logik neu-rhythmisiert, oder bestimmte Silben werden verschluckt oder einfach aufgrund des sinnlichen Potentials von Melos und Timbre, das Sinn temporär außer Kraft setzt oder in den Hintergrund drängt.³⁴

³¹ Vgl. Vernallis (2002); vgl. Frith (1996, S. 159).

³² Siehe hierzu u.a. Krämer (2006).

³³ Der Zusammenhang von Materialhaftigkeit und Zeichenhaftigkeit einer Aufführungssituation nimmt allgemein eine zentrale Position in den Diskussionen zum Performativitätsbegriff ein. Allgemein ist in den kulturwissenschaftlichen Diskursen eine verstehenskategorische Verschiebung zugunsten der Begriffe *Ereignis* oder *Präsenz* zu beobachten. Für den Bereich der performativen Ästhetik siehe hierzu Fischer-Lichte (2004). Im Hinblick auf epistemologische Fragestellungen siehe Gumbrecht (2004).

³⁴ Vgl. Krämer (2006, S. 280 ff.).

Die genannten Aspekte treffen auch auf den Spezialfall von (*gesungener*) *Sprache in populärer Musik* zu. Dem ist jedoch als wesentlicher Aspekt anzufügen, dass das Vokale in populären Bereichen auf den spezifischen Traditionslinien von Blues, Jazz und Schlager aufbaut.³⁵ Bereits in jenen historischen Sparten war der Gesang von einer eigentümlichen Verschmelzung von sozialem Verbalisationsgestus und traditionellen gesanglichen Elementen durchdrungen. Eine weitere Ausdehnung des vokalen Repertoires im Sinne fortwährender Synthetisierung fand dann mit Beginn der Rock'n'Roll-Ära und in den Folgejahrzehnten statt. Insbesondere die dem Alltäglichen entstammenden Formen der Verbalisierung wie z.B. Schreien, Sprechen, Flüstern – direkt zu veranschaulichen an Genres wie Heavy Metal, Hip Hop und Chanson – haben mehr und mehr zu einer stilistisch-ästhetischen Etablierung von Pop-Sprache und -Stimme jenseits des traditionellen Gesangs³⁶ geführt. Das Phänomen *Rap* ist diesbezüglich nur eines unter vielen repräsentativen Beispielen, das aufzuzeigen vermag, dass Sprache und Stimme im Sinne einer eigenen poplarmusikalischen Idiomatik funktionieren.

Wendet man nun die Ausführungen zur Spezifik der vokalen Gestaltung in populärer Musik zurück auf die analytischen Anforderungen von populärer Musik, so wird deutlich, dass der *Gesangsbegriff* aufgrund seiner präskriptiven Anteile nur begrenzt von Nutzen ist. Demgegenüber vermeidet die Bezeichnung *vokale Performanz* begriffsbedingte Vor-Aussagen. Eine begriffliche Neu-Orientierung an ebenjener Bezeichnung sei an dieser Stelle empfohlen, denn sie führt über die Zwischenstufe einer Ausweitung des analytischen Vokabulars schließlich zu einem höheren reflexiven Schärfegrad. Hiernach ist durch *vokale Performanz* eine *übergeordnete Verstehenskategorie* geschaffen, entlang derer es möglich ist, klanglich-vokale Ereignisse auf das eigentümliche Wechselspiel von gesanglichen Anteilen und alltagssprachlichen Gesten hin zu durchleuchten.³⁷

³⁵ Vgl. Wicke (1992, S. 452 ff.). Die genannte Trias rekurriert auf den historischen Tatbestand, dass populäre Musik im Allgemeinen und das Vokale im Speziellen sich in großen Teilen aus folkloristischen Elementen, vor allem afro-amerikanischer Provenienz, speisen. Im gleichen Zuge darf aber der Einfluss der europäischen Spiel- und Singtradition – hier repräsentiert durch den Schlager – nicht marginalisiert erscheinen. So zeichnen sich im Schlager die Traditionslinien der populären europäischen Musiktradition ab (vgl. Middleton 2001, S. 63).

³⁶ Die Bezeichnung ‚traditioneller Gesang‘ wird an dieser Stelle nicht als polares Gegenmodell zu vokaler Performanz in populärer Musik aufgestellt. Vielmehr gilt es die Bezeichnung als Eingrenzung von vokalen Gattungen zu begreifen, die in Herkunft, Aufführungspraxis und Stilistik eine geringe Nähe zur populären Musik aufweisen. Hierzu zählt zum einen der Klassische Gesang, der u.a. durch Begriffe wie Konzertgesang, Operngesang, Lied- und Oratorien-gesang repräsentiert wird. Zum anderen ist hiermit auf das Gros der volksmusikalischen Gattungen weltweit verwiesen, die auf orale Tradierung und organisch-akustische Klang-erzeugung aufbauen.

³⁷ Simon Frith bewertet die Integration von Umgangssprache in die ästhetische Form des Songs als eigentümliche ‚Erhöhung‘ ebenjenes Sprachstils. Die Verschränkung von Umgangssprache und Musik ist infolgedessen Keim fortwährender Spannungsproduktion (vgl. Frith 1996, S. 168). Bei Dieter Baacke finden sich Ansätze einer gesamtheoretischen Fun-

Mit dem Verweis auf die humanen respektive sozialen Eigenschaften von vokaler Performanz in populärer Musik ist ein erster Schritt zur analytischen Neu-Vermessung vollzogen worden. Der zweite Schritt erfolgt über den Rekurs auf ein weiteres determinierendes Bezugssystem: *die Aufnahme-technik*. Das Aufnahme-Studio hat mit seinen sich ständig vervollkommnenden technischen Möglichkeiten zur Differenzierung und Beeinflussung von Klang, insbesondere seit Einführung der Mehrspurtechnik, einen immer größeren Stellenwert in der Entwicklung der populären Musik bekommen.³⁸ Die Aufnahme selbst gründet sich auf den Dreischritt Tonaufnahme, Abmischung und Mastering. Die musikalische Performance im Studio wird in Form einer Tonspur fixiert. Mehrere Tonspuren übereinandergelegt ergeben dann den Song bzw. den Tonträger. Auf der Basis von Soundprozessoren (auch Klangeffekte genannt) und Mastering-Einstellungen kann die klangliche Wahrnehmung entscheidend beeinflusst werden.³⁹

Die Aufnahme vermag es, den Interpreten bzw. die Interpreten präsent erscheinen zu lassen, wenn beispielsweise durch das Zusammenspiel von Echo- und Kompressor-Effekten in der Wahrnehmung des Hörers bestimmte Vorstellungen von Körperlichkeit und Räumlichkeit hervorgerufen werden. Dem ist aus Sicht des Performativitätsdiskurses anzufügen, dass die Stimme auf dem Tonträger nicht im Sinne tatsächlicher Präsenz erfahren wird. Vielmehr wird auf Grundlage von Präsenz-Effekten der *Schein von Gegenwärtigkeit* erzeugt.⁴⁰ Nichtsdestoweniger bleibt der ursprüngliche Präsenzkontext, in dem Stimme entsteht, auf dem auditiven Speichermedium subjektiv imaginierbar. Die Aufnahme ist gewissermaßen mit performativen Anteilen aufgeladen. Simon Frith leuchtet ebenjene Aspekte stimmlicher Medialität und Performativität weiter aus. Er formuliert diesbezüglich die Ansicht, dass die grundlegende Vorstellbarkeit stimmlicher Ereignis-

dierung des populären Felds, die eine ähnliche Richtung einschlagen. So geht Baacke in Anlehnung an den Bricolage-Begriff davon aus, dass das kulturelle Territorium der populären Musik geprägt ist durch eine ständige Neuordnung und Rekontextualisierung von Objekten, Tönen und Kombinationen, mit dem Ziel, neue Bedeutungen herzustellen und damit das vorhandene Gesamtsystem von Bedeutungen neu zu ordnen (vgl. Baacke 1998, S. 45). Dass diese Entstehung kultureller Bedeutung nicht allein Produkt symbolischer Kommunikation zwischen Akteur und Rezipient ist, sondern sich auch auf ökonomische Belange zurückführen lässt, muss an dieser Stelle ins Gedächtnis gerufen werden. Die Identifikation der sich dahinter verborgenden Mechanismen ist Aufgabe musiksoziologischer Forschung und kann entsprechend im Rahmen dieses Beitrags nicht erfolgen. Es sei diesbezüglich auf die Arbeit Peter Wickes verwiesen, der sich rezeptions- und markttheoretisch dem populären Gesamtfeld annähert und aus dieser Perspektive zu der Einsicht einer zunehmenden Pluralisierung und Fragmentarisierung populärer Musik gelangt (vgl. Wicke 1993, S. 36 ff.).

³⁸ Vgl. Wicke/Ziegenrucker (2001, S. 515).

³⁹ Vgl. Hawkins (2000). Hawkins gelangt in seiner Materialanalyse zu der Einsicht, dass die harmonische Komplexität des betrachteten Songs auf die spezifische Bearbeitung und Abmischung der Tonspuren zurückgeht, d.h. die Medialität des Tonträgers rahmt Produkt und Rezeption gleichermaßen.

⁴⁰ Vgl. Fischer-Lichte (2004, S. 174).

kontexte ein zentrales theoretisches Apriori zum Phänomen Pop-Stimme bildet. Darauf aufbauend apostrophiert er *vier Imaginationsebenen*, die die Interpretationsmöglichkeiten von Stimme in populärer Musik markieren. Hiernach wird die Stimme in einem Song vorgestellt als musikalisches Instrument, Körper, Person und Charakter.⁴¹ Alle vier Ebenen werden analytisch-interpretativ relevant gesetzt. Für unseren methodologischen Blick auf das Themenspektrum Text-Ton ist dies dahingehend bedeutsam, als dass hierdurch das Erfordernis artikuliert wird, um das Phänomen Stimme herum einen transdisziplinär verständlichen Begründungszusammenhang zu entwickeln, der sowohl spezifisch musikalischen als auch performativen Perspektivierungsansätzen standhält.

Führt man nun die ursprüngliche Frage nach der Bedeutung des Texts in Popsongs mit den vorangegangenen Ausführungen zusammen, so lässt sich als eine Art *Zwischenfazit* festhalten, dass die Bedeutungsebene Text in der übergeordneten Kategorie des Vokalen aufgeht. Das Vokale selbst ist als ein Netzwerk von musikalischen Elementen, Stimm-Körper, personal-biographischen Spuren, sozio-kulturellen Charakter-Figuren, audiotecnischen Prozessen und semantischen Feldern zu abstrahieren – ein Netzwerk, das im Vortrag einzelner Solisten materiell greifbar wird.⁴² Für die Analyse ist hiermit angedeutet, dass es sich im Falle stimmlicher Performanz um eine klangliche Hervorbringung handelt, der das ‚Äußerliche‘ innerlich zu sein scheint. Das gesungene Wort und die Geste können folglich als Variablen eines dynamischen Ausdruckskomplexes verstanden werden, die je nach Aufführung und Aufführungskontext bestimmte Ausprägungen zeitigen.⁴³

Die entscheidende Weichenstellung im analytischen Umgang mit populärer Musik offenbart sich in dem Moment, in dem man das konkrete *Material* vorliegen hat. Zweierlei ist hier zu berücksichtigen: Zum einen existieren poplarmusikalische Formate wie Song und Clip als mediale Produkte – ersterer (vornehmlich) als Tonträger, letzterer als Synthese aus Tonträger

⁴¹ Vgl. Frith (1996, S. 187). Auf das mediale Bezugssystem Technik/Aufnahmetechnik wird in Friths Modell auf der ersten Ebene des musikalischen Instruments Stimme eingegangen. So wird der gekonnte Umgang mit dem Mikrophon als wesentlicher Bestandteil musikalisch-stimmlicher Kunstfertigkeit verstanden (vgl. ebd., S. 188).

⁴² Siehe hierzu auch die Ausführungen Christian Bielefeldts zum Bedeutungsnetzwerk des Phänomens Stimme im populären Kontext (vgl. Bielefeldt 2008).

⁴³ Der genannte Begründungszusammenhang von gesungenem Wort, Stimme, Geste, Aufführung und Kontext geht auf die Ausführungen Paul Zumthors zum Oralitätsbegriff zurück. Hierin setzt Zumthor ferner die Variablen Akteur und Medien zentral. In diesem Sinne beschreibt er einen umfassenden Deutungskomplex, der auch im Hinblick auf die nicht originär musikalischen Materialaspekte populärer Musik seine Gültigkeit behält (vgl. Zumthor 2002, S. 243). Zumthor gibt aber an anderer Stelle zu bedenken, dass es problematisch wäre, jede gestische und mimische Qualität in ein Zeichensystem zu verorten. Allerdings kann die Geste Zeichen in dem Maße sein, als sie kulturell bedingt ist oder in einem bestimmten Milieu eine konventionelle Bedeutung trägt (vgl. Zumthor 1988, S. 712).

und Bewegungsbildspur.⁴⁴ In Entsprechung ihrer jeweiligen technischen Dispositionalität bringen sie die Inhalte, die sie übertragen, auf eine bestimmte Art und Weise hervor.⁴⁵ Zum anderen meint die Begriffskomponente ‚Ton‘, dass verschiedene Tonspuren in Gestalt von musikalisch-klanglichen Sukzessionsverläufen zusammengefasst werden. Was sich angesichts dieser Unterscheidung andeutet, ist die Tatsache, dass die performativen, sprachsemantischen und medialen Deutungshorizonte mit Blick auf den (popular-)musikalischen Vortrag nur von begrenzter Reichweite sind. Besonders augenscheinlich wird dies entlang der Bedeutungsebene des rein Instrumentalen. Diese führt in eine originär musikalische Konstitutionslogik ein, die sich Fragen etwa der Gestenhaftigkeit oder Denotativität nicht grundsätzlich verschließt, aber nur unzureichend durch diese einzufassen ist. An einem einfachen Experiment lässt sich die eigentümliche Sonderstellung von Musik demonstrieren: Stellt man bei der Betrachtung des musikbasierten Audiovisionsformats Clip den Ton stumm, so entsteht auf der Bildebene eine sonderbar anmutende Reihung von visuellen Ereignissen, seien es Handlungs- und Bewegungsabläufe, Gesten, mimische Ausdrücke oder Setting-Arrangements. Zudem wäre es möglich, dass der Betrachter eines solchen Experiments gar nicht anders kann, als eine musikalische Tonspur zu imaginieren. In den theoretischen Diskurs zurückübertragen heißt dies, dass die Musikebene einen übergeordneten Sinnzusammenhang herzustellen scheint. Unterstützt wird diese Annahme dadurch, dass die Musik, in Form des Songs, beim bloßen Hören schon als sinnhaft erfahren wird.⁴⁶ Mit Blick auf das oben erwähnte Ausgangskonstrukt eines dynamischen Aus-

⁴⁴ In Clips mit diegetischem Sound wird zudem eine musikfremde Tonspur integriert.

⁴⁵ Bei Sybille Krämer wird diese Auffassung als kleinster gemeinsamer Nenner der Medialitätsdebatte identifiziert (vgl. Krämer 2004, S. 23). Aufgrund der diesbezüglich vorliegenden Fülle an Positionen kann im Rahmen dieses Beitrags nicht auf einzelne Medialitätskonzepte eingegangen werden. Aufzuzeigen ist lediglich, dass Tonträger und Musikclip auch als Teil eines umfassenden Medialitätsdiskurses zu erfassen sind. Fragen der Medialität sind im Bereich der Musikwissenschaft bislang nur mit Zurückhaltung diskutiert worden sind. So besteht bezüglich des Zusammenhangs von Elektroakustik und musikalischer Ästhetik noch grundsätzlicher Aufklärungsbedarf (vgl. Böhme-Mehner/Mehner/Wolf (Hg.) 2008).

⁴⁶ Insgesamt sei an dieser Stelle eine inhaltliche Abgrenzung gegenüber der Gattung Filmmusik vorgenommen – aus diesem Grund auch die terminologische Festlegung ‚musikbasierte Audiovisionsformate‘. Neben dem Musikclip zählen zu den musikbasierten Formaten im populären Sektor noch die musikalischen Großformen Rock-/Pop-Konzert, Musical und Musikfilm. Der Hauptunterschied zwischen vorgenannten Formaten und Filmmusik besteht darin, dass die Musik im Film der Untermalung und Verstärkung der bildlichen Ebene dient (vgl. Bullerjahn 2001, S. 20 ff.). Gemäß der oben genannten Terminologie würde es sich hier um ein bildbasiertes Audiovisionsformat handeln. Nichtsdestoweniger lässt sich auch an Filmmusik aufzeigen, dass musikalische Klanglichkeit ab einem gewissen Grad inhaltlich-struktureller Konsistenz im Hörer sinnhafte respektive sinnliche Bezüge herzustellen vermag. So wird dies am Markt-Phänomen der sogenannten Original-Soundtracks evident. Diese Tonträger enthalten zumeist nur die jeweilige Musik, die im Film erklingt. Bemerkenswert ist, dass sie in nicht unerheblichem Maße vertrieben und konsumiert werden, also eine ‚sinnlich-sinnhafte‘ Unterhaltungssituation herstellen.

druckskomplexes ist also zu präzisieren, dass die Musik hierin zentral gesetzt werden muss. *Die sich im Song realisierenden musikalischen Strukturen formen den Nukleus des gesamten Ausdruckskomplexes (Bild-)Text/Stimme-Ton. Das bedeutet für die Analyse, dass alle Materialebenen nur unter Bezugnahme zur Musik zu beschreiben und deuten sind.*

Eine solche analytische Prämisse soll nicht dazu führen, dass der Musikwissenschaft im Hinblick auf das interdisziplinäre Forschungsobjekt Musikclip eine exponierte Stellung zuteil wird. Es gilt vielmehr, den analytischen Raum im Spannungsfeld von Musik-, Sprach- und Medienwissenschaft so zu vermessen, dass eine Ordnung von Analysefeldern, Konzepten und Begriffen entsteht, die in finaler Wendung eine gegenstandsangemessene Betrachtung des Musikclips ermöglicht. Insgesamt ist anzuregen, dass die genannten Disziplinen ihre Zusammenarbeit systemisch begreifen, d.h. im Sinne einer *materialbezogenen kulturwissenschaftlichen Reflexion*.⁴⁷ Zuletzt muss es sowohl für Musikwissenschaft als auch für Sprach- und Medienwissenschaft darum gehen, die jeweils im Song oder Clip enthaltenen kulturellen Codes analytisch-interpretativ zu repräsentieren. Der Aspekt kultureller Konventionalisierung ist disziplinübergreifend von erkenntnisleitendem Interesse. *Kulturelle Konventionen* entstehen als Folge des reziproken Beziehungsgeflechts sozialer Praxis und ästhetischer Produktion. Sie bilden den Nährboden, aus dem die Objekte der Populärkultur ihr spezifisches Verständigungs- und Wirksamkeitspotential ziehen.⁴⁸ Alle Disziplinen, die sich mit Gegenständen populärer Kultur beschäftigen, stehen somit vor der Aufgabe, kollektive Bedeutungsfelder in den Verästelungen unterschiedlicher Materialzusammenhänge hervor zu fördern.

4. Materialanalyse aus musikwissenschaftlicher Perspektive: Exkursivität und Rekursivität

Im Folgenden soll ein globaler analytischer Ansatz zu populärer Musik vorgestellt werden – die Betrachtung von Text-Ton-Relationen ist diesem immanent. Eingedenk der im vorangegangenen Abschnitt angedachten inter-

⁴⁷ Jobst P. Fricke entwirft einen systemischen Ansatz aus dem Blickwinkel der Musikwissenschaft heraus. Gefordert wird eine grundlegende Besinnung auf den gemeinsamen Forschungsgegenstand. Um diesen bündeln die beteiligten Disziplinen ihre Kompetenzen. Möglich erscheint dies, da Musik auf dreierlei Weise Gegenstand wissenschaftlicher Betrachtungen sein kann: als notiertes Werk, als akustisches Phänomen und als erlebter Klang. Daraus folgt, dass Musik als menschliches Phänomen nur von den Eigenschaften, Möglichkeiten und Zielen des Menschen aus zu verstehen ist. Systemisch zu arbeiten bedeutet danach, dieses komplexe Gefüge von Wirkungsgrößen in seinem dynamischen Verhalten zu betrachten (vgl. Fricke 2003).

⁴⁸ Nach Hans Otto Hügel entfalten die Objekte der Populärkultur ihren spezifischen Reiz dadurch, dass sie fortwährend zwischen den Registern des Sozialen und Ästhetischen changieren (vgl. Hügel 2003).

disziplinären Zusammenarbeit ließe sich die Ausführung dieses Ansatzes als primär musikwissenschaftliches Unterfangen begreifen, da als Analyse-material zunächst der Tonträger vorliegt. Nichtsdestoweniger ist bereits auf der musikanalytischen Ebene eine Vermittlung zwischen performativen, medialen und sprachsemantischen Deutungshorizonten vorgesehen. Der Gedanke der Vermittlung kann indes nicht darüber hinwegtäuschen, dass die andauernden Bedeutungsverschiebungen im Spannungsfeld von Sozialität und Ästhetizität eine Herausforderung für die Analyse darstellen. Abhilfe schafft diesbezüglich der Umweg über die Abstrahierung des binären Codes sozial/ästhetisch auf der Folie des Objektes selbst. So ist für das theoretische Verständnis von Popsongs die der Luhmann'schen Systemtheorie entlehnte Vorstellung einer Zwei-Seiten-Form hilfreich.⁴⁹ Deren zwei Seiten – denkbar als *Innen- und Außenseite* – stehen in keinem eindeutigen Zuordnungsverhältnis. Während die Innenseite im Medium des Klangs durch eine bestimmte Kopplung an Elementen konstituiert wird, wird die Außenseite durch die Möglichkeiten, die das Medium des sozialen Sinns bietet, generiert. Ein Phänomen wie Stimme fungiert in diesem Relationsgefüge als bedeutendes Verbindungsstück, da es Melos und sozialen Verbalisationsgestus aufeinander bezieht. Sie ist qua ihrer spezifischen Position prädestiniert, ein Übermaß an Bedeutungszusammenhängen herzustellen.

Das bedeutet nunmehr, dass die Forschenden angeleitet sind, materielle Zusammenhänge so offenzulegen, dass ein grundlegendes Verständnis für das Verhältnis von musikalischen und nicht-musikalischen Bedeutungsfeldern möglich wird. Es gilt hierbei zu berücksichtigen, dass das jeweilige Klangmaterial durch äußere Faktoren wie Medienwahl, Aufführungskontext, Genrezugehörigkeit und Oeuvrebezug determiniert wird. Im Zusammenschluss fungieren sie als eine Art Produktionsgedächtnis, welches daran erinnert, dass die vielfältigen Materialsynthesen in den medialen Settings einem ästhetischen Ereignishorizont entspringen, der mit einem bestimmten Künstler oder Genre in Verbindung gebracht wird.⁵⁰ An ebenjenen Faktoren können alsdann materialspezifische Interpretationslinien vorbereitet werden. Peter Wicke leitet hieraus ein methodologisches Verdikt ab, nämlich dass

die klangstrukturellen Konfigurationen eines Stückes kontextabhängig vorstellbar sind, ohne ihre internen strukturellen Determinanten deshalb an eine Konstellation aus situations- und wahrnehmungsabhängigen Beliebigkeiten zu überschreiben. (Wicke 2003, S. 122)

⁴⁹ Vgl. Luhmann (1995, S. 169).

⁵⁰ Diesbezüglich sei auf Dietrich Helms verwiesen, der insgesamt für einen erweiterten Materialbegriff optiert, der sich vom traditionellen Verständnis der Werkimmanenz emanzipiert (vgl. Helms 2002).

Die Bedeutung des *Kontextbegriffs* für die Analyse kann nicht unterschätzt werden. An ihm manifestiert sich die doppelseitige Fragerichtung des Nach-Innen und Nach-Außen.⁵¹ Die Folgen für die Analyse sind derart, dass das konkrete Klangmaterial entlang verschiedener Analysefelder enggeführt werden muss. Dabei sind zwei Betrachtungsebenen zu beachten: eine *exkursive*, die vom reinen Materialbezug ‚abschweift‘ und soziale, performative und mediale Deutungshorizonte auf sich vereint und eine *rekursive*, die sich klangimmanent kapriziert. Mit dem Begriff der Exkursivität sind die Analysefelder Genre⁵², Performance/Performanz⁵³ und Medium⁵⁴ impliziert.

⁵¹ Die Bedeutung des Kontextbegriffs lässt sich beispielsweise am Phänomen der Coverversion demonstrieren. In einem Coversong werden musikalische Strukturen in andere zeitliche und ästhetische Kontexte übertragen. Fragen wie etwa jene der Wirksamkeit der Musik oder der interpretativen Gestaltung können nur hinreichend entlang eines entsprechenden Kontextverständnisses geklärt werden.

⁵² Nach einer Definition von Franco Fabbri besagt Genre, dass musikalische Ereignisse (reale oder mögliche) sich zu Netzwerken verbinden, deren Aneignung durch sozial verbindliche Regeln definiert wird (vgl. Fabbri 1982, S. 52). Dem ist die Auffassung von Genre als kommunikatives Konzept anzufügen, welches impliziert, dass die Reziprozität von Interpret und Publikum stets von neuem exklusive Beziehungszusammenhänge generiert (vgl. Hamm 2000, S. 298). Was an beiden Definitionen zu Tage tritt, ist die Tatsache, dass die Zugehörigkeitskontexte musikalischer Zeichen als transzendierend und sich-stets-dynamisierend gelesen werden müssen. Angesichts einer solchen Unübersichtlichkeit musikalisch-sozialer Verständigung sei für die konkrete Analyse empfohlen, den Genrebezug in Form von ‚Genre-Indikationen‘ zu identifizieren. Hierdurch wird das ‚Transitorische‘ von Genre allgemein betont. So ist für die Materialanalyse die Einsicht bedeutsam, dass hier im Sinne qualitativer Methodik vorgegangen wird, d.h. sozial-kommunikative Bezüge können nur ‚stichprobenartig‘ innerhalb eines zeitlich und sachlich gerahmten Kontextes bestimmt werden. Darüber hinaus soll angeregt werden, Oeuvrebezüge und sprachsemantische Bedeutungsfelder als Teil des Genrezusammenhangs aufzufassen. Dies kann in Form einer Bestimmung von Topoi, musikalisch und sprachlich, geschehen, mit denen der/die Künstler auf eine bestimmte Art und Weise umgeht bzw. umgehen.

⁵³ Im deutschsprachigen Diskurs zur populären Musik lassen sich die Begriffe Performance und Performanz ergänzend gebrauchen. Ersterer hat sich im deutschen Sprachraum allgemein im Hinblick auf die Beschreibung des (künstlerischen) Akts der Darstellung, den Moment der Aufführung und die Präsenz von Akteuren durchgesetzt (vgl. Schumacher 2002, S. 384). Dieser umfassende Aspekt des Ereignischarakters von Musik ist für den populärmusikalischen Diskurs zentral, da es sich bei populärer Musik nicht selten um eine musikalische Interaktionshandlung handelt, deren Vollzug gemäß der Möglichkeiten des jeweiligen medialen Settings inszeniert wird. Fokussiert eine Forschungsfrage diesen Ereignischarakter des Zusammenspiels von Akteuren (Live oder im Tonstudio), so scheint der Gebrauch des Performance-Begriffs naheliegend. Hingegen werden die Vorzüge von Performanz in dem Moment sichtbar, da man sich der Beschreibung einzelner Materialebenen zuwendet – das englische ‚Performance‘ wäre hier aufgrund seines integrierenden semantischen Gehalts irreführend. Im Ganzen muss es in der Analyse darum gehen, musikalisch-klangliche Hervorbringung auf performative Bezüge hin zu untersuchen. Dieses Unterfangen behielte auch im Falle der Feststellung der Negation performativer Bezüge seine Legitimation – eine erste Aussage zum Material wäre sodann hergestellt.

⁵⁴ Es wird hier von einem weitgefassten Medienbegriff ausgegangen, der die Vielfalt des gesellschaftlichen Mediengebrauchs widerspiegelt. In ihm werden Gattungs- und Distributionsaspekte integrierend behandelt. Diese Auslegung erscheint erforderlich, da es insbesondere im Fall von populärer Musik deutlich zu machen gilt, dass der Medienbegriff unterschiedliche

Jene Felder werden explizit als *musikanalytisch* relevant gesetzt. An ihnen erfolgt eine kontextuelle Verortung des Materials, durch die der Schärfeegrad der Beobachtung erhöht wird. Der Kontextbegriff versucht der eigentümlichen Dynamik Herr zu werden, dass sich entlang musikalischer Handlungen, hervorgebracht in medialen Settings, zeitlich bedingte Verständigungsverhältnisse zwischen Publikum und Interpret konstituieren. Die Deskription der kontextuellen Ausgangslage erscheint angesichts der mannigfachen Materialtransformationen, die sich aufgrund ständiger Medienwechsel in der populären Musikpraxis vollziehen, unentbehrlich. Wenn beispielsweise ein Song in Form einer Aufnahme betrachtet wird, werden Fragen der Genrebezogenheit und Performativität zunächst im Hinblick auf seine spezifische Medialität als Tonträger zu beantworten sein, es sei denn, die Forschungsfrage intendiert ausdrücklich etwas anderes. Auch können eine Aufnahme der Band X und eine Live-Performance der Band Y nur unter Bedingung einer entsprechend ausformulierten Forschungsfrage beobachtet werden.

Mit der *rekursiven* Betrachtungsebene bewegt sich die Analyse auf der Ebene der Musikimmanenz – dem originären Terrain der Musikwissenschaft. Verschiedene musikalische Dimensionen wie Formverlauf, Stimme, Instrumentierung, Harmonien, Motive, Sound, Rhythmus, und die Bedeutungsebene Text werden hier aufeinander bezogen. In einem Transkriptionsschema gilt es diese vertikal zu schichten. Dieses Vorgehen wird dem Postulat der Ausführlichkeit aber auch jenem der Gegenstandsangemessenheit gerecht. So weisen Begriffe wie die des *Sounds* oder des *Rhythmus*⁵⁵ darauf hin, dass die musikalischen Strukturen im Pop-Kontext nicht im Sinne einer Musik als Text⁵⁶, sondern als klangsinnliche Texturen zu lesen sind.⁵⁷

mediale Settings und Formate impliziert, in denen das klangliche Ereignis nicht nur vermittelt sondern auch in ein bestimmtes Licht gerückt wird, d.h. es wird im jeweiligen Medium von neuem kontextualisiert und infolge re-materialisiert. Die wesentlichen medialen Settings und Formate populärer Musik sind: Tonträger, Konzert, Clip, Film und Internet.

⁵⁵ Als Umsetzungsoption ist anzudenken, die Dimensionen Motive, Sound und Rhythmus zu einer Betrachtungsebene zusammen zu ziehen. Daran soll deutlich werden, dass es bei der rekursiven Analyse nicht darum gehen kann, einen Materialaspekt wie Rhythmus grundsätzlich durchgehend zu transkribieren. Die finale Entscheidung zugunsten der ausführlichen Visualisierung eines bestimmten musikalischen Ereignisses sollte auf die spezifische Beschaffenheit des Materials und die zu erwartenden Aussagen durch die Transkription zurückgeführt werden.

⁵⁶ Dietrich Helms postuliert eine allgemeine Abkehr von der klassischen Musiktheorie. Er begründet dies mit Unbehagen gegenüber dem Musikverständnis, das der Musiktheorie zugrunde liegt. Hiernach stützt sich diese auf das Paradigma einer *musica poetica*, einer werk gewordenen Musik, die in der Anordnung des Materials eine Aussage des Komponisten codiert. In diesem Sinne spiegelt sie nur einen zeit- und kulturbedingten Ausschnitt musikalischer Praxis wieder (vgl. Helms 2003b, S. 200).

⁵⁷ Bei Martin Pfeleiderer wird Sound als Klanggeflecht gedeutet und in ein dreigliedriges hierarchisches Ordnungssystem eingefügt, das zwischen Mikro-, Meso- und Makroebene unterscheidet. Entsprechend dieser Reihenfolge sind es Einzelsounds, Klangfolgen/-gestalten

Darüber hinaus zieht die Forderung nach Rekursivität auch eine Neu-Gewichtung des musikanalytischen Blicks nach sich. Gemeint ist die Fokussierung auf die Relationalität der klanglichen Elemente.⁵⁸ Rekursivität verinnerlicht die Zielvorstellung, dass das Klangmaterial in seinem begrenzten Inventar an Elementen durch sich selbst definiert wird, d.h. jeder beliebige Punkt durch das musikalisch-klangliche Strukturgeflecht selbst verständlich gemacht werden kann. Mit dem Begriff der *Rekursivität* sind einerseits die Grenzen einer solchen Betrachtung mit formuliert – die Komplementierung daher durch die exkursive Betrachtungsebene –, andererseits wird dadurch, dass die Wirkungszusammenhänge der Elemente hervorgefördert werden, der Blick auf ebenjene Aspekte des Klangmaterials geschärft, die für die Bedeutungskonstitution in bestimmten Kontexten eine Rolle spielen können. Dies ist nicht zuletzt für die Identifikation und Interpretation audiovisueller Korrespondenzverhältnisse im Musikclip eine wesentliche Vorstufe.

Zum Abschluss dieses globalen methodologischen Umrisses gilt es die Einsicht zu festigen, dass Musikwissenschaft und alle Disziplinen, die an der Beobachtung populärer Musikformate beteiligt sind, vor der Aufgabe stehen, kollektive Bedeutungsfelder in unterschiedlichen Materialkontexten hervor zu fördern. Es ist davon auszugehen, dass sich zukünftige Materialanalysen als sehr aufschlussreich erweisen werden, wenn es darum geht, die kulturellen Kontexturen innerhalb des sich ausdehnenden Netzwerks populäre Musik näher zu beleuchten. Durch den besprochenen Teilaspekt Text-Ton wird diese Annahme unterstützt. So entwickelt Stimme in der Kontaktzone zur Musik ein Spiel mit Vorstellungen von Körperlichkeit,

und Klangtexturen die den Gesamtkomplex Musik und dessen Rezeption prägen (vgl. Pfeleiderer 2003).

⁵⁸ Eine nähere Erläuterung des Rekursivitätsbegriffs scheint an dieser Stelle erforderlich, da im Folgenden im Sinne methodischer Heuristik argumentiert wird. Hierzu sei auf den Song als die zentrale klanglich-materielle Hervorbringung populärer Musikpraxis und außerdem auf seine Funktion als ein Unterhaltung ermöglichendes Objekt eingegangen – durch die unterhaltungstheoretische Aspektualisierung soll allerdings nicht die Kunstfertigkeit populärer Musikproduktion in Frage gestellt wird. Als Unterhaltungsobjekt schafft der Song eine „Sonderrealität der Unterhaltung“ (Luhmann 2004, S. 100), indem er bestimmte musikalisch-klangliche Ereignisse aufeinander bezieht und miteinander verschmilzt, andere hingegen ausschließt. Als musikalische Kleinform wiederum setzt er einen spezifischen Umgang mit musikalischer Sukzessivität voraus. Für den Hörer bedeutet der Umstand der Standardisierung im Kleinformat, dass er sich grundsätzlich in diesem Format zurechtfindet. Vor dem eigentlichen Hörakt hat der Hörer eine Ahnung davon, in welchem klanglichen Rahmen sich der Song bewegen wird – Irritationen inbegriffen. Die Regelmäßigkeit ist die Voraussetzung dafür, dass der Song als Unterhaltungsobjekt überhaupt seine Wirkung entfalten kann. Durch den Aspekt der Wirksamkeit ist gleichsam aufgezeigt, dass Regelmäßigkeit nicht mit der Herstellung von Einförmigkeit und Wiederholung gleichzusetzen ist. Der Song kann folglich als Schema vorgestellt werden. Ein solches Schema zwingt nicht zu Wiederholungen, es legt Handeln nicht fest, da seine Funktion gerade darin liegt, Spielraum für frei gewähltes Verhalten zu generieren (vgl. Luhmann 2004, S. 193). Die Folge hiervon ist, dass sich im einzelnen Song Ähnlichkeits- bzw. Differenzverhältnisse zu Song-Paradigmen abzeichnen (vgl. Shave 2008, S. 42 ff.).

Persönlichkeit und Charakterlichkeit – ein Mehr an Bedeutungen ist die Folge. Der Ausblick auf das Zusammenspiel von Bild, Text und Ton im Musikclip lässt erahnen, dass hier Bedeutung von neuem konstituiert wird. Ein tiefergehendes Verständnis für die musikalischen Materialzusammenhänge wird helfen, diese Bedeutungsproduktion zu entziffern.

Literatur

- Altrogge, Michael (2000): Tönende Bilder. Interdisziplinäre Studie zu Musik und Bildern in Videoclips und ihre Bedeutung für Jugendliche. Bd. 1–3. Berlin.
- Aufderheide, Pat (1986): Music videos: The look of sound. In: *Journal of Communication* 36, S. 57–78.
- Baacke, Dieter (1998): Neue Ströme der Weltwahrnehmung und kulturelle Ordnung. In: Baacke, Dieter (Hg.): *Handbuch Jugend und Musik*. Opladen, S. 29–57.
- Barth, Michael/Neumann-Braun, Klaus (1996): Augenmusik. Musikprogramme im deutschen Fernsehen – am Beispiel von MTV. In: *Landesanstalt für Kommunikation Baden-Württemberg (LFK) (Hg.): Fernseh- und Radiowelt für Kinder und Jugendliche*. Villingen, S. 249–265.
- Bechdorf, Ute (1996): Music video histories. In: Hackl, Christiane/Prommer, Elizabeth/Scherer, Brigitte (Hg.): *Models und Machos? Frauen- und Männerbilder in den Medien. (= Kommunikation audiovisuell 21)*. Konstanz, S. 277–299.
- Bergermann, Ulrike (2003): Videoclip. In: Hügel (Hg.), S. 478–482.
- Bielefeldt, Christian (2008): Voices of Prince. In: Bielefeldt/Dahmen/Grossmann (Hg.), S. 201–219.
- Bielefeldt, Christian/Dahmen, Udo/Grossmann, Rolf (Hg.) (2008): *PopMusicology. Perspektiven der Popmusikwissenschaft*. Bielefeld.
- Bódy, Veruschka/Weibel, Peter (Hg.) (1987): *Clip, Klapp, Bum. Von der visuellen Musik zum Musikvideo*. Köln.
- Böhme-Mehner, Tatjana/Mehner, Klaus/Wolf, Motje (Hg.) (2008): *Elektroakustische Musik – Technologie, Ästhetik und Theorie als Herausforderung an die Musikwissenschaft*. Essen.
- Bowman, Rob (2003): The determining role of performance in the articulation of meaning: the case of „Try a Little Tenderness“. In: Moore, Allan F. (Hg.): *Analyzing Popular Music*. Cambridge, S. 103–130.
- Bullerjahn, Claudia (2001): *Grundlagen der Wirkung von Filmmusik*. Augsburg.
- Chion, Michel (1994): *Audio-vision: sound on screen*. New York.
- Deutsches Filmmuseum Frankfurt (Hg.) (1993): *Sound & Vision. Musikvideo und Filmkunst*. Frankfurt a.M.
- Fabbri, Franco (1982): A theory of musical genres: two applications. In: Horn, David/Tagg, Philip (Hg.): *Popular music perspectives*. Göteborg/Exeter, S. 52–81.
- Fischer-Lichte, Erika (2004): *Ästhetik des Performativen*. Frankfurt a.M.

- Fix, Ulla (2001): Aspekte der Intertextualität. In: Brinker, Klaus/Burkhardt, Armin/Ungeheuer, Gerold et al. (Hg.): Text- und Gesprächslinguistik/Linguistics of Text and Conversation. (= Handbücher zur Sprach- und Kommunikationswissenschaft/Handbooks of Linguistics and Communication Science 16.2). Berlin/New York, S. 449–457.
- Fricke, Jobst P. (2003): Systemische Musikwissenschaft. In: Niemöller, Klaus Wolfgang/Gätjen, Bram (Hg.): Perspektiven und Methoden einer systemischen Musikwissenschaft. Frankfurt a.M., S. 13–21.
- Frith, Simon (1996): *Performing rites. On the value of popular music.* Cambridge.
- Genette, Gérard (2004): *Palimpseste: die Literatur auf zweiter Stufe.* Frankfurt a.M.
- Großmann, Rolf (1998): Wo beginnt Intermedialität? Latente Prämissen und Dimensionen eines klärungsbedürftigen Konzepts. In: Helbig, Jörg (Hg.): *Intermedialität: Theorie und Praxis eines interdisziplinären Forschungsgebiets.* Berlin, S. 108–119.
- Gumbrecht, Hans Ulrich (2004): *Diesseits der Hermeneutik. Die Produktion von Präsenz.* Frankfurt a.M.
- Hamm, Charles (2000): Genre, performance, and ideology in the early songs of Irving Berlin. In: Middleton (Hg.), S. 297–306.
- Hausheer, Cecilia (1994): Werbende Klangaugen. In: Hausheer/Schönholzer (Hg.), S. 186–197.
- Hausheer, Cecilia/Schönholzer, Annette (Hg.) (1994): *Visueller Sound. Musikvideos zwischen Avantgarde und Populärkultur.* Luzern.
- Hawkins, Stan (2000): Prince: Harmonic analysis of ‚Anna Stesia‘. In: Middleton (Hg.), S. 58–70.
- Helms, Dietrich (2002): Musikwissenschaftliche Analyse populärer Musik? In: Rösing, Helmut/Schneider, Albrecht/Pfleiderer, Martin (Hg.): *Musikwissenschaft und populäre Musik. Versuch einer Bestandsaufnahme.* (= Hamburger Jahrbuch für Musikwissenschaft 19). Frankfurt a.M., S. 91–103.
- Helms, Dietrich (2003a): In Bed with Madonna. Gedanken zur Analyse von Videoclips aus medientheoretischer Sicht. In: Helms/Phleps (Hg.), S. 99–117.
- Helms, Dietrich (2003b): Vom System Ton zum System Sound. In: Phleps/Appen (Hg.), S. 197–228.
- Helms, Dietrich/Phleps, Thomas (Hg.) (2003): *Clipped Differences. Geschlechterrepräsentationen im Musikvideo.* (= Beiträge zur Populärmusikforschung 31). Bielefeld.
- Hügel, Hans-Otto (2003): Einführung. In: Hügel (Hg.), S. 1–22.
- Hügel, Hans-Otto (Hg.) (2003): *Handbuch Populäre Kultur. Begriffe, Theorien und Diskussionen.* Stuttgart.
- Jacke, Christoph (2003): Kontextuelle Kontingenz. Musikclips im wissenschaftlichen Umgang. In: Helms/Phleps (Hg.), S. 27–40.
- Keazor, Henry/Wübbena, Thorsten (2005): *Video thrills the radio star. Musikvideos: Geschichte, Themen, Analysen.* Bielefeld.
- Kerscher, Gottfried/Richard, Birgit (2003): *MoVi und MuVi. Zur Interpretation bewegter Bilder in Film und Musikvideoclip als Bildwissenschaft und ‚kritische Stil-*

- analyse'. In: Ehrenspeck, Yvonne/Schäffer, Burkhard (Hg.): Film- und Fotoanalyse in den Erziehungswissenschaften: ein Handbuch. Opladen, S. 203–225.
- Kittler, Friedrich A. (1999): Gramophone, film, typewriter. Stanford, CA.
- Klein, Richard (2008): Das Narrative der Stimme Bob Dylans. In: Bielefeldt/Dahmen/Grossmann (Hg.), S. 220–240.
- Krämer, Sybille (2006): Die „Rehabilitierung der Stimme“. Über die Oralität hinaus. In: Kolesch, Doris/Krämer, Sybille (Hg.): Stimme. Annäherung an ein Phänomen. Frankfurt a.M., S. 269–295.
- Krämer, Sybille (2004): Was haben „Performativität“ und „Medialität“ miteinander zu tun? In: Krämer, Sybille (Hg.): Performativität und Medialität. München, S. 11–32.
- Luhmann, Niklas (1995): Die Kunst der Gesellschaft. Frankfurt a.M.
- Luhmann, Niklas (2004): Die Realität der Massenmedien. 3. Aufl, Wiesbaden.
- Lüders, Christian (1991): Deutungsmusteranalyse. Annäherungen an ein risikoreiches Konzept. In: Garz, Detlef/Kraimer, Klaus (Hg.): Qualitativ-empirische Sozialforschung. Opladen, S. 377–408.
- Lüders, Christian/Meuser, Michael (1997): Deutungsmusteranalyse. In: Hitzler, Ronald/Honer, Anne (Hg.): Sozialwissenschaftliche Hermeneutik: eine Einführung. Opladen, S. 57–80.
- Lull, James (1987): Popular music and society. London.
- Leschke, Rainer (2007): Einführung in die Medientheorie. München.
- Meuser, Michael/Sackmann, Reinhold (Hg.) (1991): Analyse sozialer Deutungsmuster. Paffenweiler.
- Middleton, Richard (Hg.) (2000): Reading pop. Approaches to textual analysis in popular music. Oxford.
- Middleton, Richard (2001): Musikalische Dimensionen. Genres, Stile, Aufführungspraktiken. In: Wicke, Peter (Hg.): Rock- und Popmusik. Handbuch der Musik im 20. Jahrhundert. Bd. 8. Laaber, S. 63–106.
- Neumann-Braun, Klaus (Hg.) (1999): Viva MTV! Popmusik im Fernsehen. Frankfurt a.M.
- Neumann-Braun, Klaus/Schmidt, Axel (1999): McMusic. Einführung. In: Neumann-Braun (Hg.), S. 7–42.
- Oevermann, Ulrich (2001a): Zur Analyse der Struktur von sozialen Deutungsmustern. In: Sozialer Sinn 1 S. 3–34.
- Oevermann, Ulrich (2001b): Die Struktur sozialer Deutungsmuster – Versuch einer Aktualisierung. In: Sozialer Sinn 1, S. 35–82.
- Paech, Joachim (1994): Bilder-Rhythmus. In: Hausheer/Schönholzer (Hg.), S. 46–63.
- Pfleiderer, Martin (2003): Sound. Anmerkungen zu einem populären Begriff. In: Phleps/Appen (Hg.), S. 19–29.
- Phleps, Thomas/Appen, Ralf von (Hg.) (2003): Pop Sounds. Klangtexturen in der Pop- und Rockmusik. (= Texte zur populären Musik 1). Bielefeld.
- Rajewsky, Irina O. (2002): Intermedialität. Tübingen u.a.

- Reichertz, Jo (1992): Der Morgen danach. Hermeneutische Auslegung einer Werbe-fotographie in zwölf Einstellungen. In: Hartmann, Hans A./Haußl, Rolf (Hg.): Bil-derflut und Sprachmagie. Fallstudien zur Kultur der Werbung. Opladen, S. 141–164.
- Reichertz, Jo (2005): Wissenssoziologische Verfahren der Bildinterpretation. In: Mi-kos, Lothar/Wegener, Claudia (Hg.): Qualitative Medienforschung. Ein Handbuch. Konstanz, S. 141–152.
- Rösing, Helmut (2003): Bilderwelt der Klänge – Klangwelt der Bilder. Beobachtungen zur Konvergenz der Sinne. In: Helms/Phleps (Hg.), S. 9–26.
- Schmidbauer, Michael/Löhr, Paul (1996): Das Programm für Jugendliche: Musikvideos in MTV Europe und VIVA. In: *Television* 9, S. 6–32.
- Schmidt, Axel/Neumann-Braun, Klaus/Autenrieth, Ulla (2009): Viva MTV! reloaded. Musikfernsehen und Videoclips crossmedial. Baden-Baden.
- Schröter, Jens (1998): Intermedialität. In: *Montage a/v* 7, 2, S. 129–154.
- Schumacher, Eckhard (2002): Performativität und Performance. In: Wirth, Uwe (Hg.): Performanz. Zwischen Sprachphilosophie und Kulturwissenschaft. Frankfurt a.M., S. 383–402.
- Shave, Thomas (2008): Communicative contract analysis: an approach to popular music analysis. In: *Organised Sound* 13, S. 41–50.
- Spielmann, Yvonne (1998): Intermedialität: das System Peter Greenaway. München.
- Vernallis, Carol (2002): The functions of lyrics in music video. In: *Journal of Popular Music Studies* 14, S. 11–31.
- Weibel, Peter (1987): Was ist ein Videoclip? In: Bódy/Weibel (Hg.), S. 274–275.
- Wicke, Peter (1992): Jazz, Rock und Popmusik. In: Stockmann, Doris (Hg.): Volks- und Populärmusik in Europa. (= Neues Handbuch der Musikwissenschaft 12). Laaber, S. 445–477.
- Wicke, Peter (1993): Vom Umgang mit Popmusik. Berlin.
- Wicke, Peter/Ziegenrucker, Kai-Erik/Ziegenrucker, Wieland (2001): Handbuch der populären Musik. 4. Aufl. Mainz.
- Wicke, Peter (2003): Popmusik in der Analyse. In: Vendrix, Philippe (Hg.): *Acta Musicologica*. Bd. 75. Basel, S. 107–126.
- Wulff, Hans-Joachim (1999): The Cult of Personality – authentisch simulierte Rock-videos. In: Neumann-Braun (Hg.), S. 262–278.
- Zumthor, Paul (1988): Körper und Performanz. In: Gumbrecht, Hans Ulrich/Pfeiffer, Karl Ludwig/Elsner, Monika (Hg.): *Materialität der Kommunikation*. Frankfurt a.M., S. 703–714.
- Zumthor, Paul (2002): Mündlichkeit/Oralität. In: Barck, Karl-Heinz/Frontius, Martin/Schlenstedt, Dieter/Steinwachs, Burkhard/Wolfzettel, Friedrich (Hg.): *Ästhetische Grundbegriffe. Historisches Wörterbuch in sieben Bänden*. Bd. 4: Medien – Populär. Stuttgart, S. 234–256.