

„The evil of the thriller“:

Multimodalität und Musikvideosemiotik am Beispiel von Michael Jacksons
THRILLER

Stephan Brüssel

1. Einführung: Zum Vorhaben einer multimodalanalytischen Semiotik des Musikvideos

Der Musikvideoclip stellt besondere Anforderungen an eine semiotisch orientierte Analyse, vor allem deshalb, da ihn die Kombination mehrerer Zeichensysteme konstituiert, deren Einzelcharakteristik, Hierarchisierung und Zusammenspiel für die Mediatisierung von Inhalten entscheidend sind. Sollte sich aber der Ansatz einer Musikvideosemiotik als nützlich erweisen und den Gegenstandsbereich in sinnvoller Weise beschreib- und greifbar machen – und das darf inzwischen als konsensfähige Grundannahme gelten –, dann würde die vornehmliche Herausforderung in Weiterführung bereits bestehender Forschungsansätze darin bestehen, einen genaueren Blick auf die multimodale Beschaffenheit dieses Formats zu werfen. Denn das Vorhandensein von Multimodalität im Musikvideo ist offenkundig.

Ursprünglich ist das Phänomen als Gegenstand der Linguistik bekannt, inzwischen aber über die Untersuchung natürlicher Sprachen hinaus auch in diversen anderen Zusammenhängen von Zeichenpraktiken in Medienkontexten erkannt und erforscht worden. So hat Stöckl zunächst ganz grundsätzlich den semiotischen Grundansatz bei der theoretischen Fassung bestärkt, von anderer Seite aus haben Bateman und Sachs-Hombach für eine medientheoretische Konzipierung plädiert und Bateman und Schmidt eine multimodale Filmanalyse entwickelt.¹ Die Frage an dieser Stelle wäre nun, wie sich dies auf das Musikvideo und ent-

¹ Vgl. Hartmut Stöckl, „Multimodalität – Semiotische und textlinguistische Grundlagen“. In: Nina-Maria Klug/Hartmut Stöckl (Hgg.), *Handbuch Sprache im multimodalen Kontext*. Berlin/Boston 2016, S. 3-35; Nina-Maria Klug/Hartmut Stöckl, „Sprache im multimodalen Kontext“. In: Eckehard Felder/Andreas Gardt (Hgg.), *Handbuch Sprache und Wissen*. Berlin/Boston 2015, S. 242-264; John Bateman/Klaus Sachs-Hombach, „Multimodalität im Schnittbereich von Medientheorie und Semiotik“. In: *Zeitschrift für Semiotik* 41/1-2 (2019), S. 11-36; John Bateman/Karl-Heinrich Schmidt, *Multimodal Film Analysis. How Films Mean*. London 2012.

sprechende Herausforderungen einer medienkulturwissenschaftlichen Arbeit niederschlägt.

Umso dringlicher stellt sich diese Frage im Fall von Michael Jacksons im Jahr 1983 erstmals auf MTV ausgestrahlten „paradigmatic monster video“² THRILLER (USA), das im Kontext einer „palimpsestic nature of the modern music video“³ beschrieben und als „hybrid form“⁴ diskutiert wird: durch seine Nähe zum Film, den Umgang mit ‚Wirklichkeit‘, die Verschmelzung von Narration und Performance. Das Video ist aus verschiedenen Perspektiven mit unterschiedlichen Schwerpunktsetzungen untersucht worden, etwa im Hinblick auf das Problemgefüge ‚Männlichkeit/Sexualität/Monströsität‘⁵ oder als Zeitdokument mit Bezug auf sozialpolitische Problemkomplexe,⁶ und es gilt inzwischen als fester Bestandteil des popkulturellen Gedächtnisses.⁷ Der hier anklingende repräsentative Charakter von THRILLER darf aber auch aus systematischer Sicht angenommen werden, und zwar, was den Status des Videos als Vertreter der Textsorte ‚Musikvideo‘ anbelangt, eine Perspektive, die sich in Anbetracht der aufwendigen Machart ja geradezu aufdrängt.⁸

Aufgabe und Ziel dieses Beitrags soll sein, die Semiotik des Musikvideoclips unter besonderer Berücksichtigung des Teilphänomens der Multimodalität auszubauen. Denn zwar ist unmittelbar einleuchtend und offensichtlich, dass die Textsorte multimodal verfährt, wie genau aber der semiotische Stellenwert zum einen theoretisch fassbar, zum anderen analytisch auswertbar ist, muss derzeit noch als offener Punkt gelten. Im Zuge der Aufarbeitung einiger Fundierungen der Textsorte wie auch des Phänomenbereichs wird sich zeigen – und dieser These möchte ich im Folgenden nachgehen –, dass Multimodalität mit Medialität verschränkt ist, sich in Form von (interagierender) Semioseoperatoren manifestiert und die Präsentationsform maßgeblich determiniert. Damit liefert sie, so die Annahme, im Musikvideo einen entscheidenden Zugang zur Bedeutungskonstitution und müsste in analytisch-interpretatorischer Perspektive sehr viel stärkere Berücksichtigung finden als bislang geschehen, da mit ihr auf mikro- wie auf

² Drago Momcilovic, „Music Video Gothic: Fragmentary Form at the Dawn of MTV“. In: *Gothic Studies* 23/2 (2021), S. 148-162; hier S. 149.

³ Andrew Goodwin, *Dancing in the Distraction Factory: Music Television and Popular Culture*. Minneapolis 1992, S. 78.

⁴ Momcilovic, „Music Video Gothic“, S. 158.

⁵ Vgl. Kobena Mercer, „Monstrous Metaphors: Notes on Michael Jackson’s *Thriller*“. In: Dies., *Welcome to the Jungle: New Positions in Black Cultural Studies*. New York/London 1994, S. 33-52.

⁶ Vgl. Peter Childs, „Pop-Video: Michael Jackson’s ‚Thriller‘ and ‚Race‘“. In: Ders., *Contemporary Culture Texts and Critical Approaches*. Edinburgh 2006, S. 40-48.

⁷ Vgl. Christoph Jacke, „Who Cares about Music in Music Videos? Toward a Multiperspectival Pop Cultural Study of Music Videos“. In: Henry Keazor/Thorsten Wübbena (Hgg.), *Rewind, Play, Fast Forward. The Past, Present and Future of the Music Video*. Bielefeld 2010, S. 179-194; hier S. 181, u. Harmony Bench, „Monstrous Belonging: Performing ‚Thriller‘ after 9/11“. In: Melissa Blanco Borelli (Hg.), *The Oxford Handbook of Dance and the Popular Screen*. Oxford 2013, S. 393-411; hier S. 394.

⁸ Dies legen bspw. die Ausführungen von Christopher Lynch nahe (vgl. ders., „Ritual Transformation through Michael Jackson’s Music Video“. In: *Journal of Communication Inquiry* 25/2 (April 2001), S. 114-131; insb. S. 115-118).

makrostruktureller Ebene eines gegebenen Medientextes feingliedrigere Prozesse der Semiotisierung beschrieben werden können.

2. Musikvideoclip: Medienformat, Forschungszugänge, Grundprämissen

Wurde das Musikvideo zunächst von ideologie- und kulturkritischen Positionen aus auf einen Blick auf seine kommerzielle Nutzung hin heruntergebrochen und als kulturell bedeutungslos diffamiert, ist es inzwischen als wesentlicher Bestandteil der massenmedialen Kommunikation, der (pop-)kulturellen Selbstdarstellung und -verständigung nicht nur anerkannt,⁹ sondern (Stand heute) zudem recht gut erforscht – obwohl der Gegenstandsbereich zugleich aufgrund der Produktionsvielfalt kaum zu überblicken ist.

Definieren ließe sich das Format als filmische, d.h. audiovisuelle Textsorte, deren auffälligste Eigenschaft in einem besonderen Hierarchieverhältnis seiner multimedialen/-modalen Beschaffenheit besteht: Gegenüber anderen Formaten wie insbesondere dem Spielfilm dominiert die Tonebene, der Videoclip inkorporiert die „[a]udiovisuelle Aufzeichnung eines Popsongs“,¹⁰ bei dem auf intermediale und semiotisch komplexe Weise Kohärenz hergestellt und Sinn evoziert wird, konventionelle Muster und Verfahren aus Film und Fernsehen einerseits Eingang finden, sie andererseits innovativ umspielt werden. Das Musikvideo zeichnet sich ferner durch einen Doppelcharakter aus, insofern es als Werbeträger für Song und Interpret*innen dient, darüber hinaus jedoch – etwa seit Mitte der 1970er-Jahre und Queens *BOHEMIAN RHAPSODY* (UK 1975) – zunehmend auch ästhetischen Charakter, eine Clip-Ästhetik¹¹ ausgebildet hat, die mit der Loslösung von der bloß abgefilmten Performance einhergeht. Vor diesem Hintergrund verwundert es nicht, dass prägende Signa und thematische Kulminationspunkte das Starimage, Medien- und Selbstreflexivität sowie Musik darstellen.

Erste Zugänge, z.B. zu spezifischen Vertextungsstrategien,¹² Rezeption¹³ und Ästhetik¹⁴, ergaben sich in der Kommunikations- und Medienwissenschaft gleich nach Sendeantritt von MTV in den 1980er-Jahren; weitere, insbesondere theore-

⁹ Vgl. Simon Rehbach, *Medienreflexion im Musikvideo. Das Fernsehen als Gegenstand intermedialer Beobachtung*. Bielefeld 2018, S. 61.

¹⁰ Jan-Oliver Decker/Hans Krahl, „Videoclip“. In: *Reallexikon der deutschen Literaturwissenschaft*. Bd. 3. Berlin/New York 2007, S. 781-782; hier S. 781.

¹¹ Vgl. Friedemann Malsch/Dagmar Streckel, *Künstler-Videos. Entwicklung und Bedeutung*. Ostfildern-Ruit 1996.

¹² Vgl. Hans-Jürgen Wulff, „Die Ordnungen der Bilderflut. Konstellationen medialer Kommunikation als strukturbildendes Prinzip in Performance-Videos“. In: *Rundfunk und Fernsehen* 37/4 (1989), S. 435-446. <http://www.derwulff.de/files/2-21.pdf>; Abruf am 19.08.2022.

¹³ Vgl. Marsha Kinder, „Music Video and the Spectator“. In: *Film Quarterly* (Herbst 1984), S. 2-15, und Dolf Zillmann/Norbert Mundorf, „Image Effects in the Appreciation of Rock Video“. In: *Communication Research* 14/3 (1987), S. 316-334.

¹⁴ Vgl. Veruschka Bódy/Peter Weibel (Hgg.), *Clip Clap Bum. Von der visuellen Musik zum Musikvideo*. Köln 1987.

tische Impulse gewann die Forschung in der Folge durch die Cultural Studies¹⁵ und die Narratologie¹⁶. Auffällig in neuesten Beiträgen, wie beispielsweise dem von Simon Rehbach, ist eine doppelläufige Blickrichtung auf die medienreflexive und die historische Dimension des inzwischen immens breitenwirksam produzierten und distribuierten Formats.¹⁷

Einen gewichtigen Beitrag – nach wegweisenden Vorarbeiten etwa von Winfried Nöth¹⁸ in Kassel und Hans-Jürgen Wulff¹⁹ in Kiel – hat zuletzt ebenfalls die Passauer Mediensemiotik geliefert, deren Arbeiten hier als Ausgangspunkt dienen und daher in aller Kürze kartografiert werden sollen. In seinem einschlägigen Forschungsbeitrag macht Jan-Oliver Decker – angelehnt an Michael Titzmann²⁰ – einen weiten Textbegriff geltend und fasst das Musikvideo als eigenständiges Medienprodukt und audiovisuellen Text auf, dessen oberflächenstrukturelle Verfasstheit paradigmatische Ordnungen (d.h. eine semantische Tiefenstruktur) codiert. Methodisch bewegt sich der Ansatz im Rahmen von Propositions- und Merkmalsanalysen mit dem Ziel, diese tiefenstrukturellen Ordnungen zu rekonstruieren. Genauer geht es um die Analyse der spezifischen Mechanismen der Bedeutungsproduktion, die Decker im gegebenen Fall als „kombinierte[s] Prinzip“²¹ durch drei, jeweils aufeinander referenzialisierende semiotische Systeme Musik, Sprache und Film ansieht, wobei Bedeutung einerseits durch verschiedene Träger auf der Oberfläche codiert wird, andererseits (und parallel dazu) ein Bedeutungsträger an der Oberfläche ebenso verschiedene Bedeutungen in der Tiefenstruktur aufweisen kann. In der Regel nimmt der Pop-song den hierarchisch höchsten Platz beim Bedeutungsaufbau ein,²² darin wiederum erscheint die Sprache als dominantes System, auf das der musikalische Kotext ausgerichtet ist. Decker weist jedoch zugleich auch auf den Umstand hin, dass neben Videos mit dominanten auditiven Kotexten ebenfalls solche Spielarten vorliegen, in denen „die musikalisch-sprachlichen Kotexte textintern nur den Status eines Kontextes für die privilegierte Bedeutungsproduktion in den Filmbildern innehaben“.²³ Kohärenzbildung, Signifikationsprozesse und damit die Herstellung von Sinnhaftigkeit verlaufen ihm zufolge illustrativ, strukturell, explikativ

¹⁵ Vgl. Simon Frith u.a. (Hg.), *Sound and Vision. The Music Video Reader*. London/New York 1993.

¹⁶ Vgl. die Beiträge von Andreas Blödorn und Oliver Krämer in: Susanne Kaul/Jean-Pierre Palmier/Timo Skrandies (Hgg.), *Erzählen im Film. Unzuverlässigkeit, Audiovisualität, Musik*. Bielefeld 2009, erschienen in der Reihe Medienkulturanalyse beim Transcript-Verlag, herausgegeben von Reinhold Göring.

¹⁷ Vgl. Rehbach, „Medienreflexion im Musikvideo“, S. 9.

¹⁸ Vgl. Winfried Nöth (Hg.), *Semiotics of the Media. State of the Art, Projects, and Perspectives*. Berlin/New York 1997.

¹⁹ Vgl. Wulff, „Die Ordnungen der Bilderflut“.

²⁰ Vgl. grundlegend Michael Titzmann, *Strukturelle Textanalyse. Theorie und Praxis der Interpretation*. 3. Aufl. München 1993 sowie ders. „Semiotische Aspekte der Literaturwissenschaft: Literatursemiotik“. In: Roland Posner/Klaus Robering/Thomas A. Sebeok (Hgg.), *Semiotik. Ein Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur*. Berlin 2004, S. 3028-3103.

²¹ Vgl. Jan-Oliver Decker, *Madonna: Where's that girl? Starimage und Erotik im medialen Raum*. Kiel 2005, S. 28.

²² Vgl. ebd., S. 44.

²³ Ebd., S. 71.

und konnotativ, wobei er an Madonnas Videos – in diesem Fall in erster Linie an LUCKY STAR (USA 1984) – aufzeigt, dass derartige Sinnbildungsformen in der Regel im Verbund auftreten. Letztlich zentral erscheint bei alledem das, was Decker im Zusammenhang mit der konnotativen Gesamtbedeutung ‚Bedeutungsknoten‘ nennt:

So, wie auf elementarer Ebene in solchen Textstellen illustrative, explikative und strukturelle Verfahren der Kohärenzbildungen Bedeutung aufbauen, indem musikalisch-sprachliche Ebene und Filmbilder sich zusammen zu übergeordneten Zeichen ergänzen, genauso ergänzen sich auch diese Bedeutungsknoten innerhalb des Gesamtextes zu einem textintern geknüpften semantischen Netz.²⁴

Und dieses Netz garantiere die Verbindung zwischen primären konnotativen Bedeutungen und der „Verarbeitung textexterner Kontexte“:²⁵ eben die „übergeordnete, kohärente, konnotative Gesamtbedeutung eines Musikvideos“.²⁶ Zwei Konstituenten der Bedeutungsproduktion, so arbeitet Decker richtiggehend heraus, seien in diesem Zusammenhang als Meta-Indices an erster Stelle relevant: Erstens die Konzeption der ‚Person‘ und zweitens der medial im Video vorgeführte und der vorgeführte mediale Raum ‚Musikvideo‘. Das gelte exemplarisch für die Videos Madonnas, jedoch auch für die Textsorte insgesamt – jedenfalls seit ihres ‚Ästhetisierungsschubs‘ in den 1980er-Jahren.

Auch für Hans Krahl bilden das Konzept des ‚Künstler-Selbst‘ und Selbstreferenzialität die entscheidenden Signa, wobei die Textsorte konkret aufgrund „genrehärenter Qualitäten erstens eine grundlegende Nähe“ zu selbstreferenziellen Phänomenen aufweist und sie zweitens „einen genuinen Beitrag bei der Bildung eines ‚Künstler-Selbst‘“ leistet.²⁷ Sie „dient der Selbstdarstellung des Interpreten“²⁸ und ist daher aufgrund ihrer genuin dispositiven Struktur selbstreflexiv grundiert. In dieser Hinsicht habe man es zwar mit einem Genre mit Werbecharakter zu tun, das stets in einem funktionalen Kontext verortet und auf spezifische mit sozialen Praktiken vernetzt sei, gegenüber ‚herkömmlicher‘ Werbung jedoch a) Provokation und Ereigniswert nicht nivelliert werden, sondern konstitutiv seien, b) die appellative Darstellungsfunktion von der emotiven (und poetischen) überlagert werde und c) das Video als Werbetext in besonderer Beziehung zum Beworbenen stehe, nämlich in ikonisch-synekdochischer, da es in Teilen das Beworbene selbst darstellt. Insofern ließe sich sagen, dass das Musikvideo Werbung und ästhetischen Selbstzweck in sich vereint.²⁹

Krahl, das sei nun weiterführend ebenfalls aufgegriffen, fasst das Musikvideo in Einklang mit Decker als Genre mit ausgeprägt *multimodaler* Struktur in Form

²⁴ Ebd., S. 144.

²⁵ Ebd.

²⁶ Ebd., S. 144f.

²⁷ Hans Krahl, „Das Künstler-Selbst. Referenz und Image im Musikvideo“. In: *Zeitschrift für Semiotik* 42/1-2 (2020), S. 153-174; hier S. 156.

²⁸ Ebd., S. 157.

²⁹ Vgl. Decker/Krahl, „Videoclip“, S. 781.

einer Kombination der Medien Musik, Sprache, Film auf.³⁰ Allerdings wird hier wie dort weder näher expliziert, was unter ‚Modalität‘ verstanden wird, noch erläutert, welche Konsequenzen der Aspekt für die theoretische Konturierung und methodische Vorgehensweise hat. In allgemein semiotischer Fassung könnte damit beispielsweise auf eine Matrix von Handlungslogiken in narrativen Strukturzusammenhängen abgehoben werden, wie sie Algirdas Greimas in der modalen und aspektuellen Grammatik seiner Textsemiotik entwickelt hat – aktantielle Modalitäten des ‚Tuns‘ und deskriptive Modalitäten des ‚Seins‘ –,³¹ oder mit Leonard Linsky auf die Erweiterung der Referenzsemantik durch die Dimension der Modalität, die unterschiedlich geartete Referenzmodi in gegebenen Texten zu separieren und analysieren erlaubt.³² Beides erscheint zwar nicht gänzlich abwegig, ist aber an dieser Stelle, in Annäherung an das ‚semiotische Potenzial‘ der kombinierten Einzelmedien – für sich genommen und im Zusammenspiel – vermutlich nicht gemeint. Doch was bedeuten Modalität und Multimodalität im Rahmen eines semiotischen Ansatzes wie dem vorliegenden und wie genau müssten sie Berücksichtigung finden und eingebunden werden, um eine Semiotik des musikalischen Videoclips medientheoretisch zu präzisieren? Das wären unmittelbare und dringende Fragen in diesem Kontext.

Der Forschungsauftrag des vorliegenden Bandes besteht darin, einer ‚neuen‘ Medialität nachzugehen, die im Multimodalen begründet liegt. In Übertragung auf den Untersuchungsgegenstand hier und mit Blick auf die Frage nach der Konsolidierung der eigenen medialen Konstitution und Verfahrensweisen im Umgang mit audiovisuellen Zeichensystemen im Musikvideo, scheint mir der Punkt einer „bedeutungs- und effizienzsteigernde[n] Kombination“³³ verschiedener Zeichenmodalitäten entscheidend zu sein. Anders formuliert: Semiotische Modi sind enorm wichtig zur Erfassung der Bedeutungsorganisation im Musikvideoclip – sie sind nicht nur ontologisch gegeben, sondern eben funktional eingebunden.

3. Modalität und Multimodalität: Interdisziplinär und semiotisch. Theoretische Ausgangslage und methodologisch-analytische Operationalisierung

Doch treten wir noch einmal einen Schritt zurück: Sind Musik, Geräusche, Sprache und Bewegtbilder als Einzelmedien zu behandeln, die das Musikvideo im Verbund ausmachen, oder als distinkte Zeichensysteme – oder wiederum als Zeichenmodalitäten? Alle diese Fassungen sind vorgenommen oder zumindest nahegelegt worden. Oder kommt es allein auf die Perspektive an, den (medien-)theoretischen Background, vor dem man agiert, die entsprechende Prämissensetzung? Wollen wir einstweilen davon ausgehen, dass mit der Annahme einer

³⁰ Vgl. Krah, „Das ‚Künstler-Selbst‘“, S. 156.

³¹ Vgl. Algirdas Julien Greimas, *Du sens II*. Paris 1983 (engl. *On Meaning*. Minneapolis 1987).

³² Vgl. Leonard Linsky, *Referring*. London 1967, S. 125f.

³³ Bateman/Sachs-Hombach, „Multimodalität im Schnittbereich von Medientheorie und Semiotik“, S. 16.

multimodalen Struktur des Musikvideos in heuristischer Hinsicht etwas gewonnen wäre. Was müsste vorab geklärt werden, wie äußert sich das Phänomen im gegebenen Fall und wie ließe sich der Befund analytisch operationalisieren? Bei THRILLER stellt sich diese Frage verschärft, haben wir es schließlich mit einem Medienprodukt zu tun, das offensichtlich über den bloßen Status als Musikvideo hinausreicht, das offensichtlich Spielfilmelemente mit Videoclipen kombiniert, komplex erzählt und intermedial referiert und dadurch nicht nur die Starpersona überdeterminiert signifiziert, sondern zudem auch medienhistorische Botschaften transportiert. Aber zunächst einmal: Wie wäre, um nur einmal einen Aspekt des Beispiels zu nennen, der Musikscores vor Songeinsatz modalanalytisch zu unterscheiden und was damit gewonnen? Warum ist eine Differenzierung zwischen Medium, Code und Modalität tatsächlich sinnvoll, wenn wir über Fälle wie Jacksons Hit-Klassiker sprechen?

Bei den Bemühungen der etwa letzten zwanzig Jahre, semiotische Kommunikations- und Medientheorien zu revidieren und neu auszurichten, nimmt das Verhältnis von Multimodalität und Medialität einen zentralen Platz ein.³⁴ „Keine Kommunikation ohne Medialität“³⁵, so ein Credo Werner Hollys einerseits. Andererseits proklamiert Sigrid Norris: „All interactions are multimodal“³⁶, ebenso wie Nina-Maria Klug/Hartmut Stöckl konstatieren: „Menschliche Kommunikation ist grundsätzlich multimodal“³⁷, und: „Multimodalität ist der natürliche Urzustand unserer kommunikativen Ökologie“³⁸. Allgemeiner Konsens scheint zu sein, dass Phänomene wie ‚Medium‘ und ‚Medialität‘, ‚(Zeichen-)Modalität‘, ‚Materialität‘, ‚Kommunikation‘, ‚soziokulturelle Konventionen und Zeichenpraktiken‘ in Theoriebildung und analytischer Operationalisierung als Kategorien bereitgestellt und in ihrem Zusammenspiel beachtet werden sollten.³⁹ Doch die Gesamtlage zeigt auch terminologische, taxonomische und heuristische Probleme und Lücken auf, resultierend aus unsystematischen Begriffsdefinitionen, unzureichenden Objekterfassungen und interdisziplinären Übersetzungen oder deren Hemmungen.⁴⁰ Und Probleme ergeben sich gar innerdisziplinär, zum Beispiel bei Bateman/Sachs-Hombach und ihren anregenden Überlegungen zur ‚Multimodalität im Schnittbereich von Medientheorie und Semiotik‘. Ausgehend von der medienphilosophisch-anthropologischen Annahme, dass der Mensch sich durch

³⁴ Zur Übersicht vgl. Jan Georg Schneider/Hartmut Stöckl (Hgg.), *Medientheorien und Multimodalität. Ein TV-Werbespot – Sieben methodische Beschreibungsansätze*. Köln 2011; zum Verhältnis von (Multi-)Medialität und Multimodalität, das Schneider/Stöckl differenziert und kritisch darstellen, trotzdem zugleich aber auch als „untrennbare Verbindung“ erkennen vgl. ebd. S. 24-28.

³⁵ Werner Holly, „Medien, Kommunikationsformen, Textsortenfamilien“. In: Stephan Habscheid (Hg.), *Textsorten, Handlungsmuster, Oberflächen. Linguistische Typologien der Kommunikation*. Berlin/New York 2011, S. 144-163; hier S. 144.

³⁶ Sigrid Norris, *Analyzing Multimodal Interaction. A Methodological Framework*. New York/London 2004, S. 9.

³⁷ Klug/Stöckl, „Sprache im multimodalen Kontext“, S. 242.

³⁸ Ebd., S. 244.

³⁹ Vgl. auch in größerer Fassung als die bislang genannten Gunther Kress/Theo van Leeuwen, *Multimodal Discourse. The Modes and Media of Contemporary Communication*. London 2001.

⁴⁰ Stöckl, „Multimodalität – Semiotische und textlinguistische Grundlagen“, S. 4, und Klug/Stöckl, „Sprache im multimodalen Kontext“, S. 247 und 261.

komplexe Fähigkeiten auszeichne, er nicht nur in situationsunabhängiger Weise zu kommunizieren imstande sei, sondern zudem auch in vielfältiger Weise kommuniziere, entwerfen sie ein Konzept von Zeichenmodalitäten – verstanden als Konfiguration bestehend aus semiotischer Ressource und Materialitäten, die diese Ressourcen einsetzen – und verbinden dies mit der Kategorie des Mediums – als ebendiese Materialität bzw. als Materialitätslieferant. Für Bateman/Sachs-Hombach formiert ‚Materialität‘ die entscheidende Größe, da Zeichenmodalitäten nicht ohne Materialität existieren, Medien und Genres wiederum nicht ohne Zeichenmodalitäten. Auf dieser Grundidee aufbauend entwickeln sie ein Schichtenmodell, bei dem das Medium differenziert mit Zeichenmodalitäten in Verbindung gesetzt wird und das gleichzeitig technologische und soziokulturelle Aspekte zusammenführt:

In dem hier eingeführten Modell der Zeichenmodalität bildet [...] die feste Einbeziehung der Materialität eine geeignete Brücke zu einer multimodalen Konstruktion von ‚Medium‘ im semiotischen Sinn. [Medien bieten] den unmittelbaren Kontext, in dem Zeichenmodalitäten verwendet werden können: Bei Zeichenmodalitäten handelt es sich also um keine ‚frei schwebenden‘ Instanzen, sondern diese bedürfen eines Mediums, um sich zu materialisieren. Medien gruppieren in dieser Sichtweise folglich Zeichenmodalitäten dynamisch zu soziokulturell und historisch situierten Konfigurationen.⁴¹

Innerhalb dieses theoretischen „Grundgerüst[s]“⁴² konzipieren sie ein dreischichtiges Konzept von ‚Zeichenmodalität‘, verstanden als Komplex bestehend aus 1) dem Material, das zum kommunikativen Ausdruck beansprucht wird, 2) „Systemen von Kontrasten“⁴³ – d.h. ‚Grammatiken‘ – und 3) der Diskurssemantik, die für die „strukturellen Konfigurationen der vorherigen Ebene mit kontextuellen Interpretationen verantwortlich“⁴⁴ ist, zum Beispiel Textsortenmuster oder Gattungsregeln. Das Medium begreifen sie dabei als Container, der Zeichenmodalitäten konfiguriert und auf einer ‚Projektionsfläche‘ (*canvas*) präsentiert. Die Regularitäten eines Mediums, so fügen sie an, werden materiell wie semiotisch trianguliert: „Das Material (‚canvas‘) eines Mediums besteht [...] zu jedem Zeitpunkt der Existenz dieses Mediums aus den Materialitäten, die die beteiligten Zeichenmodalitäten einbringen“.⁴⁵

So einleuchtend dieser Entwurf erscheint, ihm mangelt es an Präzision im Detail. Wenn Zeichenmodalitäten *ausschließlich* materialiter zutage treten, warum dann überhaupt die Differenzierung in Zeichen und ihr Material? Wäre nicht vielmehr der Zeichenbegriff um seine materielle Dimension zu erweitern? Worin genau besteht der Unterschied zwischen *Zeichenmodalität* als „konventionali-

⁴¹ Bateman/Sachs-Hombach, „Multimodalität“, S. 24.

⁴² Ebd., S. 13.

⁴³ Ebd., S. 23.

⁴⁴ Ebd., S. 24.

⁴⁵ Ebd., S. 26.

sierte Form[] der Gestaltung von Materialität für kommunikative Zwecke“⁴⁶ und *Zeichen*, auf die diese Definition doch ebenfalls zutrifft? Von medientheoretischer Warte aus ist die Notwendigkeit differenzierter Kategorien unbestritten, denn fraglos beeinflussen „Medien Prozesse der Bedeutungskonstitution“, somit auch die „materiellen Bedingungen“ und Regularitäten von Medien.⁴⁷ Und Modi in ihrer (bei Bateman/Sachs-Hombach aufgefächerten) Mehrdimensionalität generieren ebenso fraglos in ihrer „funktionalen Verschränkung [...] eine eigenständige Bedeutungsebene“.⁴⁸ Doch die Grundsatzfragen bleiben bestehen – ebenso wie die Frage nach der Operationalisierbarkeit und Heuristik des entwickelten Modells, die über eine deskriptive Erfassung eines gegebenen Gegenstandes hinausreicht.⁴⁹

Möchte man daher die angerissenen Vorarbeiten produktiv machen und den konkreten Nutzen für eine Musikvideosemiotik ermitteln – und das soll hier geschehen –, sollte ein genauerer Blick auf die wesentlichen Objektteilbereiche nicht schaden, zumal die Frage nach tragfähigen Definitionen und Klassifikationen sogar in Fachkreisen bis dato ebenfalls hoch akut ist.⁵⁰ Dies an dieser Stelle allerdings mit gewisser Vorsortierung: Denn ‚(Zeichen-)Modalität‘ und ‚Multimodalität‘ sind in den unterschiedlichsten Bereichen erforscht worden, so vor allem in textlinguistischen – die Forschungslage zeigt sich reichhaltig.⁵¹ Für uns von Interesse können indessen nur diejenigen Beiträge sein, die erstens einen transmedialen Ansatz liefern, also über einen bloß sprachbezogenen Bezug hinausreichen, zweitens insbesondere av-Medien in den Fokus rücken, drittens semiotisch ausgerichtet sind und viertens Validität bewiesen haben, allgemein als konsensfähig gelten und heuristische Ergiebigkeit in Aussicht stellen.

‚Medium‘ in einer an dieser Stelle hinreichenden Fassung – und hiermit übernehme ich das Begriffsverständnis von Bateman/Sachs-Hombach – sei definiert als materielle Trägereinheit zeichenhafter Prozesse; das Medium ist diejenige Größe, in der Zeichen realisiert werden, Semiose stattfindet, und mit dessen Hilfe kommuniziert wird: die Stimme, Schriftzeichen auf Papier, das Telefon, der Film. Das Medium ist ein Container mit Projektionsfläche (*canvas*). Medien in diesem ‚engen‘ Begriffsverständnis wären in Anlehnung an Gunther Kress und Theo van Leeuwen „materielle Ressourcen, die bei der Herstellung semiotischer Produkte und Ereignisse verwendet werden und sowohl die Werkzeuge als auch die verwendeten Materialien einschließen“⁵². Gleichwohl damit zugegebener-

⁴⁶ Ebd., S. 27.

⁴⁷ Ebd., S. 14.

⁴⁸ Ebd., S. 16.

⁴⁹ Vgl. ebd., S. 29f.

⁵⁰ Vgl. Ulrich Schmitz, „Multimodale Texttypologie“. In: Nina-Maria Klug/Hartmut Stöckl (Hgg.), *Handbuch Sprache im multimodalen Kontext*. Berlin/Boston 2016, S. 327-347; hier S. 330.

⁵¹ Vgl. bspw. die Reihe *Routledge Studies in Multimodality*, herausgegeben von Kay L. O’Sullivan oder das *Handbuch Sprache im multimodalen Kontext*, herausgegeben von Nina-Maria Klug und Hartmut Stöckl (2016).

⁵² Bateman/Sachs-Hombach, „Multimodalität“, S. 21f., und Gunther Kress/Theo van Leeuwen, „Multimodality and multimodal research“. In: Eric Margolis/Luc Pauwels (Hgg.), *The Sage Handbook of Visual Research Methods*. Los Angeles 2011, S. 549-569.

maßen andere Begriffsverständnisse ausgeblendet sind, erscheint diese Fassung jedoch insoweit hinreichend, als sie es uns erlaubt, derartige Prozesse wie auch die Mechanismen ihres Zustandekommens in den Blick zu nehmen. Man müsste demnach, was unser Vorhaben anbelangt, differenzieren: Ton, gesprochene und geschriebene Sprache, Musik, Filmbild stellen semiotisch-modale Ausdrucksformen dar, die in ihrer Kombination, ihrem Verbund die spezifische Medialität des Musikvideoclips kennzeichnen, nämlich eine filmische. Die Textsorte basiert somit auf dem Medium Film.

Weiter wäre ein ‚Text‘ – im weiten Begriffsverständnis nach Decker und im Sinne der Literatur- und Mediensemiotik – als konkret vorliegendes und fixiertes, begrenztes, strukturiertes und informationshaltiges Medienprodukt zu begreifen⁵³ und Multimodalität wiederum dann gegeben, wenn „zumindest zwei Zeichenmodalitäten strukturell und funktional integriert werden“.⁵⁴ Folglich ließe sich ein multimodaler Text

als Zusammenschluss mehrerer unterschiedlicher Zeichenmodalitäten zu einem kohäsiven und kohärenten Ganzen definieren, dessen musterhafte Inhalts- und Handlungsstruktur sowie Verwendungsweisen der Modalitäten typisierten Gebrauchssituationen entspringen und bestimmte kommunikative Funktionen haben.⁵⁵

‚Modus‘ bezeichnet dabei „die Art der Präsentation und Wahrnehmung“⁵⁶ bzw. die Art und Weise, wie etwas unter Zuhilfenahme bestimmter Materialitäten (verstanden als materielle Substrate) kommuniziert wird.⁵⁷ Differenziert werden können die Modi der gesprochenen und geschriebenen Sprache, des stehenden Bildes, der bewegten Bilder und Audio.⁵⁸ Filme und damit auch Musikvideos sind demzufolge genuin multimodal – nämlich audiovisuell und multipel-codiert – grundiert, während beispielsweise Briefe, E-Mails und Tagebucheinträge in der Regel monomodal – visuell-schriftsprachlich – konstituiert sind.

Das Konzept der Zeichenmodalität, das ja nun für die Beschreibung multimodaler Texte fundamental ist, begründet sich dementsprechend auf Sinnesfundierung und Wahrnehmungskanal, in denen Zeichen prozessiert werden, sowie auf den Aspekten der Codiertheit und Medialität von Zeichen.⁵⁹ Entscheidend bei der Abgrenzung von Zeichenmodalitäten sei, so Hartmut Stöckl in Anlehnung an Charles Morris, die Zugehörigkeit von Zeichen zu einem Code, sinnfällig auch der

⁵³ Zur Problematik der Textdefinitionen vgl. Michael Klemm, „Ausgangspunkt. Jedem seinem Textbegriff? Textdefinitionen im Vergleich“. In: Ulla Fix u.a. (Hgg.), *Brauchen wir einen neuen Textbegriff? Antworten auf eine Preisfrage*. Frankfurt am Main u.a. 2002, S. 17-29. Zum hier vorliegenden Begriffsverständnis vgl. auch Jurij M. Lotman, *Struktur literarischer Texte*. München 1972, S. 83-86.

⁵⁴ Stöckl, „Multimodalität“, S. 4.

⁵⁵ Ebd., S. 20.

⁵⁶ Schmitz, „Multimodale Texttypologie“, S. 332.

⁵⁷ Vgl. Bateman/Schmidt, *Multimodal Film Analysis. How Films Mean*, S. 77.

⁵⁸ Vgl. Schmitz, „Multimodale Texttypologie“, S. 332.

⁵⁹ Vgl. Stöckl, „Multimodalität“, S. 6.

konstitutive Einfluss des Mediums auf Struktur und Gebrauch von Modalitäten.⁶⁰ Sie müssen demnach „materiell-medial realisiert [...] und in einer raumzeitlichen und sozialen Situation verwendet“ werden und „verfügen über eine interne Strukturierung, die Bedeutungen, Kombinationsmöglichkeiten und Gebrauchsfunktionen ihrer Zeicheninventare regelt“.⁶¹

Allerdings wäre damit bis hierhin nicht allzu viel gewonnen: Zeichenmodalitäten im Musikvideo sind ja durchaus differenzierbar und, wie zu sehen war, auch entsprechend differenziert worden. Auch ihr Zusammenspiel – siehe Decker – ist bestimmt und Mechanismen der Bedeutungsgenerierung auf Basis des Zusammenspiels – Stichwort: Bedeutungsknoten – ergründet. Man trägt also Eulen nach Athen: Der Texttyp im Fall von THRILLER kombiniert gesprochene, gesungene und geschriebene Sprache, bewegte Bilder, Geräusche und Musik; es lassen sich strukturelle, illustrative, konnotative und andere Relationen zwischen Bild, Ton und Sprache benennen und auf diese Weise der spezifischen Profilierung des Starimages im gegebenen Fall nachgehen. Worin also genau besteht das heuristische Potenzial einer Präzisierung der Begriffe ‚Zeichenmodalität‘ und ‚Multimodalität‘? Die Antwort im Anschluss an die obigen Überlegungen: Das Konzept ‚Zeichenmodalität‘ soll dazu dienen, Zeichenrealisierungsformen zu erfassen, um auf Grundlage dessen komplexe semiotische Anordnungen in multimodalen Texten und ihre Bedeutungen zu analysieren, in Texten, die durch mehr als ein Zeichensystem und in mehr als nur einem Modus gestaltet sind.

Mit einer solchen Zielsetzung geht es an dieser Stelle – erneut mit Stöckl – erstens darum, die syntaktischen, semantischen und pragmatischen Eigenschaften der verschiedenen Zeichenmodalitäten zu registrieren und entsprechend in der Decodierung der semiotischen Struktur zu berücksichtigen, die interne Struktur von Sprache (als „Zeichensystem par excellence“⁶²), Bewegtbild, Musik, Geräusch auf einer Skala zwischen ‚stark konventionell/kaum motiviert‘ bis ‚schwach konventionell/stark motiviert‘, ihre (semantischen) Ausdruckspotenziale und die Frage nach den kommunizierten Bedeutungen und Aussagen, wie auch die (pragmatischen) Kommunikationsfunktionen und -aufgaben etwa mit Michael Hallidays Metafunktionen⁶³ oder – wie bei Kraus – mit Roman Jakobsons Sprachfunktionen. Das Ziel besteht in unserem Zusammenhang indes nicht darin, die beteiligten Zeichenmodalitäten isolativ auf ihren Charakter hin zu bestimmen, auch nicht darin, sie in ihrer Spezifik von anderen abzugrenzen oder mit jenen zu vergleichen, sondern es besteht darin, die jeweilige Modalität in ihrer konkret gegebenen Spezifik hinsichtlich Syntax, Semantik und Pragmatik zu analysieren und in ihrer Interaktion mit anderen beteiligten Modalitäten zu sehen, um dem Musikvideo als kohäsiven und kohärenten Text gerecht werden zu können. Das erscheint zwar aufwendig, ist aber in Anbetracht der strukturell-semiotischen Komplexität des Texttyps unumgänglich.

⁶⁰ Vgl. ebd., S. 8.

⁶¹ Ebd., S. 9.

⁶² Ebd., S. 11.

⁶³ Vgl. Michael A. K. Halliday, *Language as Social Semiotic. The Social Interpretation of Language and Meaning*. London 1978.

Dazu ist zweitens ein Modell zur multimodalen Textsortenanalyse vonnöten, das von zwei Seiten aus aufgestellt werden kann: Zum einen ausgehend von Stöckl selbst, der fünf Beschreibungsebenen listet und als heuristisch aufeinander aufbauende Analyseschritte ansetzt: (1) *Gliederung/Abgrenzung*. Die analytische Ausrichtung auf die internen Signale der Textstrukturierung, auf die paradigmatische Ordnung und die nach außen abgrenzende Rahmung. (2) *Handlungsstruktur*. Die Analyse der pragmatischen Leistung der Modalitäten, ihre funktionale Verteilung in Abfolgen und Anordnungen des Textes als Kommunikationsakt. (3) *Themenstruktur*. Die Analyse des jeweiligen Beitrags einzelner Zeichenmodalitäten innerhalb der thematischen Gesamtkonfiguration. (4) *Multimodale Verknüpfung*. Laut Stöckl der „Kern einer multimodalen Textsortenanalyse“.⁶⁴ In den Blick gerät die gegenseitige Bezugnahme der Zeichenmodalitäten untereinander, um zu bestimmen, „welche Elemente sich in welcher Weise kohäsiv oder kohärent verhalten und welche pragmatischen oder rhetorischen Funktionen Modalitäten in wechselseitiger Bezogenheit füreinander übernehmen“⁶⁵. (5) *Intertextualität*. Die Frage nach und die Analyse von expliziten und impliziten Bezugnahmen auf außertextuelle Zeichenangebote sowie ihre bedeutungsgenerierende Integration in die gegebene Konfiguration.

Zum anderen ließe sich ein dezidiert filmanalytischer Ansatz einbringen: Bateman und Karl-Heinrich Schmidt erheben Multimodalität zum „essential feature of filmic meaning making“⁶⁶ und nehmen in Fortführung von Christian Metz' *Großer Syntagmatik* eine „combination of distinct semiotic codes“⁶⁷ an, die – theoretisch – in zwei Gruppen aufgeteilt werden können: *Reality codes* auf der einen Seite und *representation codes* auf der anderen. Die einen – *reality codes* – formieren die Gesamtheit solcher (ikonischen, indexikalischen oder symbolischen) Zeichen(-systeme), die die dargestellte Welt in ihrer semiotischen Gesamtheit modellieren und in ihrer ‚realen‘, sozialen, physikalischen usw. ‚Welthaftigkeit‘ wahrzunehmen erlauben. Codes wie soziale Umgangsformen, Systeme des Fühlens und Denkens, Kleiderordnungen, Farbkodierungen und ideologische Codes, Codes des Pflanzen- und Tierreichs oder auch maschinelle Codeformen usw. sind (in der Regel) profilmisch verankert und nicht filmspezifisch, werden aber filmisch verwendet.⁶⁸ Anders im Fall der *representation codes*, die als „essential mechanisms of filmic discourse construction“⁶⁹ die filmische Mediatisierung grundieren: *Mise-en-cadre* und *Mise-en-chaîne* („technical features of film“), Musik und Sprache bzw. die filmische Einstellung und größere Syntagmen („audio-visual elements“).⁷⁰ Leitend ist dabei die Idee, die filmische Ebene der

⁶⁴ Stöckl, „Multimodalität“, S. 23.

⁶⁵ Ebd.

⁶⁶ Bateman/Schmidt, „Multimodal Film Analysis“, S. 91.

⁶⁷ Ebd., S. 134.

⁶⁸ Ebd., S. 142. Vgl. auch Étienne Souriau, „La structure de l'univers filmique et le vocabulaire de la filmologie“. In: *Revue internationale de Filmologie* 2/7-8 (1951), S. 231-240 (dt. „Die Struktur des filmischen Universums und das Vokabular der Filmologie“. In: *Montage/av* 6/2 (1997), S. 140-157).

⁶⁹ Bateman/Schmidt, „Multimodal Film Analysis“, S. 143.

⁷⁰ Ebd., S. 158.

Präsentation (der Erzählung), wie andernorts für literarische Texte bereits unternommen,⁷¹ als „a multiply articulated and articulable audiovisual stream“,⁷² als kombinierte Gestaltung mehrerer distinkter Einheiten zu konturieren. Die Codes bilden in ihrer Gesamtanlage das materielle audiovisuelle Substrat zur Bedeutungsgenerierung und Informationsvermittlung, sie formieren demnach, in ihrer „complexity of interaction“⁷³ die Grundlage für komplexe Bedeutungsstrukturierung, wie an anderer Stelle ebenfalls Jurij Lotman postuliert: Der Film

enthält auch unmittelbar sprachliche Mitteilungen, musikalische Mitteilungen, er aktiviert außertextliche Beziehungen, die vielfältige Bedeutungsstrukturen an den Film anschließen. Alle diese semiotischen Schichten sind auf komplexe Weise miteinander verzahnt, und ihre wechselseitigen Beziehungen ergeben wiederum semantische Effekte. Diese Fähigkeit des Films, die verschiedenartigsten Typen der Semiose ‚aufzusaugen‘ und in einem einzigen System zu organisieren, ist gemeint, wenn vom synthetischen oder polyphonen Film gesprochen wird.⁷⁴

Synthetischer oder polyphoner Film im Sinne Lotmans ist hier gleichbedeutend mit seiner Eigenschaft als multimodale Textsorte. Kehren wir damit zum medien-theoretisch-semiotischen Ansatz und seiner Applikation in unseren Untersuchungszusammenhang zurück und komplettieren das transmediale Gesamtmodell durch filmanalytische Zugriffsweisen, die für eine Beschäftigung mit einem multimedialen Texttyp wie dem Musikvideo essenziell sind und ebenfalls reflektiert werden müssen,⁷⁵ so ließe sich der Ansatz einer multimodalen Textanalyse fruchtbar machen. Das hier vorgeschlagene Teilmodell sieht ganz konkret zwei, auf die multimodale Verfasstheit gegebener Texte ausgerichtete und aufeinander aufbauende Analyseschritte vor, die umstandslos in generelle wie spezielle Interpretationstheorien überführt werden können.⁷⁶

⁷¹ Vgl. das Ebenenmodell von Erzählungen im erzähltheoretischen Entwurf Wolf Schmid. Schmid spricht bei der Präsentation der Erzählung, der „Phäno-Ebene“ eines gegebenen Erzähltextes, von der „Materialisierung der medial noch nicht manifestierten Erzählung im Medium“ (Wolf Schmid, *Elemente der Narratologie*. 3. Aufl. Berlin/Boston 2014, S. 224 und 242-244). Eine auf diesem Entwurf aufbauende und auf filmische Zusammenhänge übertragene Ebenendifferenzierung habe ich andernorts vorgeschlagen vgl. Stephan Brössel, *Filmisches Erzählen. Typologie und Geschichte*. Berlin/Boston 2014, S. 46f.

⁷² Bateman/Schmidt, „Multimodal Film Analysis“, S. 162.

⁷³ Ebd., S. 144.

⁷⁴ Jurij M. Lotman, *Probleme der Kinoästhetik. Einführung in die Semiotik des Films*. Frankfurt am Main 1977, S. 145.

⁷⁵ Vgl. dazu grundsätzlich Hans J. Wulff, „Semiotik der Filmanalyse. Ein Beitrag zur Methodologie und Kritik filmischer Werkanalyse“. In: *Kodikas/Code* 21 (1998), S. 19-26. <http://www.derwulff.de/2-84>; Abruf am 06.09.2022.

⁷⁶ Grundsätzlich zur multimodalen Textsortenanalyse Stöckl, „Multimodalität“, S. 22-32. Zur Interpretationstheorie allgemein vgl. Josef Simon (Hg.), *Zeichen und Interpretation*. Frankfurt am Main 1994; Albert Menne, *Einführung in die Methodologie. Elementare allgemeine wissenschaftliche Denkmethode im Überblick*. Darmstadt 1980, und Titzmann, *Strukturelle Textanalyse*. Zur Methodik der Film- und Mediensemiotik vgl. Wulff, „Semiotik der Filmanalyse“, Dennis Gräf u.a.,

- (1) *Bestimmung von multimodaler Spezifik und semiotischen Signifikanzen*. Textsortenbestimmung und *canvas*-Gestaltung, Signifikanzen des Umgangs mit Modalitäten, Teilsemiosen und evozierte Zeichenqualitäten, handlungs- und themenstrukturelle, intertextuelle Aspekte der Kohärenzstiftung, globale Spezifik der multimodalen Verknüpfung
- (2) *Makro- und mikrostrukturelle Erfassung bedeutungstragender Einheiten*. Aspekte der Gliederung/Abgrenzung, multimodale Verknüpfung und Paradigmenbildung, Propositionsanalyse

Die Liste der Analyseaspekte, so übersichtlich sie hier erscheint, ist bewusst offen und ausbaubereit gehalten, und nicht nur textsemiotisch kompatibel, sondern zugleich anschlussfähig sowohl an Arbeiten zur Textsorte ‚Musikvideo‘ als auch an die semiotische Medientheorie. Denn der Zugriff wird erstens allgemeinen Annahmen zur Multimodalität und der Analyse multimodaler Texte gerecht, indem er zweitens übergreifende transmediale Basistheoreme mit filmmedienspezifischen Theoremen kombiniert. Drittens liefert er auf einer übergeordneten (methodologisch-interpretationstheoretischen) Ebene Handwerkswerkzeuge zur Erfassung multimodaler Strukturanordnungen und ihrer Funktionalisierung in av-Formaten.

4. Multimodale (Über-)Determinationen zwischen ‚Musikvideo‘ und ‚Horrorfilm‘: Michael Jacksons THRILLER

Das hier gewählte Beispiel, das klang nun mehrfach an, genießt sicherlich einen Sonderstatus, und das in mehreren Hinsichten. Zum einen *historisch* besehen: Zu Beginn der 1980er-Jahre produziert und ausgestrahlt, gilt es als Meilenstein der Musikvideogesichte, und das nicht zuletzt auch aufgrund seiner reflexiven (Selbst-)Einbettung in einen medienkulturgeschichtlichen Kontext – Stichwörter: globale Ausdifferenzierung, Erneuerungsbewegungen, ‚visuelle Lust‘ (Mulvey), ‚Cinema of Urban Crises‘, Rückbesinnung auf die 1950er-Jahre usw.⁷⁷ –, indem es, unter der Regie des für seine Horrorfilme bekannten Regisseurs John Landis produziert, eine Vielzahl an Referenzen auf die Textsorte ‚Spielfilm‘ im weiteren und auf den Horrorfilm im engeren Sinne anlegt. Zum zweiten – und das steht damit in Zusammenhang – in *struktureller* Hinsicht: Die Länge des Clips ist mit knapp 14 Minuten deutlich exponiert. Flankiert wird die syntagmatische Einheit

Filmsemiotik. Eine Einführung in die Analyse audiovisueller Formate. Marburg 2017, und Jan-Oliver Decker, „Strukturalistische Ansätze in der Mediensemiotik“. In: Martin Endres/Leonhard Herrmann (Hgg.), *Strukturalismus heute. Brüche, Spuren, Kontinuitäten*. Stuttgart 2018, S. 79-105. Zur Musikvideanalyse Christoph Jacke, „Who Cares about the Music in Music Videos?“. In: Henry Keazor/Thorsten Wübbena (Hgg.), *Rewind, Play, Fast Forward. The Past, Present and Future of the Music Video*. Bielefeld 2010, S. 179-194.

⁷⁷ Vgl. Willem Strank, *Handbuch Filmgeschichte. Von den Anfängen bis heute*. München 2021, S. 145, 157 und 168f.

des Songs durch zwei Teilsegmente mit erkennbar narrativem Charakter: Die Initialsequenz, die eine Film-im-Film-Struktur präsentiert, und die Schlusssequenz, die das vorangegangene Geschehen retrospektiv als Traum markiert und wiederum handlungs- wie zeichenlogisch an die erste Sequenz anschließt. Man darf also davon ausgehen, dass THRILLER eine Herausforderung für Theorie und Praxis darstellt, gleichzeitig aber auch als Chance begriffen werden darf, den Versuch eines semiotischen Ansatzes auf seine Tragfähigkeit hin zu prüfen und heuristisch abzusichern.

4.1 Komplexer ,Thrill': Diegetische und narrative Ebenenstrukturen und inner-semiotisches Verweissystem

Als Verständigungsbasis für meine Ausführungen möchte ich die folgende, heuristisch motivierte Segmentierung vorschlagen, die ich aus einem basalen Sequenz- und Handlungsprotokoll ableite (vgl. Abb. 1) und die unterschiedliche filmische Wiedergabemodi (Modus ,Musikvideo' u. Modus ,Spielfilm') und Ebenenstrukturen (narrative u. diegetische Ebenen) berücksichtigt.

Segment 1: Vorspann u. Titel

Segment 2: Rendezvous, Michaels Transformation zum Werwolf und Verfolgung durch den Wald

Segment 3 mit Teilsegmenten a u. b: Kinobesuch mit vorzeitigem Abbruch (a), vor dem Kino, Angstthema, musikalischer Einsatz des Songs (b)

Segment 4 mit Teilsegmenten a–f: Fußweg u. Songperformance der Strophen 1–3 (a); ,Rap' u. ,Auferstehung' (b), Choreografie der Untoten (c), Songperformance des Refrains (in Wiederholung) u. Choreografie (d). *Sonderteilsegmente e u. f:* Bedrohungsszenerie u. Michaels Transformation zum Untoten (e), Flucht zum Haus und akute Bedrohungsszenerie (f)

Segment 5: Aufwachen auf der Couch, Beruhigung, Michaels Blick in die Kamera

Segment 6 mit Teilsegmenten a–c: Abspann u. Standbild (*freeze frame*), *Zoom-In* von Michaels Gesicht (a), Songperformance mit Choreografie (b), Abspann und Standbild der letzten Einstellung mit Heranfahrt an Gesicht eines anonymen Untoten (c)

Sequenzprotokoll THRILLER (USA 1983, J. Landis)

Handlungssegment	Sequenznr.	Inhalt	Anzahl der Einstellungen	Gesamtzeit/Sequenzzeit (Sek.)
-	1	Vorspann: „Due to my...“; Michael Jackson's Thriller	2	00:00:00-00:00:20/ 21
I	2	Date: Michael und Freundin fahren nachts oder am späten Abend in einen Park. Sie werden ein Paar. M möchte ihr ein Geheimnis gestehen, verwandelt sich jedoch plötzlich in einen Werwolf. Sie flüchtet.	49	00:00:21-00:03:01/ 160
	3	Verfolgungsjagd im Wald: Zunächst vermag sie vor ihm zu flüchten. Schließlich schneidet er ihr den Weg ab und bedroht sie.	16	00:03:02-0:03:41/ 39
	4	M und F im Kino: Horrorfilm im Kino, offenbar stellen die Sequenzen 1-3 diesen Film dar. Sie fürchtet sich vor dem Film, er amüsiert sich. Sie verlässt das Kino, bevor der Film endet. Er folgt ihr.	9	00:03:42-00:04:13/ 31
II	5	Vor dem Kino: Gespräch über Angst. Er neckt sie etwas. Sie streitet ab, obwohl sie weiß, dass sie tatsächlich Angst hatte. Sie machen sich auf den Weg.	4	00:04:14-00:04:39/ 25
	6	Song beginnt: Sie gehen. M umtanzt sie (und bezieht sich damit auf den Text) und singt zwei Strophen des Songs.	14	00:04:39-00:06:23/ 104
III	7	Friedhof: Sie passieren einen Friedhof und verlassen die Szenerie. Auf dem Friedhof regt sich etwas. Nach und nach öffnen sich Gräber und Gruften. Untote erscheinen. Ebenso aus der Kanalisation tauchen Untote auf.	15	00:06:24-00:07:53/ 89
	8	Bedrohung: Die Protagonisten werden von den Untoten umzingelt. Stehen Rücken an Rücken. Plötzlich ist auch M ‚verwandelt‘ und ein Untoter.	13	00:07:54-00:08:29/ 36
	9	Performance und Refrain: Die Untoten tanzen eine Choreografie. Es folgt der Refrain des Songs. Anschließend nehmen die Untoten abermals die Protagonisten in den Blick und nähern sich ihr.	51	00:08:30-00:10:35/ 125
IV	10	Flucht zum Haus und Bedrohung: F flüchtet und läuft auf ein Haus zu, erreicht die Tür. Sie hastet in ein Zimmer. Die Untoten, angeführt von M, brechen in das Haus ein und erreichen sie. Sie nähern sich wieder F; M streckt den Arm nach ihr aus und berührt sie.	32	00:10:35-00:11:46/ 71
V	11	Realitätenwechsel?: F erwacht schreiend – es war nur ein Traum. Vor ihr steht M, der sie anlächelt (er trägt die Kleidung aus Handlungssegment I), seinen Arm auf ihrer Schulter. Er hilft ihr auf und legt ihr den Arm um die Schulter. Er blickt in die Kamera. Seine Augen sind verändert wie in Handlungssegment I (bei der Verwandlung zum Werwolf).	4	00:11:47-00:12:04/ 17
-	12	Abspann: Freeze Frame. Credits. Refrain (mit Tanz der Untoten). Tanzeinlage dreier Untoter. Ein Einzelner.	29	00:12:04-00:13:42/ 98

Abb. 1: Sequenzprotokoll THRILLER (J. Landis, USA 1983)

Zur Frage der Übertragung des Protokolls in die Segmentierung: S1 und 6 sind deckungsgleich mit den Sequenzen 1 und 12. S2 – die subordinierte narrative Ebene des ‚Films im Film‘ – fasst die Sequenzen 2 und 3 zusammen. S3 und 4 umfassen zwar eine kohärente narrative und diegetische Ebene, sind aber deshalb sondiert, da S3 nicht zur Songperformance zählt, gleichwohl – und darauf wird einzugehen sein – das Intro des Songs hier bereits einsetzt. Alle Segmente 2 bis 4 wiederum sind insofern geklammert, als sie eine subordinierte diegetische (Traum-)Ebene präsentieren, während S5 den Wechsel zur übergeordneten, film-internen Wirklichkeit anzeigt. Die Sondersegmente umfassen Moduswechsel, auf die ich noch zu sprechen komme.

Die offensichtliche Komplexität des Videos rührt, so deutet es diese Übersicht an, aus der mehrfachen Ebenenstruktur mit internem semiotischem Verweissystem, der hohen intermedialen Referenzendichte und dem Verhältnis von Narration und (narrationsobstruierender) Performanz, wobei dem ersten und zweiten Aspekt sicherlich der größte Anteil zukommt: Die ‚Film-im-Film‘-Struktur, die als solche retrospektiv markiert wird, einerseits und die ebenfalls retrospektiv markierte Darstellung eines Traums andererseits. In dem einen Fall etabliert THRILLER zwei divergente *narrative* Ebenen: die des Films, der im Kino läuft und die Zuschauerschaft – unter ihnen Michael und seine Freundin – in einen Zustand zwischen Schrecken und Vergnügen versetzt, und die des dargestellten Geschehens nach dem Kinobesuch. Der Kinofilm offenbart sich im Nachhinein als Binnener-

zählung. So sehr aber auch die Thematik dieser beiden Ebenen bis hin ins kleinste Detail semantischer Setzungen und Entsprechungen hinein kongruiert und damit an der übergeordneten, makrostrukturellen Paradigmatik des Videos beteiligt ist, so verwunderlich ist innerhalb dieser Anlage, dass sie innerdiegetisch nur begrenzt von Relevanz ist: Zwar ist der Film Anstoß für das weitere Geschehen und entpuppt sich syntagmatisch als dessen strukturelle Analogie, aber dass die Figuren sich im Film gewissermaßen selbst zuschauen, wird von ihnen weder erkannt, noch problematisiert und reflektiert. Daher muss wohl angenommen werden, dass die Ebenenstruktur *narratorial* relevant gesetzt wird, nicht aber *figural*: Es geht dem Video hier um präsentationsaktbezogene Aussagen, nicht um solche, die auch die Figuren thematisieren müssten. Allerdings sind Letztere freilich durchaus von dieser Strategie betroffen, denn wie den Figuren im Kinofilm ergeht es letzten Endes ja auch ihnen: Der Horrorfilm holt sie gewissermaßen in ihrer eigenen, außerfilmischen und vermeintlichen sicheren Welt ein. In dem anderen Fall indiziert der in Sequenz 11 vorliegende Traummarker seinerseits wiederum den Wechsel zwischen zwei *diegetischen* Ebenen, während der einen Ebene (dargestellt in den Segmenten 2–4) der Status einer introspektiv präsentierten, subjektiv-psychomentalen ‚Innenwelt‘ mit (in intersubjektiver Hinsicht) nichtfaktischer Ontologie zugeschrieben wird, entspricht die andere genau dem: der textinternen, intersubjektiv erfahr- und erlebbaren Wirklichkeit, in der – so der Twist an dieser Stelle – banalerweise eben auch geträumt werden kann. Doch auch hier unterläuft eine Strukturanalogie das etablierte Gefüge, denn mit seinem Blick in die Kamera, der Direktadressierung der Zuschauerschaft offenbart Michael diejenigen gelbleuchtenden Augen wie in Sequenz 2,⁷⁸ was zu zwei Schlüssen herausfordert: Zum einen, dass es um die Beständigkeit von ‚Realität‘, wie sie vom Video modelliert wird, nicht zum Besten bestellt ist, ihr Aufbau stets auch ihre Unterminierung impliziert, zum anderen, dass doch zumindest die Figur Michael um die Botschaft des Videos weiß, nämlich dass er selbst es ist, die Starpersona, die das entscheidende Element in diesem Zusammenhang inkorporiert, während die anderen Figuren entweder nicht wissen – wie im Fall der namenlosen Freundin – oder in ihrem Tun genau diesen Umstand noch unterstreichen – wie die Untoten mit ihrem Tanz mit Michael in ihrer Mitte. Das Ganze wird zusätzlich durch unterschiedliche Vor- und Rückverweise zeichenhaft gesättigt und festgezurr. Die gelb leuchtenden Augen, die Kleidung Michaels, die Rollenbesetzung und Namensgebung, das sprachliche Zeichen ‚Vincent Price‘ und dessen Stimme (im Kinofilm wie im ‚Rap‘) – alles das verdichtet sich zu einem Netz der semantischen Überdetermination, in deren Zentrum das Künstlerimage steht.

Zu erkennen ist demnach schon auf den ersten Blick, dass das Video einerseits oberflächenstrukturell auffallend umständlich gestaltet ist und dabei – im späteren Verlauf der Filmgeschichte der 1990er-Jahre enorm populär werdende –

⁷⁸ Zur Direktadressierung über den figürlichen Blick vgl. Francesco Casetti, *Inside the Gaze. The Fiction Film and Its Spectator*. Bloomington 1998. Zur „metaphorische[n] Rolle des Sehens und des Auges“, die sich an dieser Stelle auch für THRILLER anbietet, vgl. Thomas Elsaesser/Malte Hagener, „Auge und Blick“. In: Dies. *Filmtheorie zur Einführung*. 3. Aufl. Hamburg 2011, S. 103-135; insb. S. 108 und 117-121.

Formen komplexen Erzählens geltend macht, die die Forschung unter den Slogan der Unzuverlässigkeit und des Mind-Bender aufgearbeitet hat.⁷⁹ Andererseits erscheint es tiefenstrukturell, also hinsichtlich seiner paradigmatischen Ordnung und Ideologie – seiner vermittelten Botschaft –, äußerst homogen, ja geradezu tautologisch-redundant. Wiederholt geht es um (erotisch konnotierte, heterosexuelle) Paarbeziehungen, immer ist es dabei Michael Jackson, der (gemeinsam mit der Schauspielerin Ola Ray) in einer solchen Beziehung gezeigt wird und agiert, wiederholt wird eine ‚Realität‘ aufgebaut, dann aber in ihrer ontologischen Grundkonstitution ausgehebelt, wobei abermals Jackson als entscheidender Faktor vom Menschlichen ins Nichtmenschliche transformiert, sich ‚entfremdet‘, die Partnerin existenziell bedroht und das Realitätsgefüge ins Schwanken bringt, dabei aber bezeichnenderweise stets im Zentrum allen Gezeigten steht.

4.2 ‚Thriller‘ in oszillierendem Gesamtmodus: Textsorte, *canvas* und Multimodalität

Komplex zeigt sich THRILLER aber auch in seiner (multi-)modalen Signatur, die semiotisch gesprochen den Konnex zwischen textueller Oberflächenstruktur und reduzierbarer Tiefenstruktur ausmacht, die Schaltstelle der Zeichenprozessierung und damit das Fugenelement zwischen Form- und Inhaltsebene. Und hierin, so ließe sich behaupten, besteht wohl auch der Reiz, den das Video bis heute ausstrahlt: Es ist ein in seiner Machart ‚besonderer‘ Videoclip. Doch inwiefern?

Bereits bei der Textsortenbestimmung ergeben sich Schwierigkeiten, die nicht außer Acht gelassen werden dürfen, ja, die gar ganz bezeichnend sind im Hinblick auf die Spezifik, die Jacksons Video ausmacht. Auf der einen Seite ist der Fall klar: Das Medienprodukt ist unmissverständlich als Musikvideo intendiert und vermarktet worden. Zudem weist es entsprechende, für die Textsorte ganz charakteristische Eigenschaften auf: Performance des Songs (des gleichnamigen Albums aus dem Jahr 1982), Auftritt der Starpersona, strukturelle und thematische Reflexion derselben. Auf der anderen Seite allerdings reicht der Text in seiner Machart über generische Grenzen hinweg, denn er weist integrierte Spielfilmanteile auf und ist demnach nicht allein in quantitativer Hinsicht deutlich länger als andere Musikvideos, sondern tritt auch in qualitativer Hinsicht hervor, indem er seine systemreferenzielle Strukturgebung auf der einen Seite thematisch an die Songperformance anbindet, auf der anderen Seite darüber hinaus ein metatextuelles Moment geltend macht, nämlich die reflexive Selbstverortung in der Medienkulturgeschichte und Popkultur der 1970er- und 1980er-Jahre. Die Forschung ist sich richtiggehend einig, von einem Musikvideo zu sprechen, aber es

⁷⁹ Vgl. Steven Johnson, *Everything Bad is Good for You*. London 2006, und Warren Buckland (Hg.), *Puzzle Films. Complex Storytelling in Contemporary Cinema*. Malden/Mass. 2009. Vgl. auch die Überblicksartikel in *Mind-Bender: Begriffe – Forschung – Problemfelder (= Paradigma. Studienbeiträge zu Literatur und Film 1 [2017])*. <https://www.uni-muenster.de/Germanistik/ffm/Paradigma/paradigma1/index.html>; Abruf am 21.11.2022.

handelt sich eben um ein Video mit ausgeprägter Spielfilmästhetik und selbst- und metareflexiver Gestaltung. Und in diesem seinem Hybridcharakter weicht es vom (damals) kulturell Erwartbaren deutlich ab und hat in dieser Anlage ebenso Konsequenzen für die multimodale Analyse. Denn die Gestaltung der medialen Projektionsfläche (des *canvas* nach Bateman/Sachs-Hombach) zeigt sich nicht nur vielfältig, sondern zugleich auch variativ: Der generelle Modus der Informationsvergabe, das Zusammenspiel aller Zeichenmodalitäten – im Folgenden ‚Gesamtmodus‘ – ändert sich auf den ersten Blick nicht, er bleibt audiovisuell. Und doch scheint im Wechsel zwischen ‚spielfilmartig‘ und ‚musikvideoartig‘ etwas auf, bei dem im Anschluss an die Idee des vorliegenden Bandes von einer ‚Medialität zweiter Ordnung‘ gesprochen werden könnte. Ein modal oszillierendes Moment bricht sich etwa Bahn, sobald der Song (in S3) einsetzt – eine multimodale Strategie, auf die noch einzugehen sein wird –, der Gesamtmodus changiert schließlich gänzlich mit Einsatz der Performance (in S4). Multimodalität als ‚Medialität zweiter Ordnung‘ zeigt sich gerade in diesem Wechselspiel der modal gesteuerten Informationsvergabe. In Anbetracht dessen ließe sich, einstweilen grob, für die ‚Ausstattung‘ des *canvas* festhalten, dass am Zusammenspiel aus Bild und Ton die für Musikvideos üblichen Zeichenmodalitäten beteiligt sind, wobei wiederum der Gesamtmoduswechsel zwischen ‚Spielfilm‘ und ‚Musikvideo‘ sinnfällig ist, bedingt durch eine alternierende Dominanz beteiligter Modi, je nachdem, welches Maß an Relevanz ihnen bei der Informationsvergabe zukommt. Und die Art und Weise der Präsentation, das fügt sich wie aufgezeigt ein, ist ja zugleich auch implizites Thema des Textes.

Kohärenzstiftend wirkt dabei insgesamt vor allem das angesprochene semiotisch engmaschige Verweissystem sowie natürlich das Lied „Thriller“: Das in der Figurenkonstellation verankerte Paradigma der Paarfindung, das in allen Handlungssegmenten, auf allen vorhandenen Ebenen variiert wird, in Rekurrenz und Modifikation identischer Zeichen und Zeichenträger überdeterminiert in Erscheinung tritt und demnach als hochgradig relevant zu erachten ist. Michael Jackson und Ola Ray in unterschiedlichem Setting und verschiedenen Outfits, verbal, non- und paraverbal kommunizierend, interagierend, zusätzlich Jackson als generisch aufgeladener Zeichenträger, indem er sich in ‚Monster‘ verwandelt, Ray als ‚final girl‘. Daneben werden, um nur ein weiteres wesentliches Element dieses Referenznetzes zu nennen, Name und Gesicht von Vincent Price grafisch bzw. visuell gezeigt (vgl. S3), im ‚Rap‘ (S4b) dann stimmlich komplettiert. Der Song ist gegenüber der Albumversion unverkennbar umgestaltet, determiniert das Bildgeschehen und implementiert mit dem Gesang eine Textschicht, die innerdiegetisch angebunden wird: Jackson besingt spielerisch und erotisch konnotiert ‚sein Mädchen‘, entwirft dabei ein Bedrohungsszenario eines ‚Bösen kurz vor Mitternacht‘, dem das besungene Du ausgesetzt ist, exponiert im Refrain schließlich einen Zustand des ‚Thrillers‘, der nach seiner Verwandlung in einen Zombie gar real wird. Die Abwandlung des Songs ist in multimodal-analytischer Sichtweise signifikant und soll ebenfalls noch Berücksichtigung finden. Was aber auch hier festzuhalten ist: Trotzdem der *canvas* derart vielgestaltig und variativ konstituiert ist, hindert es uns nicht daran, THRILLER als kohärenten Text zu rezipieren. Die

Frage wäre aber, wie Bedeutung konkret aufgebaut und eine Botschaft, vielleicht auch mehrere Botschaften transportiert werden – und zwar unter besonderer Beachtung seiner multimodalen Verfasstheit. Dabei wäre den bisherigen Erkenntnissen ein weiteres Spezifikum anzufügen, denn nicht nur *kombiniert* THRILLER unterschiedliche Gattungscharakteristika und wechselt dementsprechend im Gesamtmodus, auch *hybridisiert* er sie und stiftet mit Hilfe dieses Prinzips Sinn.

4.3 ‚Thriller‘ in Überdetermination: Makro- und mikrostrukturelle Hybridisierungsprozesse

Damit gelangen wir zum nächsten Analyseschritt, mit dem die mikro- und makrostrukturelle Erfassung bedeutungstragender Einheiten ins Zentrum rückt – die Stratifikationen in den angesprochenen Segmenten haben hierbei entscheidenden Anteil. Doch zunächst zur Rahmung, die für sich genommen bereits beachtenswert ist: 6b ist ja einerseits als ein solches Rahmensegment angelegt, wird aber neben der Präsentation der Credits für eine erneute Darbietung der Choreografie nebst gesungenem Refrain genutzt. Die pragmatisch-textuelle Begrenzung wird demnach doppelt funktionalisiert, nämlich zum einen als Rahmen, die sie medial (ersichtlich am ablaufenden Timecode) und konventionell (angezeigt durch die Credits) darstellt, und zum zweiten in diegetisch lose angeschlossener Form als Ort der wiederholten Darstellung des Attraktionsmomentes in Form von Performance und deskriptiven Syntagmen (Einstellungen von Gesichtern tanzender Untoter usw.). In anderer Weise auffällig ist das initiale Segment 1. Zunächst die schriftliche und apologetische Aussage des Künstlers Jackson, mit der er sich vom Okkulten distanziert, was in Anbetracht der Sachlage zunächst einmal seltsam anmuten mag, sind die Fiktionssignale schließlich derart deutlich, dass sich ein entsprechender Hinweis eigentlich erübrigt hätte. Andererseits besteht ja, wie aufgezeigt, eine der wesentlichen Textstrategien genau darin, die ‚Person‘ Michael Jackson (als die er schließlich inszeniert wird) mit dem Okkulten semantisch zu korrelieren, und zwar rekurrent und exponiert. Es folgt der Titel mit ausgeprägt multimodaler Spezifik – man könnte auch sagen, ein Titel, der das semiotische Programm ankündigt und gleichsam vorwegnimmt: „Michael Jackson’s Thriller“, alles in blutigen Lettern, demnach bereits für sich genommen multimodal, insofern hier ein für den Horrorfilm zentrales ‚Material‘ – Blut – als Zeichenträger zum Einsatz kommt. Dieses offenkundig nicht unübliche Verfahren – man denke im Horrorfilm an die zahlreichen Zeichen und Textbotschaften in Blut – dient im vorliegenden Fall klar der referenziellen Anbindung an ebendieses Genre: THRILLER ‚dockt‘ hier gewissermaßen gleich zu Beginn an den Horrorfilm an. Das in der Tonspur angelegte Atemgeräusch ist auf semantischer Ebene paradigmatisch: Wie das Blut der Lettern indiziert es das ‚Leben‘, das auf seiner Grenze zum Tod schließlich Thema des Videos insgesamt gerät. In programmatischer Sicht zeigen demnach also modale Semioseoperatoren in ihrer Kombinati-

on an, worum es dem Video über seine thematische Ausrichtung hinaus auch geht, nämlich, wie beschrieben, um modale ‚Wechselspiele‘ zwischen Bild und Ton.

Die Segmente 3b, 4b und 6b sind subtiler gestaltet und dienen letztlich der genannten allgemein prägenden Textstrategie der generischen Hybridisierung. In S3a setzt der Song ein und untermalt das Geschehen: das Paar vor dem Kinoeingang, er ist belustigt ob ihrer Schreckhaftigkeit, sie beteuert das geringe Ausmaß ihrer Angst – bei alledem: keine Signale darauf, dass eine Songperformance beginnt, vielmehr findet lediglich ein Moduswechsel auf auditiver Ebene statt, bei dem Jacksons Musik den diegetischen Ton der Kinoatmo komplettiert, der wiederum syntagmatisch ein (fiktionsintern) nichtdiegetischer Ton des Filmscores (Filmmusik) vorangeht. Das mag mit Blick auf die Alltagsrezeption des Clips überhaupt nicht nennenswert zu sein, ist der Text schließlich in eine Kommunikationssituation eingebettet, bei der klar ist, dass man Jacksons Video schaut und der Song folgt. Dennoch ist der syntagmatische Aspekt der Verkettung verschiedenartiger Tonmodi und die jeweiligen synchronen Bild-Ton-Relationen, möchte man den konstituierenden Verfahren dieses Filmtextes auf den Grund kommen, nicht zu übergehen. 4b reiht sich dahingehend ein, wenn auch hier modal abermals anders verfahren wird. Der ‚Rap‘ (intoniert durch Price) und das Bildgeschehen werden parallel geführt und ergänzen sich gegenseitig:

Darkness falls across the land
 The midnight hour is close at hand
 Creatures crawl in search of blood
 To terrorise y’awl’s neighbourhood
 And whosoever shall be found
 Without the soul for getting down
 Must stand and face the hounds of hell
 And rot inside a corpse’s shell
 The foulest dench is in the air
 The funk of forty thousand years
 And grizzly ghouls from every tomb
 Are closing in to seal your doom
 And though you fight to stay alive
 Your body starts to shiver
 For no mere mortal can resist
 The evil of the thriller.⁸⁰

Es ist, wie gesagt, immerhin Vincent Price, der hier spricht, ein insbesondere durch seine Darstellungen in Poe- und Lovecraft-Adaptionen zur Ikone des Horrorfilms gewordener Mime, dessen Stimme im allgemeinen kulturellen Wissen der Zeit vor allem für eines steht: den populären Horror der alten, klassischen Schule der 1960er-Jahre. Price hat akustisch wie visuell einen festen wie prominenten Platz im Archiv. Vor Einsatz dieser Zeilen verweilt die Kamera auf einem

⁸⁰ Zitiert nach Childs, „Pop-Video“, S. 47.

Friedhof und verlässt das Paar. So, wie der gesprochene Text eine (u.a. historische) Dimension der dargestellten Welt eröffnet, die über das von Jackson Mitgeteilte hinausreicht, so zeigt sich auf Bildebene die konkrete Manifestation in Form diverser Untoter, die ihre Gräber aufbrechen und verlassen, um ungelenkt in die urbane Welt der Lebenden einzufallen. Das Segment führt damit aufs Anschaulichste vor Augen, worin der emergente Effekt der Hybridisierung besteht: die Überdeterminierung der Horribilität – ‚Thriller‘ in größtmöglich repräsentativer Form, auch, und dies gewissermaßen fortführend, im erneuten Moduswechsel zum ‚Spielfilm‘ samt Rundfahrt um das Paar, Twist bei Michaels Transformation und Vertigo-Effekt zur Unterstreichung des massiven Irritations- und Schockmomentes der jungen Frau, die – auch das entspricht zwar der Logik des Musikvideos, unterläuft aber die Konventionen des Spielfilms – der choreografierten Darbietung beiwohnt, bevor sie die Flucht ergreift, oder auch im kleinen, aber doch im gegebenen Zusammenhang wichtigen Detail, dass während der Performance mit großer musikalischen Untermalung der Atmo-Sound (der Schritte auf dem Asphalt) beigefügt (und eben nicht getilgt) ist. Rekurrent umspielt der Text demnach also die Grenze zwischen beiden Gesamtmodi, umspielt sie bis zur Auflösung in einem gänzlich neuen Format, das er selbst zu repräsentieren beansprucht. Vermehrt lassen sich dahingehend im Syntagma Daten ausmachen, die auf die multimodale Verknüpfung abheben, um auf diese Weise Zeichenmodalitäten zu kombinieren und zu hybridisieren.

Die Paradigmenbildung und damit die Bedeutungskonstituierung im vorliegenden Fall ist demzufolge aufs Engste mit der multimodal-semiotischen Verfasstheit verknüpft. Glückende vs. scheiternde Paarbeziehung, Leben vs. Tod, Mensch vs. Nicht-Mensch, Selbst vs. Selbstentfremdung, ‚Realität‘ vs. Realität – alles dies sind topologische Teilordnungen, mit denen das Video operiert, die es in Zusammenhang bringt und darüber die Problemverhandlung installiert. Dabei ist das Umspielen einer Grenze als maßgebliche Textstrategie zu erachten. Das gilt gleichermaßen für die Ebene des Dargestellten wie für die Ebene der Darstellung. Es deutete sich mehrfach an, dass die genannten Oppositionen einerseits insofern *gesetzt* werden, als das jeweilig oppositionelle Teilsystem stets als ‚Einbruch‘ des Fantastischen inszeniert wird, ausgelöst durch die Figur Michael. Er ist der Grenzgänger, diejenige Figur, die die als gültig angenommene Ordnung infrage stellt bzw. unterminiert und substituiert. Andererseits wird dieser neu eingeführte Systemzustand denn immer auch wieder (zumindest temporär) *zurückgenommen* und *relativiert*, eine ‚alte‘ Ordnung wieder hergestellt, bevor sie aufs Neue gefährdet wird und zusammenbricht. Eine Grenze hat demnach also immer nur losen Bestand und dem Protagonisten ist dies durchaus bewusst, während im Gegensatz dazu in Form der Twist-Strukturen des Videos eine Beständigkeit der Grenzziehung wiederholt re-inszeniert wird und die andere Figur wie auch die Zuschauerschaft in falscher Sicherheit gewogen werden. Dasselbe gilt nun allerdings auch für die Paradigmen ‚Spielfilm‘ und ‚Musikvideo‘, die vom *canvas* abzuleiten sind: Die Kombination setzt die Grenzziehung zwischen beiden, ihre Disjunktheit voraus, ihre Hybridisierung ‚umspielt‘ diese Setzung, indem aus ihr

ein ‚Drittes‘ hervorgeht: Eine Art exponiert ‚spielfilmartiges‘, eben spezifisch fremdreferenzielles, hybrides Medienformat.⁸¹

Der propositionale Gehalt des Ganzen und damit auch die Bedeutungsproduktion erscheinen daher vielschichtig. Zwei Teilpropositionen könnten lauten: *Das Musikvideo ist eine flexible Textsorte (P1), die mit Hilfe verschiedener Formen der System- und Einzeltextreferenz (auf den Horrorfilm), fremdgenerische Elemente semantischer, segmentaler und modaler Art zu transponieren und zu integrieren vermag (P2)*. Und bezogen auf den relevantesten Term: *Michael Jackson ist Dreh- und Angelpunkt der Problemverhandlung, funktionales Bindeglied zwischen Darstellung und Dargestelltem und Kulminationspunkt der referenziellen Bedeutungsdimension (P3)*. P3 ist für sich genommen eine banale Aussage, die neben dem Offensichtlichen – Jackson als Protagonist – im Gesamtkomplex des Gegebenen aber auf einen sinngebenden strukturellen Nexus abhebt, denn die Figur wird schließlich durch den Star selbst verkörpert, um den herum eine reflexive Ebene des Videos angelegt ist: Mit Decker und Krahl der reflexive Entwurf eines Künstler-Image. Eine weitere, damit korrelierte Reflexionsebene ergibt sich aus der Eigenschaft der Figur, Grenzen zu überschreiten und Codes des Horrorfilms zu realisieren. Die Konstitution der Figur, ihre Fähigkeit zur Transformation wird zum Anlass genommen (und durch eine Reihe von Relevanzsignalen angezeigt), dies auch auf Präsentationsebene deutlich zu exponieren und damit Bezugnahmen auf den Horrorfilm konkret zu überführen: Die Verwandlung in Sequenz 2 als Einzeltextreferenz referiert unübersehbar auf Landis' Film *AN AMERICAN WEREWOLF IN LONDON (USA/GB 1981)*. P1 wiederum ist im Grunde genommen ebenso banal und bestätigt ausgehend vom vorliegenden Einzelfall den Befund der Forschung hinsichtlich der Offenheit der Textsorte in Richtung Film und Fernsehen. Anders gewendet jedoch, nämlich im Hinblick auf die filmische Artikulation referenzieller Einheiten, weist sie auf einen hochkomplexen Vorgang hin, vor allem eingedenk von P2 sowie Quantität und Qualität referenzieller Struktur(anordnungen). Denn mag auch generische Offenheit sui generis textsortenkonstitutiv sein, es kommt letzten Endes vor allem darauf an, zu ergründen, wie sie als Möglichkeit konkret genutzt wird – und nach Auseinandersetzung mit unserem Beispiel ließe sich als Ergebnis anbringen: Sie ist mit der multimodalen Anlage des Genres aufs Engste verschränkt. Das heißt, die Transposition fremdgenerischer Strukturen und Elemente verläuft im Einzelfall mannigfaltig und mehrdimensional und bedarf einer gesonderten Analyse, die, so konnte aufgezeigt werden, für eine semiotische Untersuchung unabdingbar ist. Die für den gegebenen Untersuchungszusammenhang entscheidende Proposition ließe sich demnach wie folgt fassen: *Multimodalität im Musikvideoclip weist Funktionen*

⁸¹ Zum Phänomen der Hybridität (als Kombination heterogener Einheiten) ist bekanntermaßen in der Intermedialitätstheorie gearbeitet worden; ich verzichte daher an dieser Stelle auf weitere Erläuterungen (vgl. Irmela Schneider, „Von der Vielsprachigkeit zur ‚Kunst der Hybridation‘. Diskurse des Hybriden“. In: Dies./Christian W. Thomsen (Hgg.), *Hybridkultur. Medien, Netze, Künste*. Köln 1997, S. 13-66, und Yvonne Spielmann, „Intermedialität und Hybridisierung“. In: Roger Lüdeke/Erika Greber (Hgg.), *Intermedium Literatur. Beiträge zu einer Medientheorie der Literaturwissenschaft*. Göttingen 2004, S. 78-102). Zur Untersuchung im Musikvideo wie auch zur einhergehenden Problematik vgl. den Überblick bei Rehbach, „Medienreflexion“, S. 66f.

ersten und zweiten Grades auf (P4). Die Funktionen ersten Grades sind auf basal pragmatischer Ebene angesiedelt: Es geht hierbei um die filmische Medialisierung als kommunikativer Akt, der im gegebenen Fall eben multimodal konstituiert und als Kommunikat entsprechend komplex gestaltet ist. Veranschaulicht werden konnte darüber hinaus aber auch, dass Multimodalität im gegebenen Medienformat eine Art ‚Medialität zweiter Ordnung‘ installiert, indem sie neben ihrer pragmatischen Funktionalität ebenfalls *semantisch* funktionalisiert ist: Denn sie erzeugt nicht nur eine ‚Mehrkanaligkeit‘ des Kommunikationsaktes, sondern ist selbst Bedeutungsträger; je nach ‚Set‘ des Gesamtmodus, verlagert sich die Semantik der Präsentation – mit der Folge, dass das Musikvideo vor allem in seiner multimodalen Anlage sein maßgebliches Konstituens (siehe P1 und 2) codiert und so Hybridisierungsmomente realisiert. Im vorliegenden Fall wird der lexikalische Term ‚Thriller‘ überdeterminiert und andersherum (mit Decker) der thematische Verhandlungskomplex, das heißt die paradigmatische Ordnung und der Umgang mit ihr in Überdeterminierung repräsentiert.⁸²

5. Fazit: Aussichten einer multimodalanalytischen Mediensemiotik des Musikvideos

Für eine medientheoretisch fundierte Semiotik mit Ausrichtung auf die multimodale Verfasstheit von Texten ist in den letzten Jahren viel getan worden. Das Ziel dieses Beitrags war es, für den gesonderten Bereich des Musikvideos einzustehen und Ansätze zur Multimodalität und multimodalen Textanalyse in eine Semiotik des Musikvideoclips zu implementieren.

Wie ein Aufriss der aktuellen Forschungslage zeigte, liegt eine Musikvideosemiotik zwar in elaborierter Form vor, allerdings mangelt es ihr bislang an einer Berücksichtigung der multimodalen Beschaffenheit der Textsorte. Abhilfe sollte die Aufarbeitung kompatibler Vorarbeiten zur Modalität und Multimodalität sowie die Integration eines zweigliedrigen Analyseansatzes in den zeichentheoretischen Theorierahmen schaffen. Wie erörtert, eröffnet der Blick auf Multimodalität zunächst ganz generell eine Sicht auf Zeicheninteraktionsprozesse, das Zusammenspiel der in Bild und Ton kombinierten Einheiten, ihre (zeichenmodale) Grundierung und ‚geschichtete‘ ‚Mehrwert‘-Erzeugung; und damit auch die Perspektive auf Bedeutungsproduktionsoperatoren, d.h. das Zusammenwirken beteiligter Einheiten beim Aufbau von Bedeutung. Der hier vorgeschlagene Ansatz speist sich aus modal-, medien- und filmtheoretischen Vorarbeiten und strebt die Einbindung von vor allem zwei Aspekten multimodaler Semiose in die Analyse ein: Erstens die Bestimmung von multimodaler Spezifik und semiotischen Signifikanzen, die Auseinandersetzung mit Fragen nach der konkreten Gestaltung des *canvas*, danach, welche Modalitäten auf welche Weise zum Einsatz kommen und wie sie auf Basis ihrer spezifischen Verknüpfung Kohärenz stiften. Eine Art modalanalytische Bestandsaufnahme also. Zweitens die Erfassung be-

⁸² Vgl. Decker, „Madonna“, S. 64f.

deutungstragender Einheiten der multimodalen Textebene in mikro- und makrostruktureller Hinsicht, die Verschaltung von Modalität, Multimodalität und Paradigmenbildung, die wiederum in einen propositionsanalytischen Zugriff zu übersetzen ist. Der Zulaufpunkt des Ansatzes besteht demnach nicht allein in der Analyse pragmatischer Leistungen von Zeichenmodalitäten und ihrer Beteiligung innerhalb der thematischen Gesamtkonfiguration, sondern zusätzlich in der Analyse der semantischen Funktion ihrer Interaktion.

Mit THRILLER wurde ein besonders instruktives Beispiel gewählt – sowohl hinsichtlich der Anwendbarkeit des Analysemodells als auch hinsichtlich erwartbarer Ergebnisse einer Musikvideountersuchung. In Auseinandersetzung mit dem Video wurde ersichtlich, wie der vorgeschlagene Ansatz operationalisiert werden könnte, gleichwohl er – und dies sei freilich nochmals angemerkt – nur eine *anteilige* Interpretationsbasis zu bieten hat und selbstredend in größer angelegte Analysen integriert werden müsste. THRILLER jedenfalls kombiniert, das wird bereits bei oberflächlicher Betrachtung deutlich, Struktureinheiten des Musikvideoclips mit solchen des Horrorfilms. Die Frage sollte aber sein, *wie* er das tut. Eine Erkenntnis: durch die verstärkte und relevante Einbindung von Zeichenmodalitäten in den Bedeutungsaufbau. Das Video bietet mehrere Rezeptionsweisen an: An erster Stelle die als Musikvideo mit dominant-reflexivem Umgang mit der Starpersona, dann auch als ‚besonders innovatives‘ Video, das über die ‚Ränder‘ der Singperformance hinaus ein komplexes Narrativ entfaltet und in struktureller Sicht sozusagen ‚mehr‘ als ein bloßes Musikvideo ist, gewissermaßen ein ‚Kurzspielfilm‘. Es kann aber zudem auch als Paradebeispiel einer integrativen Mediegeschichte gelesen werden: Seine – gerade auf zeichenmodaler Ebene – vermittelte Botschaft lautet ja schließlich auch, dass es sich um eine Textsorte handelt, die anschlussfähig und flexibel ist. Wenn die Forschung von einer „hybrid form“⁸³ spricht, von „hybrids“⁸⁴ oder gar von „monstrous hybrids“,⁸⁵ dann müsste infolge der vorliegenden Untersuchung vermerkt werden: Das genuine Gestaltungsprinzip der Hybridisierung ist in erster Linie in Auseinandersetzung mit den modalen Operatoren der Semiose zu decodieren, und eine Interpretation, die die Strategie der Grenznivellierung stark macht, müsste unter Zuhilfenahme entsprechender Ergebnisse unterfüttert werden. Ordnungen werden gesetzt, Grenzen etabliert, haben aber nur temporären Bestand, werden stets wieder durchbrochen, aufgehoben und substituiert – und das gilt für den Umgang mit Ordnungssätzen auf der Ebene des Dargestellten, der *histoire*, aber nun allerdings ebenso für die Ebene der Darstellung, des *discours*, der Präsentationsebene. Das multimodale Wechselspiel zwischen ‚Musikvideo‘ und ‚Spielfilm‘ ist entsprechend gleichbedeutend nicht allein mit mannigfaltigen, sondern zugleich auch oszillierenden Signifikationsprozessen. Wie also zu sehen war, ist multimodale Kommunikation „mehr als die Summe der jeweiligen modalen Aspekte“, be-

⁸³ Childs, „Music Video Gothic“, S. 158.

⁸⁴ Peter G. Christenson/Donald F. Roberts, *It's not only Rock & Roll: Popular Music in the Lives of Adolescents*. Cresskill/New Jersey 1998, S. 141.

⁸⁵ Dick Hebdige, „What is ‚Soul‘?“. In: Alan M. Olson/Christopher Parr/Debra Parr (Hgg.), *Video Icons & Values*. Albany/New York 1991, S. 121-133; hier S. 129.

steht die „eigentümliche Leistungsfähigkeit [dieser] Kommunikation gerade in der bedeutungs- und effizienzsteigernden Kombination unterschiedlicher Modi“. ⁸⁶ Auf einprägsame Weise gilt das für Jacksons Video wie mutmaßlich für den Musikvideoclip insgesamt.

⁸⁶ Bateman/Sachs-Hombach, „Multimodalität“, S. 16.

Filmverzeichnis

- AN AMERICAN WEREWOLF IN LONDON. John Landis (USA/GB 1981).
 BOHEMIAN RHAPSODY. Bruce Gowers (UK 1975).
 LUCKY STAR. Arthur Pierson (USA 1984).
 THRILLER. John Landis (USA 1983).

Literaturverzeichnis

- Bateman, John/Karl-Heinrich Schmidt. *Multimodal Film Analysis. How Films Mean*. London 2012.
- Bateman, John/Klaus Sachs-Hombach. „Multimodalität im Schnittbereich von Medientheorie und Semiotik“. In: *Zeitschrift für Semiotik* 41, 1-2 (2019), 11-36.
- Bench, Harmony. „Monstrous Belonging: Performing ‚Thriller‘ after 9/11“. In: Melissa Blanco Borelli (Hg.). *The Oxford Handbook of Dance and the Popular Screen*. Oxford 2013, 393-411.
- Bódy, Veruschka/Peter Weibel (Hgg.). *Clip Clap Bum. Von der visuellen Musik zum Musikvideo*. Köln 1987.
- Brüssel, Stephan. *Filmisches Erzählen. Typologie und Geschichte*. Berlin/Boston 2014.
- Buckland, Warren (Hg.). *Puzzle Films. Complex Storytelling in Contemporary Cinema*. Malden/Mass. 2009.
- Casetti, Francesco. *Inside the Gaze. The Fiction Film and Its Spectator*. Bloomington 1998.
- Childs, Peter. „Pop-Video: Michael Jackson’s ‚Thriller‘ and ‚Race‘“. In: Ders. *Contemporary Culture Texts and Critical Approaches*. Edinburgh 2006, 40-48.
- Christenson, Peter G./Donald F. Roberts. *It’s not only Rock & Roll: Popular Music in the Lives of Adolescents*. Cresskill/New Jersey 1998.
- Decker, Jan-Oliver. *Madonna: Where’s that girl? Starimage und Erotik im medialen Raum*. Kiel 2005.
- Decker, Jan-Oliver/Hans Krahl. „Videoclip“. In: *Reallexikon der deutschen Literaturwissenschaft*. Bd. 3. Berlin/New York 2007, 781-782.
- Decker, Jan-Oliver. „Strukturalistische Ansätze in der Mediensemiotik“. In: Martin Endres/Leonhard Herrmann (Hgg.). *Strukturalismus heute. Brüche, Spuren, Kontinuitäten*. Stuttgart 2018, 79-105.
- Elsaesser, Thomas/Malte Hagener. „Auge und Blick“. In: Dies. *Filmtheorie zur Einführung*. 3. Aufl. Hamburg 2011, 103-135.
- Frith, Simon u.a. (Hgg.). *Sound and Vision. The Music Video Reader*. London/New York 1993.
- Goodwin, Andrew. *Dancing in the Distraction Factory: Music Television and Popular Culture*. Minneapolis 1992.
- Gräf, Dennis u.a. *Filmsemiotik. Eine Einführung in die Analyse audiovisueller Formate*. Marburg 2017.

- Greimas, Algirdas Julien. *Du sens II*. Paris 1983 (engl. *On Meaning*. Minneapolis 1987).
- Halliday, Michael A. K.. *Language as Social Semiotic. The Social Interpretation of Language and Meaning*. London 1978.
- Hebdige, Dick. „What is ‚Soul‘?“. In: Alan M. Olson/Christopher Parr/Debra Parr (Hgg.). *Video Icons & Values*. Albany/New York 1991, 121-133.
- Holly, Werner. „Medien, Kommunikationsformen, Textsortenfamilien“. In: Stephan Habscheid (Hg.). *Textsorten, Handlungsmuster, Oberflächen. Linguistische Typologien der Kommunikation*. Berlin/New York 2011, 144-163.
- Johnson, Steven. *Everything Bad is Good for You*. London 2006.
- Kaul, Susanne /Jean-Pierre Palmier/Timo Skrandies (Hgg.). *Erzählen im Film. Unzuverlässigkeit, Audiovisualität, Musik*. Bielefeld 2009.
- Kinder, Marsha. „Music Video and the Spectator“. In: *Film Quarterly* (Herbst 1984), 2-15.
- Klemm, Michael. „Ausgangspunkt. Jedem seinem Textbegriff? Textdefinitionen im Vergleich“. In: Ulla Fix u.a. (Hgg.). *Brauchen wir einen neuen Textbegriff? Antworten auf eine Preisfrage*. Frankfurt am Main u.a. 2002, 17-29.
- Krah, Hans. „Das Künstler-Selbst. Referenz und Image im Musikvideo“. In: *Zeitschrift für Semiotik* 42/1-2 (2020), 153-174.
- Kress, Gunther/Theo van Leeuwen. *Multimodal Discourse. The Modes and Media of Contemporary Communication*. London 2001.
- Kress, Gunther/Theo van Leeuwen. „Multimodality and multimodal research“. In: Eric Margolis/Luc Pauwels (Hgg.). *The Sage Handbook of Visual Research Methods*. Los Angeles 2011, 549-569.
- Linsky, Leonard. *Referring*. London 1967.
- Lotman, Jurij M.: *Struktur literarischer Texte*. München 1972.
- Lotman, Jurij M.. *Probleme der Kinoästhetik. Einführung in die Semiotik des Films*. Frankfurt am Main 1977.
- Lynch, Christopher. „Ritual Transformation through Michael Jackson’s Music Video“. In: *Journal of Communication Inquiry* 25/2 (April 2001), 114-131.
- Malsch, Friedemann/Dagmar Streckel. *Künstler-Videos. Entwicklung und Bedeutung*. Ostfildern-Ruit 1996.
- Mercer, Kobena. „Monstrous Metaphors: Notes on Michael Jackson’s *Thriller*“. In: *Welcome to the Jungle: New Positions in Black Cultural Studies*. New York/London 1994, 33-52.
- Menne, Albert. *Einführung in die Methodologie. Elementare allgemeine wissenschaftliche Denkmethode im Überblick*. Darmstadt 1980.
- Mind-Bender: Begriffe – Forschung – Problemfelder (= Paradigma. Studienbeiträge zu Literatur und Film 1 [2017])*. <https://www.uni-muenster.de/Germanistik/ffm/Paradigma/paradigma1/index.html>; Abruf am 21.11.2022).
- Momcilovic, Drago. „Music Video Gothic: Fragmentary Form at the Dawn of MTV“. In: *Gothic Studies* 23/2 (2021), 148-162.
- Norris, Sigrid. *Analyzing Multimodal Interaction. A Methodological Framework*. New York/London 2004.

- Nöth, Winfried (Hg.). *Semiotics of the Media. State of the Art, Projects, and Perspectives*. Berlin/New York 1997.
- Rehbach, Simon. *Medienreflexion im Musikvideo. Das Fernsehen als Gegenstand intermedialer Beobachtung*. Bielefeld 2018.
- Schmid, Wolf. *Elemente der Narratologie*. 3. Aufl. Berlin/Boston 2014.
- Schmitz, Ulrich. „Multimodale Texttypologie“. In: Nina-Maria Klug/Hartmut Stöckl (Hgg.). *Handbuch Sprache im multimodalen Kontext*. Berlin/Boston 2016, 327-347.
- Schneider, Irmela. „Von der Vielsprachigkeit zur ‚Kunst der Hybridation‘. Diskurse des Hybriden“. In: Dies./Christian W. Thomsen (Hgg.). *Hybridkultur. Medien, Netze, Künste*. Köln 1997, 13-66.
- Schneider, Jan Georg/Hartmut Stöckl (Hgg.). *Medientheorien und Multimodalität. Ein TV-Werbespot – Sieben methodische Beschreibungsansätze*. Köln 2011.
- Simon, Josef (Hg.). *Zeichen und Interpretation*. Frankfurt am Main 1994.
- Souriau, Etienne. „La structure de l’univers filmique et le vocabulaire de la filmologie“. In: *Revue internationale de Filmologie* 2/7-8 (1951), 231-240 (dt. „Die Struktur des filmischen Universums und das Vokabular der Filmologie“. In: *Montage/av* 6/2 (1997), 140-157.
- Spielmann, Yvonne. „Intermedialität und Hybridisierung“. In: Roger Lüdeke/Erika Greber (Hgg.). *Intermedium Literatur. Beiträge zu einer Medientheorie der Literaturwissenschaft*. Göttingen 2004, 78-102.
- Stöckl, Hartmut. „Multimodalität – Semiotische und textlinguistische Grundlagen“. In: Nina-Maria Klug/Hartmut Stöckl (Hgg.). *Handbuch Sprache im multimodalen Kontext*. Berlin/Boston 2016, 3-35.
- Strank, Willem. *Handbuch Filmgeschichte. Von den Anfängen bis heute*. München 2021.
- Titzmann, Michael. *Strukturelle Textanalyse. Theorie und Praxis der Interpretation*. 3. Aufl. München 1993.
- Titzmann, Michael. „Semiotische Aspekte der Literaturwissenschaft: Literatursemiotik“. In: Roland Posner/Klaus Robering/Thomas A. Sebeok (Hgg.). *Semiotik. Ein Handbuch zu den zeichentheoretischen Grundlagen von Natur und Kultur*. Berlin 2004, 3028-3103.
- Titzmann, Michael. „Propositionale Analyse – kulturelles Wissen – Interpretation“. In: Hans Krah/Michael Titzmann (Hgg.). *Medien und Kommunikation. Eine interdisziplinäre Einführung*. Passau 2006, 76-92.
- Wulff, Hans-Jürgen. „Die Ordnungen der Bilderflut. Konstellationen medialer Kommunikation als strukturbildendes Prinzip in Performance-Videos“. In: *Rundfunk und Fernsehen* 37/4 (1989), 435-446. <http://www.derwulff.de/files/2-21.pdf>; Abruf am 19.08.2022.
- Wulff, Hans J.. „Semiotik der Filmanalyse. Ein Beitrag zur Methodologie und Kritik filmischer Werkanalyse“. In: *Kodikas/Code* 21 (1998), 19-26. <http://www.derwulff.de/2-84>; Abruf am 06.09.2022).
- Zillmann, Dolf/Norbert Mundorf. „Image Effects in the Appreciation of Rock Video“. In: *Communication Research* 14/3 (1987), 316-334.

