

## Multimodale Semiotik und die theoretischen Grundlagen der Digital Humanities\*

John A. Bateman, Universität Bremen

**Summary.** How the Digital Humanities (DH) will be able to do justice to the diverse disciplines and subjects of investigation they are currently pursuing is a hot topic. Again and again it is suggested that the DH face new challenges that will require new methods, tools and bridges between disciplines. It is less frequently acknowledged that DH are not the only initiative to face such a task. Indeed, models for integration across disciplines on the practical as well as the theoretical level are present in other endeavors as well. In particular, this article illustrates the relevance for DH of results of both a practical and a theoretical nature from *multimodal semiotics*. It is suggested that it is precisely in multimodal semiotics that DH could find a foundation sufficiently broad and comprehensive to advance many of the fundamental discussions currently underway.

**Zusammenfassung.** Wie die Digital Humanities (DH) ihren diversen Disziplinen und Untersuchungsgegenständen gerecht werden können, wird heiß debattiert. Dabei wird immer wieder betont, dass die DH vor neuen Herausforderungen stehen, die neue Methoden, Werkzeuge und Disziplin-übergreifendes Vorgehen erfordern. Seltener wird offengelegt, dass die DH nicht die einzige Unternehmung sind, die sich mit Aufgaben dieser Art konfrontiert sieht. In anderen Bereichen sind Modelle für eine solche Integration verschiedener Disziplinen auf praktischer sowie theoretischer Ebene durchaus zumindest ansatzweise vorhanden. In diesem Aufsatz wird die Relevanz von praktischen sowie theoretischen Ergebnissen aus der *multimodalen Semiotik* für aktuelle Aufgaben der DH diskutiert. Es wird vorgeschlagen, dass es genau die multimodale Semiotik ist, in der die DH ein ausreichend breites und umfassendes Fundament finden könnten, das für viele aktuelle Grundlagendiskussionen hilfreich wäre.

### 1. Einführung: Auf der Suche nach Grundlagen

Die ‚Digital Humanities‘ (üblicherweise mit DH abgekürzt) sind sicherlich *en vogue*. Zu ihnen gehört eine Reihe ganz unterschiedlicher Aktivitäten:

von der kunstvollen Gestaltung neuartiger Interaktionsmöglichkeiten und Darstellungsformen kultureller Daten und Praktiken aller Art bis hin zu Versuchen, neue wissenschaftliche Erkenntnisse soziokultureller Provenienz aus großen Datenmengen zu gewinnen. Für die erstgenannte Richtung spielen Design und Informationsgestaltung eine zentrale Rolle. Hier verspricht der Einsatz von immer vielfältigeren Mensch-Maschine-Schnittstellen neue Impulse für weiterführende Untersuchung des präsentierten Stoffs. Ergebnisse können hier positiv bewertet werden, wenn tatsächlich neue Impulse und Betrachtungsweisen gewonnen werden. Die zweite Richtung befindet sich in einer schwierigeren Situation. Es geht hier im Wesentlichen darum, was für Beziehungen zwischen quantitativen und qualitativen Ansätzen möglich sind. Besonders hier sind die DH heute mit großen Ansprüchen unterwegs, die oft nichts weniger implizieren als eine Neukonzipierung der Möglichkeiten von geistiger Arbeit überhaupt und infolgedessen eine Neuverhandlung der geistigen Grenzen zwischen Mensch und Maschine (vgl. Berry 2011; Evans und Rees 2012: 21). Um eine derartige Transformation und Weiterführung des ‚Kerngeschäfts‘ der Geisteswissenschaften vorzunehmen, wäre eine dementsprechend grundlegende Erneuerung in Theorie und Methoden notwendig. Aber wie die bereits umfangreiche Literatur bezüglich möglicher Definitionen des ‚Feldes‘ zeigt (vgl. McCarty 2005; Svensson 2009; Berry 2012; Burdick u.a. 2012; Liu 2013; Gold und Klein 2016), bleibt viel Grundlegendes noch unklar. Hauptziel dieses Aufsatzes wird es daher sein, einige Überlegungen zu den nötigen Bestandteilen eines angemessenen Theoriegebäudes zu liefern, das den DH der zweiten Art ein strapazierfähigeres Fundament verleihen kann.

Entscheidungen bezüglich eines solchen Fundaments müssen in den DH selbst getroffen werden; häufig wird anstelle dessen berichtet, wie sich Projekte, Professuren, Kompetenzzentren und dergleichen rasant ausbreiten. Ein Kennzeichen dieses Zustands ist die wiederkehrende Bemerkung, dass die DH „mehr noch als andere Disziplinen projektgetrieben“ (Sahle 2015) seien. Einerseits ist eine in der Tat beeindruckende Vielfalt von Projekten, die sich als DH bezeichnen, in der internationalen Forschungsszene sowie in der allgemeinen Öffentlichkeit präsent; auch die dafür benötigten Forschungsgelder und institutionelle Unterstützung wachsen schnell im Umfang. Andererseits fallen bisher gelieferte Antworten auf Fragen der Definition des Feldes relativ dünn aus. Projektbezogene Arbeit ist sicherlich nichts Schlechtes. Problematisch wird es dann, wenn Mängel im theoretischen Überbau die Bestimmung von gemeinsamen Forschungsaufgaben erschweren, die mehr als eine bloße ‚Digitalisierung‘ mit anschließender Visualisierung sind.

Wie Sahle es treffend beschreibt, trifft dies besonders für den Kern des großen DH-Vorhabens zu:

Auf dieser Grundlage besteht dann aber auch die Gefahr, dass es trotz des hohen Mitteleinsatzes nicht zu einer dauerhaften Stärkung und einem Ausbau der Digital Humanities in Deutschland kommt, sondern aus unverbundenen Teilen bestehen-

de Kooperationsprojekte jenseits ihrer lokalen Problemlösung keine nachhaltigen Effekte auf die DH insgesamt haben werden. Die Betonung der autonomen Ränder hat dann jedenfalls desintegrative Effekte auf den doch eigentlich adressierten Kernbereich und fördert eher zentrifugale als zentripetale und integrative Tendenzen (Sahle 2015).

Ein gewisser Zweifel bezüglich der recht hohen Ansprüche der DH der zweiten oben identifizierten Richtung ist daher sicherlich gerechtfertigt: Ist aus dem quantitativen Zuwachs von verwendeten Datenmengen wirklich ein qualitativer Sprung in unserem Verständnis der Phänomene zu erwarten? Wie kann ein solches Verständnis aus mehr oder weniger mechanischem Mustererkennen entstehen? Menschliche ‚Eigenarbeit‘ ist hier doch wieder notwendig, da Visualisierungen, wie häufig als wesentliche Ergebnisse der Auswertung größerer Datenmengen präsentiert, ohne weitere Interpretation oder Diskussion inhaltlich leer bleiben (vgl. Gitelman 2013; Drucker 2017: 121). Aber dann unterscheiden sich die DH von anderen modernen Untersuchungsbereichen, bei denen computergestützte Methoden angewendet werden, nicht. Wo finden sich die Prinzipien, die einen *g e i s t e s w i s s e n s c h a f t l i c h e n* Ertrag theoretisch sowie methodologisch plausibel machen?

Dass die DH als kein bloßes „kooperatives Nebeneinander von Geisteswissenschaften und Informatik“ zu verstehen sind (Sahle 2015), scheinen alle in den DH für selbstverständlich zu halten. Die sich selbst mit den DH identifizierenden Arbeiten beanspruchen meist, dass ihr Einsatz von digitalen – sprich maschinellen, d.h. mit Hilfe von Computern durchgeführten – Methoden dem Zweck folgt, die etablierten Ziele der Geisteswissenschaften, vor allem Interpretation und Erklärung, zu unterstützen und zu erweitern. Und in der Tat ist es durch die fortschreitende Digitalisierung von Medien jeder Art und die damit technisch bedingte ‚Konvergenz‘ von Medien (Jenkins 2008; Grant und Wilkinson 2009) immer leichter geworden, Fragestellungen aus früheren Zeiten der ‚literary computing‘ oder ‚humanities computing‘ mit immer größeren Mengen von Daten zu konfrontieren und für andere Formen von Daten (obwohl meist immer noch auf die visuellen beschränkt) zu übernehmen. In der Literaturwissenschaft sind zum Beispiel Fragen zur Autorschaft von Werken oder nach gemeinsamen Eigenschaften von Genres, Epochen, Erzählungen usw., die sich durch stilistisch relevante Metriken (d.h. stilometrisch) identifizieren lassen (vgl. Meister 2007), schon lange gestellt worden. In der Kunstgeschichte bietet der direkte Zugriff auf Bildmotive neue Wege für die ikonographische Erkennung von Stilentwicklungen, die Zuordnung von Kompositionselementen, die Etablierung von Traditionen usw. (vgl. Kwastek 2015; Warnke und Dieckmann 2016). Und die rasanten Verbesserungen in der Zugänglichkeit und Benutzerfreundlichkeit von automatischen Visualisierungsmethoden machen Darstellungen der Ergebnisse solcher Datenanalysen verführerisch plastisch und unmittelbar ‚greifbar‘. Dabei wird in den DH erhofft, dass sich die Komplexitäten und Nuancen hochgradig dynamischer kultureller Prozesse damit zum ersten Mal erschließen lassen. Begriffe und Methoden zu diesem Ziel,

wie etwa ‚distant reading‘ (Moretti 2013), ‚Cultural Analytics‘ (Manovich 2012), usw., gelangen immer mehr in Umlauf. Manovich (2012, 2018) argumentiert sogar, dass kulturelle Produkte wie Facebook, Blogs, Twitter und ähnliche keine alternativen Methoden zulassen, einfach der Masse wegen.

Die Möglichkeit, größere Datenmengen ‚gleichzeitig‘ betrachten zu können, erlaubt sicherlich die Fortsetzung und den Ausbau einiger traditioneller geisteswissenschaftlicher Ziele. Inwiefern dies tatsächlich etwas Neues für die Geisteswissenschaften darstellt, bleibt aber umstritten. In sich ist diese Entwicklung ein logisches und fast selbstverständliches Ergebnis des ‚computational turn‘, das kaum den DH zuzuschreiben ist (vgl. Hall 2013). Eine Reihe neuer Fragen werden durch die Möglichkeiten, große Datenmengen zu untersuchen, aufgeworfen. Jedoch betreffen viele dieser Fragen eher die soziogesellschaftlichen Prozesse der Mediennutzung als tiefergehende Interpretation von kulturellen Artefakten und Praxen, worum es im herkömmlichen Sinne den Geisteswissenschaften eigentlich geht. Kurz gesagt: technische Entwicklungen allein werden nicht ausreichen, um die anspruchsvollen Ziele der DH zu verwirklichen.

Die beiden Begriffe ‚digital‘ und ‚humanities‘ tragen unmittelbar zu diesem etwas konfusen Zustand bei. Wie häufig bei Komposita, ist die eigentliche Bedeutung ihrer Zusammensetzung kaum selbsterklärend; in diesem Falle ist sogar ein gewisser Widerspruch enthalten. Die ‚humanities‘, hier als die Geisteswissenschaften verstanden, haben traditionell recht wenig mit den Ansprüchen des ‚Digitalen‘ zu tun und umgekehrt. Bei dem Einen wird die geistige Leistung des menschlichen Interpretationsapparats hoch gepriesen und als unabdingbar für das Verständnis des Menschlichen deklariert; bei dem Anderen wird gerade die Unabhängigkeit von solchen subjektiven Vorstellungen als höchster Wert erhoben. Die überwiegend hermeneutisch ausgelegten Traditionen der Geisteswissenschaften sind nicht leicht mit dem algorithmisierten Blick der digitalen Technologie zu vereinbaren und Aufklärungsarbeit ist daher dringend gefragt.

Egal ob die DH sich als ein einheitliches Feld oder eher als ‚eine Matrix von konvergierenden Praxen‘ (Schnapp und Presner 2009) sehen, setzen die hermeneutische Ausrichtung der Geisteswissenschaften und die dem entsprechende Auffassung von DH einige Grundprinzipien der Interpretation und Analyse voraus. Dann ist es verwunderlich, wie in den DH einerseits die Möglichkeit und Notwendigkeit von *i n t e r p r e t a t i v e n* Verfahren für humanistische Überlegungen als zentral angesehen werden, andererseits aber die DH selten von *m a s c h i n e l l e n* – sprich ‚digitalen‘ – Möglichkeiten für Interpretation durch automatisches Schussfolgern Gebrauch machen. Wenn dies aber nicht auch geschieht, können die DH kaum als ein neuer Beitrag zu den Geisteswissenschaften im Allgemeinen angesehen werden, sondern beschränken sich darauf, Visualisierungsmethodik zu sein. Die tatsächlichen Interpretationen bleiben im traditionellen Rahmen. Daher muss gefragt werden, wie eine wirkliche Interpretationsleistung zu schaffen wäre. Und hier stoßen viele aktuelle Diskussionen an ihre Grenzen, weil ihnen die nötigen Grundlagen fehlen.

In diesem Sinne merken Bubenhofer und Scharloth an:

Um für kultur- und sozialwissenschaftliche Fragestellungen aber als Methode attraktiv zu sein, muss sich die maschinelle Textanalyse stärker um eine gegenstandsadäquate Modellierung des Textbegriffs bemühen und sich hierbei an Theorien der Textlinguistik orientieren; sie braucht einen integrierten Textbegriff (Bubenhofer und Scharloth 2015: 14).

Dies ist sicherlich richtig, aber für die DH im Allgemeinen greift es noch zu kurz. Durch die Textlinguistik behandelte sprachliche ‚Texte‘ bilden heute einen immer kleiner werdenden Teil der DH, weil das ‚Digitale‘ der DH schnell in Verbindung mit ‚neuen Medien‘ und der ‚digital art‘ gebracht wird. Dabei wird eine aus der Literaturwissenschaft vererbte Fokussierung auf ‚Text‘ im Sinne der natürlichen menschlichen Sprache (und meistens sogar nur der geschriebenen) besonders radikal aufgehoben. Auch innerhalb des Bereichs des Kulturerbes ist die Archivierung und Dokumentation von nicht primär textuellen Artefakten bereits längst etabliert. Primär visuell ausgerichtete Medien, wie sie in der Kunstgeschichte im Vordergrund stehen, sowie Mischungen aller Art (Text, Bild, Ton, usw.) sind auch ins Blickfeld gerückt. Die moderneren DH beschäftigen sich demzufolge nicht nur mit Texten traditioneller Art sondern mit Karten, Kunstwerken, Fotografien, Gebäuden, Filmen sowie visuellen Narrativen usw. Um dieser Datenlage gerecht zu werden, ist ein tieferliegendes Verständnis von Kommunikationsformen aller Art unabhängig von ihrer digitalen Darstellung unabdingbar. Dies ist nicht aus dem aktuellen Disziplinenrepertoire der DH abzuleiten. Clement merkt diese Lage in den DH besonders zutreffend sowie kritisch an:

Indeed, while the studies mentioned demonstrate the expanding scope of objects to read, the hermeneutical methods associated with reading remain largely unarticulated (Clement 2016: 162).

Statt nach einer theoretischen Grundlage zu suchen, wird bei vielen Arbeiten darauf gehofft, dass sich diese Problematik durch die Betrachtung von immer größeren Datenmengen auflösen wird. Dabei gibt es durchaus neue Techniken für ‚big data analysis‘, insbesondere das sogenannten ‚Deep Learning‘-Verfahren, die in der Tat wesentlich neue Möglichkeiten (vgl. Zhang u.a. 2018) bieten. Bisher ist aber unklar, wie diese Möglichkeiten sich zu der für die Geisteswissenschaften zentralen interpretativen Auseinandersetzung mit kulturellen Artefakten und Darbietungen verhalten. Deep-Learning-Verfahren erzielen bereits beeindruckende Ergebnisse bei der automatischen semantischen Klassifizierung von Daten; dabei müssen allerdings entweder bereits ausgearbeitete Klassifikationsschemata vorhanden sein, um ausreichende ‚Training Data‘ herzustellen, oder man verzichtet weitgehend auf inhaltlich interpretierbare Klassifikationskategorien: das heißt, eine Eingabe X ist zuverlässig mit einer Ausgabe Y assoziiert

worden, aber die Gründe für die Assoziation bleiben in der internen Organisation der gelernten Netzwerke verborgen. Avarado und Humphreys (2017: 740) kontrastieren zwischen ‚opaken‘ und ‚transparenten‘ Repräsentationen in einem ähnlichen Sinne. Opake Verfahren bilden dann gewissermaßen einen *Gegensatz* zu den Interpretationen der Geisteswissenschaften insofern, dass *Erklärbarkeit* keine Rolle mehr spielt: Man weiß bei einer Bild-Analyse vielleicht, dass ein Ohr und ein Auge in dem Bild vorkommen, weiß aber nicht, warum und wie diese Ergebnisse zustande gekommen sind, und noch weniger, ob diese beiden Ergebnisse etwas miteinander zu tun haben. Alles ist im Prinzip trainierbar, aber das Verfahren ist an bereits analysierte Datenmengen gebunden. Das *prinzipielle* Problem der Analyse und Interpretation ist damit nicht gelöst.

Um den Herausforderungen der hermeneutischen Interpretation verschiedenster Ausdrucksformen gewachsen zu sein, muss ein allgemeinerer als der herkömmliche ‚Text‘-Begriff angewendet werden. Hierzu schreibt bereits Hayles (2003), dass ein gewisses ‚Neudenken‘ von Textualität in den DH notwendig sei. Mithilfe prominenter literarischer Beispiele wie den bekannten Text-Bild-Kombinationen in den Werken von William Blake zeigt Hayles die problematische Grenze von Materie und Text auf. Diese Grenze bekommt heute eine besondere Brisanz z.B. wegen der raschen Zunahme künstlerischer Werke, die ganz explizit mit diesen Grenzen spielen (für viele Beispiele und Diskussion, siehe z.B. Gibbons 2011), und wegen den Herausforderungen, mit denen sich die Editionswissenschaft bei neuartigen *Digitalen Editionen* konfrontiert sieht (siehe Nantke in diesem Heft). Hayles' Verweis auf die Textualität eröffnet aber den konkreten Weg, der im weiteren Verlauf dieses Artikels besprochen wird. Umfangreiche praktische und theoretische Untersuchungen zu einer neuen Textualität, die sich über alle Ausdrucksformen und Medien erstreckt, sind in neueren Arbeiten in der Semiotik, und insbesondere in der Semiotik der Multimodalität, vorgenommen worden. Eine konsequente Anwendung von Begrifflichkeiten aus dieser Richtung der neueren Semiotik wird uns nun ermöglichen, wesentliche Fortschritte bei der für die DH erforderlichen Aufklärungsarbeit zu erzielen.

Eine solche Anwendung der Semiotik soll die Aktivitäten der DH nicht ersetzen oder mit ihr konkurrieren, sondern sie soll dazu dienen, die DH besser zu untermauern und die vielen Querverbindungen und Anschlussmöglichkeiten über Disziplinen, Medien, Theorien, Modelle und Methoden hinweg auf eine sichere Basis zu stellen. Außerdem gibt es in der multimodalen Semiotik bereits wesentliche Arbeiten, die Verbindungsmöglichkeiten und Gemeinsamkeiten mit den DH aufweisen. Einige dieser Arbeiten werden unten exemplarisch skizziert. Ohne eine solche semiotische Grundlage wird es unklar bleiben, ob die DH jemals mehr als eine allgemeine, wenn auch (noch) strategisch gut einsetzbare, Etikette für die Verwendung von computergestützten Ansätzen sein können – Ansätze, die heutzutage sowieso von den meisten Disziplinen verwendet werden und so kein Alleinstellungsmerkmal bilden.

Unsere erste Station auf dem Weg zu einer allgemeinen multimodalen Semiotik, die unten auch als Grundlage für die DH vorgeschlagen wird, beschäftigt sich mit der sogenannten **Multimodalen Linguistik**. In ihr findet man genau einen solchen erweiterten Begriff von ‚Textualität‘, der für alle Gegenstandsbereiche der DH eingesetzt werden könnte, um einen stabilen und trotzdem erweiterbaren Organisationsrahmen für Analysen bereitzustellen. Im folgenden Abschnitt wird dies erläutert werden. Insbesondere wird gezeigt, dass wir es in der multimodalen Linguistik nicht mehr ‚nur‘ mit Texten oder Bildern zu tun haben, sondern immer mit Kombinationen derselben – Kombinationen, die sogar häufig prä-theoretische Termini wie ‚Text‘ und ‚Bild‘ in Frage stellen und überholen. Dies kann dann auch ein wichtiger Schritt für die DH sein, die dringend ‚jenseits‘ des Textes operieren müssen, aber deren Methoden immer noch meistens textuell (im herkömmlichen Sinne) angelegt sind (vgl. Clement 2016: 162).

## 2. Multimodale Linguistik

Der Terminus „multimodale Linguistik“ hat seinen Ursprung in den frühen 2000er Jahren als Sammelbegriff für eine wachsende Zahl von Forschungsarbeiten, die begonnen haben, bestimmte Prinzipien und Methoden der Linguistik auf ein weitaus breiteres Spektrum von Medien anzuwenden (vgl. Bateman 2008: 38–39). Diese Forschungsrichtung hat sich Mitte der 1980er Jahre aus mehreren voneinander unabhängigen Forschungsfeldern herausgebildet, als es deutlich wurde, dass die bisherige Fokussierung auf Sprache in der Linguistik Wesentliches außer Acht ließ. Zum Beispiel war der Versuch, Werbung textlinguistisch, aber ohne die Mitbetrachtung des Bildlichen, der Typographie, der Komposition u.ä. zu analysieren, von vornherein zum Scheitern verurteilt. Um das Ziel einer zunehmenden Anwendbarkeit der Linguistik auf reale Texte zu verwirklichen, wurde es notwendig, dass Sprache in Verbindung mit anderen Ausdrucksformen wie Bild, Intonation, Gestik usw. betrachtet werden musste. Dadurch entstand aus der Textlinguistik (vgl. Klopfer 1977; Spillner 1982; Stöckl 1992), der Soziofunktionalen Linguistik (vgl. Kress und van Leeuwen 1990; O’Toole 2011 [1994]), der Gesprächs- und Interaktionsforschung (vgl. Kendon 1980; McNeill 1992) u.a. ein heterogenes Forschungsfeld mit dem Hauptziel, die effektive Wirkung von Kombinationen von Ausdrucksformen zu untersuchen und zu verstehen.

Während dieses Thema schon seit einiger Zeit in den DH als wichtig angesehen wird, bildet es für Theorien der Multimodalität das zentrale Anliegen überhaupt. Es ist dann nicht weiter verwunderlich, dass im Gegensatz zu den DH in der multimodalen Linguistik bereits eine Reihe differenzierter theoretischer Grundlagen vorliegen, die die Mechanismen der Multimodalität beleuchten. Darüber hinaus sind diese Mechanismen in jüngerer Zeit immer mehr mit empirischen Methoden – sowohl experimentell als auch korpus-basiert – verbunden worden und liegen daher noch näher an den Zielen der DH. Zu den Anwendungsgebieten dieser Multimodalitäts-

theorien gehören unter anderem: gesprochene Interaktion von Angesicht zu Angesicht (z.B. Müller 1998; Mittelberg 2006; Fricke 2012; Deppermann 2013; Mondada 2014), Bild-Text-Beziehungen (z.B. Stöckl 2006; Martinec und Salway 2005), Film (z.B. van Leeuwen 1991; Bateman 2007; Tseng 2013; Wildfeuer 2013), Comics und Graphic Novels (z.B. Cohn 2013; Bateman und Wildfeuer 2014), Musik und Ton (z.B. van Leeuwen 1999), Museenausstellungen und Architektur (vgl. Ravelli 2006; Ravelli und McMurtrie 2016) sowie Bemühungen um die Lese- und Schreibfähigkeit sowie Fragen zur Pädagogik (New London Group 2000; Jewitt und Kress 2003; Lee und Khadka 2018). Die Relevanz der Theorie sowie der Praxis von multimodaler Linguistik sollte daher für viele Fragen der DH evident sein; eine ausführliche Einführung in den aktuellen Stand der Multimodalitätsforschung ist in Bateman, Wildfeuer und Hiippala (2017b) zu finden.

Von besonderem Interesse für die aktuelle Diskussion ist die Verbindung zwischen Modellbildung und Empirie. Dieser Verbindung wird in der multimodalen Linguistik eine immer wichtiger werdende Rolle zugewiesen. Dabei ist die Auseinandersetzung mit der inhärenten Wechselwirkung von qualitativen und quantitativen Ansätzen fundamental: Modellbildung ist qualitativ, die Untersuchung und Evaluierung der Modelle häufig quantitativ. In der Linguistik ist diese Kombination längst üblich; und in der multimodalen Linguistik wird diese gleiche Methodik für verschiedenste Artefakte und alle möglichen multimodalen Kommunikationssituationen auch eingesetzt. Kurz gesagt: Während die aktuelle Praxis der DH noch überwiegend der Visualisierung von Datenmengen verhaftet bleibt (Fragen nach der Bedeutung dieser Visualisierungen schließen sich an, werden aber zu oft noch nachrangig behandelt), liegt der Schwerpunkt in der multimodalen Linguistik heute bereits auf der qualitativen Auswertung von Daten auf der Grundlage von erklärenden Modellen. Dieser Unterschied in der Fokussierung ist natürlich kein prinzipieller: beide Orientierungen könnten, und sollten, sich eigentlich gegenseitig ausgezeichnet ergänzen. Einerseits ist der Entwurf von neuartigen Visualisierungen aus guten semiotischen Gründen, wie wir unten erfahren werden, unabdingbar für die Ausarbeitung und Konzipierung neuer Ideen; andererseits müssen aber diese Ideen in Modelle, Formalisierungen und Theorien münden, um neue und allgemein gültige Erkenntnisse jenseits einzelner Projekte zu erzielen.

Im nächsten Abschnitt wird zuerst der hier angenommene theoretische Rahmen kurz skizziert und danach werden einige Beispiele diskutiert werden, die verdeutlichen, wie die allgemeine Herausforderung der Überbrückung der quantitativen-qualitativen ‚Lücke‘ im Bereich der Multimodalität behandelt wird.

## 2.1 Die Modi der Multimodalität

Zentral für ein Verständnis der multimodalen Linguistik und der Multimodalität an sich ist die Definition des Begriffes ‚semiotischer Modus‘ (engl. *semi-*

*otic mode*). Das heißt, um Multimodalität zu verstehen, ist es ratsam genauer zu spezifizieren, worauf genau das ‚Multi‘ der ‚Multimodalität‘ sich bezieht. Semiotischer Modus wird hier anstelle des in der Semiotik traditionelleren Terminus ‚Kode‘ (engl. *code*) verwendet. Wie bei vielen der Termini, mit denen man in diesem Bereich zu tun hat, ist die Bedeutung des Begriffs Kode recht breit gestreut und weckt eine Reihe eher unpassender Assoziationen, was einerseits den Austausch zwischen unterschiedlichen Bereichen der Semiotik erschwert und andererseits den Blick von der seit Charles Sanders Peirce zentralen *prozessualen* Auffassung der Semiose weggeführt hat.

Der Verwendung des Kode-Terminus für nicht sprachliche Kommunikationssysteme ist auf Roman Jakobsons Betrachtung der Semiotik in den 1950er Jahren zurückzuführen (Jakobson 1971; Eco 1981). Jakobson versucht konsequent ein Untersuchungsfeld zu konzipieren, das sich mit mehr als nur der Sprache befassen soll. Saussure und seine Nachfolger hatten (zu Recht) die unabdingbare Rolle von Konvention für die Untersuchung von Sprache betont; in Jakobsons Erweiterung der Betrachtung wurde es dann notwendig, unterschiedliche *Systeme* von Konventionen voneinander abzugrenzen. Dafür wurde der Terminus ‚Kode‘ geprägt. Kode wurde dann schnell zu einem zentralen Begriff der Semiotik. Heute ist die Situation bezüglich unterschiedlicher Fassungen des Kode-Begriffs recht komplex (vgl. Nöth 2000: 218–219). Es liegt eine Reihe definitorischer Versuche vor, die allerdings meist entweder vage im Umfang und Inhalt bleiben (der ‚Kode‘ des Theaters, der Musik usw.) oder zu einfach sind im Sinne einer über beschränktem Zeichenrepertoire operierenden Kryptologie. Häufig wird Kode als ein ‚Regelsystem für die Verwendung von Zeichen‘ verstanden, wobei die genaue Art weder der ‚Regeln‘ noch der ‚Zeichen‘ geklärt ist. Schließlich wird der Begriff ‚Kode‘ für fast jede mögliche Menge von Konventionen verwendet, die eine Rolle während der Interpretation spielen könnten. Gerade diese Fokussierung auf Konventionen und Regelwerke ist dann selbst der Kritik anheim gefallen, insbesondere wenn es sich um angeblich ‚freiere‘ Ausdrucksformen wie Bild, Poesie usw. handelt. Hier wurde häufig ein sprachzentrisches Verständnis von Zeichenprozessen unterstellt (vgl. Boehm 2004), das zumindest für visuelle Kommunikation unzumutbar scheint. Die schwer auszublenkende Konnotation von Kode in Richtung ‚Verschlüsselung‘, ‚Entschlüsselung‘ und einer mehr oder weniger direkten Verbindung zwischen Signifikant und Signifikat führt immer wieder zum (Vor-)Urteil, dass Kode unausweichlich nach dem Modell der menschlichen Sprache zu verstehen sei und daher kein angemessenes Instrumentarium für nicht-sprachliche Ausdrucksformen bereitstellen kann.

In der sozialsemiotischen Multimodalitätsforschung nach Kress und van Leeuwen (2006 [1996]) wird aus diesem Grund der Terminus ‚mode‘ präferiert. Dabei wird betont, dass multimodale Bedeutung nicht nur viel breiter als die Bedeutung von natürlicher Sprache zu fassen ist – Sprache ist dann nur eine der Ausdrucksmöglichkeiten unter vielen gleichberechtigten anderen –, sondern auch immer als ein dynamisches, kreatives Verfahren verstanden werden muss, nie als ein statisches ‚Entziffern‘ von Zeichen.

Obwohl dieser Eckpfeiler einer Theorie der Semiotik auch von denen, die den Terminus ‚Kode‘ verwenden, kaum bestritten wird, haben sich außerhalb der Semiotik in den letzten 50 Jahren dermaßen verhärtete Fronten gebildet, dass die bloße Behauptung, dass ‚Kode‘ auch semiotisch nicht als statische Entschlüsselung zu verstehen sei, kaum weiter hilft. Falsche Aussagen, wie folgende von Wharton (sogar in einem Absatz zu Peirce) sind leider all zu verbreitet:

semiotic accounts rely entirely on a code model of communication. However, a coding-decoding model is as inappropriate for analysing the intention-driven communicative interaction of humans as it is appropriate for analysing the dances of honeybees (Wharton 2009: 121).

Auf einer solchen Grundlage können produktive Umsetzungen der Prinzipien der Semiotik nicht entstehen. Deshalb werden wir hier den Begriff ‚Mode‘ verwenden, um die mit ‚Kode‘ verbundenen Assoziationen zu vermeiden.

Aber auch die Multimodalitätsforschung hat es schwierig gefunden, einen Konsens über eine einheitliche Definition von ‚semiotischen Modus‘ sowie ‚Multimodalität‘ zu erreichen. Unter den detailliertesten Darstellungen des Standes der Dinge sind die von Stöckl (2004), der von (schriftlichen) Sprache-Bild-Relationen ausgeht, und von Fricke (2012: 36–76), die bei Kombinationen von gesprochener Sprache und redebegleitender Gestik beginnt. Kress und van Leeuwen (z.B. 2006 [1996]) und ihre darauf aufbauenden Arbeiten behandeln zwar ein viel breiteres Spektrum von Phänomenen, das sich von Musik bis hin zu Architektur erstreckt, aber ihr Begriff von semiotischem Modus wird genau deswegen besonders unscharf. Im wesentlichen ist ein semiotischer Modus nach Kress und van Leeuwen nicht mehr als eine beliebige ‚semiotische Ressource‘, die von einer Benutzergemeinschaft verwendet wird, um Bedeutungen zu schaffen und zu kommunizieren (z.B. O’Halloran 1999; Bucher 2011; Kress 2014). Ein wesentlicher Unterschied zum ‚semiotischen Kode‘ ist auf dieser Basis kaum erkennbar und die daraus folgende Unschärfe ist bereits von mehreren Autoren kritisiert worden (z.B. Forceville 1999, 2007).

Als Reaktion auf diesen anhaltenden Mangel an Übereinstimmung hat Bateman (2011) eine neue Konzeption vom semiotischen Modus eingeführt, die deutlich über die bisherigen Beschreibungen hinausgeht, deren Stärken kombiniert und deren Schwächen überwindet. Im Gegensatz zum herkömmlichen Gebrauch wird semiotischer Modus wie folgt definiert:

**D e f i n i t i o n :** Ein ‚semiotischer Modus‘ ist eine sich aus dem Gebrauch innerhalb einer Nutzergemeinschaft bildende Praxis, die Kommunikation und Bedeutungskonstitution ermöglicht. Dies können wir mit den folgenden drei abstrakten Ebenen beschreiben:

- i. eine verformbare wahrnehmbare *M a t e r i e*,
- ii. eine Klassifikation (paradigmatisch) von *F o r m*einheiten und -strukturen (syntagmatisch), die die für den semiotischen Modus als pertinent anzusehenden materiellen Verformungen definiert, und

iii. eine *Diskursebene*, die dynamische Mechanismen für die abduktive Schlussfolgerung kontextbezogener Interpretationen von verwendeten Formklassifikationen bereitstellt.

Alle drei Facetten sind notwendig für die Bestimmung eines jeden einzelnen semiotischen Modus. Die vorausgesetzte enge Verbindung, die zwischen materiellen Formen und Diskursinterpretationen gezogen wird, bietet eine Grundlage für eine weit präzisere Hypothesenbildung für empirische Forschung als sie in der breiten Multimodalitätsforschung vorher möglich gewesen war. Weitere Einführungen und Beispiele sind in Bateman (2016) und Bateman u.a. (2017b) zu finden.

Diese verallgemeinerte Auffassung von semiotischen Modi erlaubt es, viele Positionen aus der früheren Multimodalitätsdiskussion zu vereinheitlichen und sinnvoll in Bezug zueinander zu setzen. Die oben erwähnten Auseinandersetzungen mit dem Kode- bzw. Mode-Begriff von Fricke (2012) und Stöckl (2004) sind zum Beispiel folgendermaßen neu zu formulieren. Der Hauptunterschied zwischen den beiden liegt in ihrer jeweiligen Ausgangsposition: in der gesprochenen Sprache und Interaktion (bei Fricke) bzw. in der schriftlichen/visuellen Kommunikation (bei Stöckl). Für Fricke ist es daher naheliegend, Fälle von Kommunikation auf eine Skala zu legen, die die Stärke des Einflusses von anderen Kodes (wie die der Gestik oder des Bildes) auf die des ‚sprachlichen‘ Kodes widerspiegelt. Wenn die involvierten Kodes beieinander stehen ohne interne Änderungen oder Anpassungen (z.B. ein *hörbares* Schreiben von einem auch gleichzeitig vorgelesenen Text in einem Theaterstück), schlägt Fricke ‚sprachliche Multimodalität‘ vor (Fricke 2012: 48). Multimodalität ‚im weiteren Sinne‘ (Fricke 2012: 47–48) bezieht sich auf Fälle, bei denen nur eine Sinnesmodalität involviert ist, aber Kodeintegration tatsächlich gerade vorliegt. Ohne Kodeintegration wäre es nur Multimodalität. Tatsächlich sind ja Bilder und Texte in Zeitungen miteinander verbunden (indirekt semantisch und oft auch formal durch direkte Verweise). Erst wenn die involvierten Kodes so eng miteinander verwoben oder verschränkt sind, dass es sinnvoll ist, von *einem* (dominanten) Kode zu reden, der über mehrere Sinneskanäle wahrgenommen wird (z.B. die Lautsprache mit redegleitender Gestik), spricht Fricke von ‚sprachlicher Multimodalität im engeren Sinn‘ (Fricke 2012: 47). Was dann bei Stöckl als der Normalfall für Multimodalität gilt – Text zusammen mit Bild –, ist bei Fricke lediglich ein Fall von Multimodalität ‚im weiteren‘ Sinn, weil nur in derselben Sinnesmodalität wahrgenommen wird (z.B. Text und Bild auf einer Zeitungsseite oder Schrift und Zeichnungen eines Comics). Bei Stöckl sind dagegen Gestik und andere redegleitende Eigenschaften eher ‚submodes‘ von gesprochener Sprache, wie die Typographie bei einem (visuellen) schriftlichen Text.

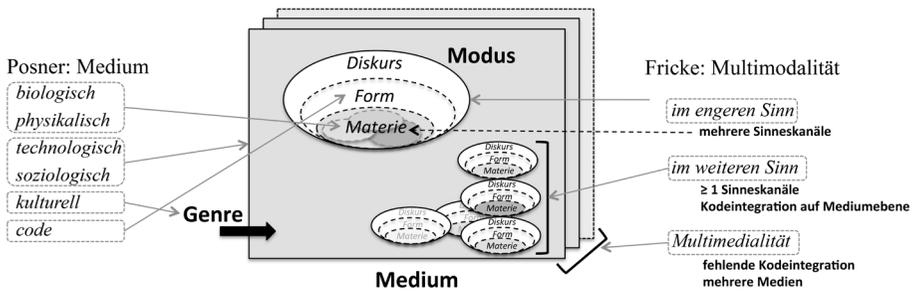
Die Verwendung der Begriffe ‚Dominante/Subordinierte‘ Kodes sowie ‚Modes/Submodes‘ erzeugt neue schwierige Fragen. Viele dieser Fragen lassen sich aber klären, wenn eine stärkere, bereits multimodal ausgerichtete Definition des Begriffs ‚Medium‘ hinzugefügt wird. In der Literatur ist

die Dringlichkeit einer medien- und kommunikationswissenschaftlich adäquaten Definition von Medium mehrfach betont worden (vgl. Krämer 2003, 2008; Stetter 2005; Elleström 2014). Eine solche Definition, die außerdem auch semiotisch adäquat wäre, gab es allerdings bis vor kurzem nicht. Ist ‚Sprache‘ ein Medium? Sind, wie bei Stetter, geschriebene Sprache, die Lautsprache und die Gebärdensprache alle separate Medien? Und wenn ja, ist dies die gleiche Art von Unterscheidung, die Schriftsprache und Bild als getrennte Medien erfassen würde? Diverse Unterschiede scheint es zu geben, aber wie sie unter einen theoretischen Ansatz zusammenzubringen sind, blieb meist unklar.

Diese Problematik ist auch Fricke sehr gegenwärtig und, um weiterzukommen, baut sie auf den Vorschlag von Posner (1986) auf, in dem sechs notwendige Facetten eines Mediumbegriffs erläutert werden: Medien sind jeweils biologisch, physikalisch, technologisch, soziologisch, kulturell und kodebezogen aufzufassen (Posner 1986: 293–296). Posners Systematisierung ist ein wichtiger Schritt für die Identifikation von relevanten Eigenschaften, die in der Kommunikation eine Rolle spielen. Allerdings stellen, wie in Bateman (2017) argumentiert, Definitionen von semiotischen Modi, von Kodes und von Multimodalität auf einer solchen Basis ein unauflösbares Problem dar, weil alle diese Kernbegriffe letztlich immer mit Hilfe der anderen Begriffe definiert werden. Dabei werden notwendige Grenzen zwischen ‚Kode‘ und ‚Modus‘, zwischen ‚Medium‘ und ‚Materialität‘, zwischen ‚Medium‘ und ‚Kode‘, zwischen ‚Medium‘ und kommunikativem ‚Zweck‘ immer wieder verwischt. Um diesen gordischen Knoten zu durchschlagen, muss die Aneinanderreihung von Medienbegriffen bei Posner eine genauer festgelegte Binnenstruktur erhalten.

Eine Voraussetzung dafür ist der hier erweiterte Begriff von semiotischem Modus. Erstens werden die biologisch und physikalisch bezogenen Medienbegriffe bereits durch die materielle Ebene des semiotischen Modus abgedeckt. Ein semiotischer Modus kann dann ‚multisensoriell‘ sein, weil eine Materie (eines Modus) sich frei aller wahrnehmbaren Eigenschaften bedienen kann. Dies ist sicherlich der Fall bei der menschlichen Sprache. Zweitens ist der kodebezogene Medienbegriff bereits durch die beiden nicht-materiellen Ebenen des semiotischen Modus dargestellt. Posners kodebezogenes Medium spaltet sich im semiotischen Modus in Form und Diskurs auf. Diese Ebenen stellen damit den kodebezogenen Medienbegriff in einer feineren Auflösung dar. Mit der Diskursebene wird darüber hinaus der direkte Anschluss an soziokulturelle Kommunikationsformen, ‚Textsorten‘ oder ‚Genres‘ ermöglicht.

Semiotische Modi als Praxen tauchen jedoch nie außerhalb soziogellschaftlicher Kontexte auf, die immer mehr oder weniger institutionalisiert sind. Erst hier finden wir den genuinen Beitrag eines angemessenen Mediumbegriffs. Ein Medium ist dann nichts anderes als ein konkreter institutionalisierter Kontext. Ein Teil dieser Institutionalisierung hat natürlich zur Folge, dass neue technologisch-bezogene Methoden für die Produktion, Verbreitung und den Konsum von Medienangeboten entwickelt werden.



**Abb. 1:** Graphische Darstellung der geschichteten Binnenorganisation von semiotischen Modi und ihren Beziehungen zur Mediumklassifikation von Posner (links) und zum Multimodalitätsbegriff von Fricke (rechts).

Winkler schlägt vor, dass ‚Zeichen‘ von Medien als Schemata zu konzipieren sind, die die Welt ‚lesbar‘ machen (Winkler 2008). Dieser Zeichenbegriff wird hier durch den viel detaillierteren Begriff des semiotischen Modus ersetzt. Deshalb besteht ein Medium üblicherweise aus mehreren Modi (vgl. Bateman 2016, 2017). Alle von Posner für Medien und Kommunikation aufgelisteten Facetten sind damit angemessen aufgenommen und zueinander in konkrete Relationen gestellt. Die Gesamtstruktur und ihre Beziehung zu Posners Kategorien und Fricke's Multimodalitätsbegriff ist in Abb. 1 graphisch dargestellt. Obwohl einige Klassifikationen dann anders als bei Fricke und Stöckl ausfallen müssen – zum Beispiel bzgl. der Frage, welche Kommunikate genau ‚mono-‘ bzw. ‚multi-‘modal sowie ‚mono-‘ bzw. ‚multi-‘medial sind –, ist der Gewinn für ein Analyseinstrumentarium substantiell. Ein solches Analyseinstrumentarium ist damit allen Fällen von multimodaler Kommunikation gewachsen und liefert eine direkte Unterstützung für empirische Forschung (vgl. Bateman und Wildfeuer 2014; Hiippala 2016; Bateman u.a. 2017a; Bateman u.a. 2017b; Bateman, Beckmann und Varela 2018; Tseng und Bateman 2018; Tseng, Laubrock und Pflaeging 2018). Einige Beispiele werden im nächsten Abschnitt aufgegriffen und mit den DH in Verbindung gebracht.

## 2.2 Anwendungsbeispiele

Nach dieser kurzen theoretischen Vorstellungsrunde soll jetzt an Hand einiger Beispiele gezeigt werden, wie dieses Modell der Multimodalität und seine methodologische Umsetzung einige problematische Trennlinien zwischen Bedeutungsformen, z.B. Text, Sprache, Bild, in den DH durchbricht. Die Beispielbereiche sind heute alle zu Untersuchungsgegenständen der DH geworden und sind daher eine gute Illustration der herrschenden Gemeinsamkeiten und Unterschiede im aktuellen Stand der DH und der multimodalen Linguistik. Es wird deutlich werden, dass viele Probleme mit Heterogenität in Ansätzen und Ergebnissen in den DH dadurch verursacht

sind, dass Abstraktionsebenen verkannt oder vermischt werden. Sprache und Bild sind zum Beispiel zwei Möglichkeiten, wie das Schema für semiotische Modi instanziiert werden kann. Sprache und Bild haben unterschiedliche Eigenschaften und keine davon ist als ‚Grund‘ für das Andere anzunehmen: die sogenannte ‚Logik‘ des Bildes ist ganz anders als die ‚Logik‘ der Sprache (vgl. Boehm 2004; Kress 2010). Nichtsdestotrotz gehorchen beide doch ‚Logiken‘ im Sinne der hier eingeführten semiotischen Ebene des Diskurses. Genau deswegen gibt es trotz stark voneinander abweichender Funktionalitäten und Wirkungsweisen Überlappungen und Ähnlichkeiten, die oft durch eine starre Trennung Bild/Text marginalisiert und zum Verschwinden gebracht werden. Sicherlich ist jede Art sprachorientierter Logo-zentrismus strikt abzulehnen; aber wenn Bild und Text wirklich als so verschieden voneinander verstanden werden, dass Verbindungen oder Ähnlichkeiten in ihren Wirkungsweisen nicht mehr bestehen können, schießt man über das Ziel hinaus, ihre Spezifika zu respektieren. Eine Brücke zwischen ihnen muss geschlagen werden, wenn die DH davon profitieren will und ein Zerfall in Einzelbereiche vermieden werden soll.

### 2.2.1 Beispiel 1: Annotationen und Multimodale Texttypentwicklung

In diesem Abschnitt wird als Beispiel die zeitliche Entwicklung eines konkreten multimodalen Texttyps, englischsprachige Tourismus-Broschüren, betrachtet. Als Vorbereitung dazu wird kurz vergleichend auf die Rolle von Annotationen zur Datenaufbereitung in den DH und in der multimodalen Linguistik eingegangen. Im Anschluss daran wird ein in der multimodalen Linguistik entwickeltes allgemeines Annotationsschema für statische seitenbasierte Dokumente, das auch für die Tourismus-Broschüren angewandt worden ist, vorgestellt.

Traditionell liegt der Fokus der Forschung zu statischen Druck- und Onlinemedien auf der sogenannten Text-Bild-Schere. Multimodal gesehen ist diese Trennung allerdings nicht ohne theoretische Bedenken: Texte als schriftliche Sprache mit Typographie, Komposition und anderen Aspekten sind auch immer visuell (vgl. Holly 2009; Kapuścińska 2017), und was genau unter Bild verstanden wird – Abbild, Diagramm, Landkarte, Photograph usw. –, ist für Analyse ebenso wichtig (vgl. Eide 2015). Durch empirisch orientierte Untersuchungen von Artefakten, in denen die visuell betrachtbare und verformbare Materie im semiotischen Vordergrund steht, kommt eine Reihe von weiteren semiotischen Modi zum Vorschein, die einerseits gemeinsam diese Materie gestalten und andererseits oft quer zu herkömmlichen Etiketten wie ‚Bild‘ oder ‚Text‘ stehen.

Die empirische Untersuchung von solchen Gegenständen erfordert in der Datensicht eine Multiperspektivierung. Dafür sind multimodale Korpora, die diese Perspektivierung unterstützen, unabdingbar. Wie aus der linguistischen Korpusarbeit gut bekannt, ist dies am besten durch Mehr-Ebenen-Annotationsschemata zu bewerkstelligen. Unterschiedliche Annotati-

onsebenen sind voneinander unabhängig und Korrelationen sind über Querweise festzulegen. Diese Entwicklung wurde auch in den DH durchgemacht. Ursprünglich wurden (schriftliche) Texte elektronisch mit Hilfe standardisierter ‚Markup‘-Schemata abgespeichert, die die Texte als inhaltsorientierte Hierarchien wiedergeben. Ein Text würde dann zum Beispiel ein ‚Buch‘ sein, das Komponenten wie ‚Kapitel‘ hat, die wiederum in ‚Absätze‘ gegliedert werden, die ‚Sätze‘ als Sequenzen von ‚Buchstaben‘ enthalten (vgl. TEI: Vanhoutte und van den Branden 2010). Schwächen der Annahme, dass eine einzelne Hierarchie ausreichen kann – auch als OHCO-Annahme („Ordered Hierarchy of Content Objects“) bekannt –, sind aber breit diskutiert worden, und heute ist diese Art der Darstellung von schriftlichen Texten weder aufrecht zu erhalten noch wünschenswert. Es ist immer möglich, ja sogar wahrscheinlich, dass unterschiedliche, voneinander unabhängige Strukturierungen für eine sinnvolle Darstellung eines ‚Texts‘ notwendig sind – bei Editionen muss man z.B. gleichzeitig Kapitel und Absätze einerseits und Buchseiten und Seitenumbrüche andererseits vorsehen, bei Sprechdaten stimmen syntaktische und prosodische Strukturen häufig nicht miteinander überein, usw. Mit multiperspektivischen Annotationen verursachen solche ineinander greifenden Strukturierungen kein Problem mehr. Allerdings ist die Verwendung solcher Annotationsschemata in den DH immer noch nicht ganz so üblich wie in der Korpuslinguistik, obwohl neue Markupstandards immer mehr in diese Richtung weisen (vgl. Bánski 2010).

Noch wichtiger für die aktuelle Diskussion ist aber der Status der angenommenen Annotationsebenen. Annahmen darüber, was genau dargestellt wird, haben sich mit der Zeit auch in den DH weiterentwickelt. Eine ziemlich grundlegende Diskussion wurde zum Beispiel von Renear (1997) geführt, in der auch betont wird, dass die Nutzung von digitalen Medien ein solideres Verständnis des ‚Text‘-Begriffes erfordert. Renear schlägt vor, dass Markup selbst eine Theorie eines Texts darstellt und darüber hinaus eine Markup-Sprache und ihre Annahmen als eine allgemeine Konzeption von Text zu sehen ist. Im Bezug auf die OHCO-Annahme differenziert Renear drei Positionen, die er auch zeitlich mit den Entwicklungen der DH in Beziehung setzt. Erstens kann Text ‚platonisch‘ gesehen werden, insofern dass das Markup eines Texts die ‚abstrakte Realität‘ des Textes als eine Hierarchie von inhaltlichen Einheiten auffasst. Zweitens wird eine eher ‚pluralistische‘ Sicht auf Text angenommen, die davon aussieht, dass ein Text aus mehreren Sichtweisen besteht. Und drittens wurde eine ‚Antirealismus‘-Position eingenommen, in der ein Text durch umfangreiches Markup, das die Annahmen über Interpretationen des Textes expliziert, erst zustande kommt: Vor Festlegung dieser Eigenschaften gäbe es seiner Position nach keinen ‚bereits existierenden‘ Text.

Renear positioniert sich selbst eher als der pluralistischen Position verpflichtet, die vielleicht auch der Auffassung von Annotationen bzw. Transkriptionen in der Linguistik am nächsten steht. Dort ist als Grundprinzip schon lange etabliert, dass es der Zweck von Transkriptionen ist, spezifi-

sche nützliche Perspektiven auf eine annotierte Einheit explizit zu machen und nicht diese Einheit genau wiederzugeben (vgl. Ochs 1979). Annotationen werden stets unternommen, um angenommene Eigenschaften und wiederkehrende Muster des Annotierten zwecks weiterer Analyse festzuhalten. Die oben erwähnte Kritik eines vermeintlichen Mangels an Rücksichtnahme auf die Materialität von Texten in den DH von Hayles (2003) führte Hayles allerdings zu einer Ablehnung aller drei von Renears identifizierten Positionen als immer noch zu linguistisch in ihrer Orientierung und daher unpassend für die DH. Für Hayles ist die Mitberücksichtigung von Materie unabdingbar:

*The materiality of an embodied text is the interaction of its physical characteristics with its signifying strategies. Centered on the artifact, this notion of materiality extends beyond the individual object, for its physical characteristics are the result of the social, cultural, and technological processes that brought it into being (Hayles 2003: 277; Hervorhebung im Original).*

Dies entspricht aber genau der Auffassung der Verbindung von Materie und Interpretationen („signifying strategies“), die sich aus dem oben eingeführten Modell des semiotischen Modus ableiten lässt. Eine Strategie, die es uns erlauben würde, diese Positionen aus der Annotationsdebatte weiterzuentwickeln, wäre daher, das durchaus überwiegend linguistische Verständnis von ‚Kodes‘ in der DH-Diskussion durch die hier gegebene Definition von semiotischen Modi zu ersetzen.

Jede Ausformulierung der zwei nicht-materiellen Ebenen (Ebenen (ii) und (iii)) eines semiotischen Modus dient dann auch für die Definition von Annotationsebenen. Die materielle Seite eines Modus (Ebene (i)) definiert dagegen die Einheiten in den Daten, die Interpretationen/Annotationen bekommen können. Genauer: Formalisierungen der ‚mittleren‘, formellen Ebene (ii) eines Modus stellen paradigmatische Klassifikationssysteme dar, die in Annotationsschemata direkt übersetzt werden können. Die oberste Diskurs-Ebene (iii) bereitet dann die Möglichkeit von automatisierten oder halbautomatisierten Interpretationsverfahren vor. Noch abstraktere Ebenen, zum Beispiel bezüglich Genre und den Eigenschaften eines Mediums (vgl. Abb. 1) lassen sich ebenfalls spezifizieren. Ob solche Annotationen eher zu Renears Klassen des Pluralismus oder des Antirealismus gehören, hängt vom Abstraktionsniveau ab. Je interpretativer die Beschreibungsebene, desto wahrscheinlicher wird es, von einer zweckgebundenen Zuschreibung von Merkmalen, die nicht intrinsisch dem annotierten Objekt zuzuordnen sind, Gebrauch zu machen – wir werden im nächsten Beispiel darauf zurückkommen. Im Prinzip ist damit bereits eine Brücke zwischen den theoretischen Strukturen und Beziehungen des hier eingeführten Modells der Multimodalität und seiner praktischen und empirischen Umsetzung in den DH geschlagen.

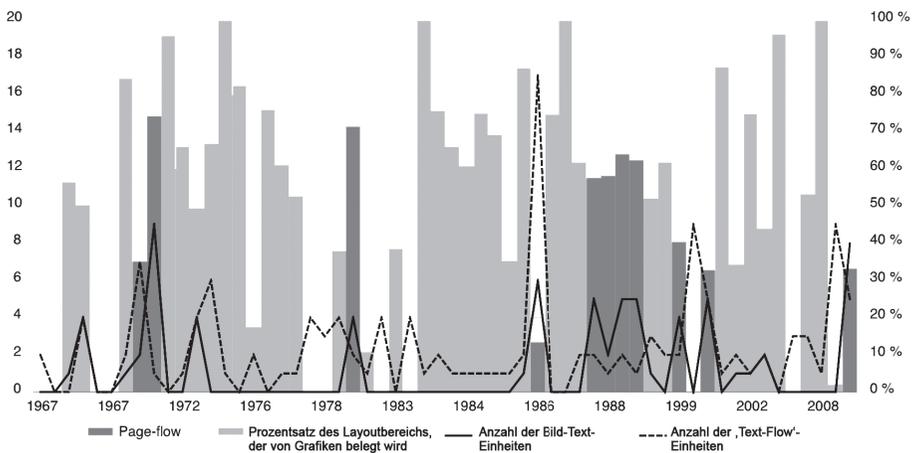
Eine wegweisende Studie dieser Art ist von Hiippala (2015) durchgeführt worden. Diese Analyse basiert auf dem von Bateman, Delin und Hen-

schel (2004) und Bateman (2008) entwickelten ‚Genre und Multimodalität‘ (GeM)-Modell, aus dem die hier eingeführte Konzeption von semiotischen Modi entstanden ist. Eine Analyse innerhalb des GeM-Ansatzes besteht darin, für jede ‚Dokumentenseite‘ innerhalb eines ausgewählten Korpus eine detaillierte Beschreibung auf jeder der fünf GeM-Annotationsebenen bereitzustellen: die in der Materie identifizierbaren Einheiten, die visuelle Komposition bzw. das Layout, die linguistische Analyse von sprachlichen Einheiten, die rhetorische Struktur des Ganzen und die Genrezuordnung. Diese Beschreibungen folgen traditionellen Methoden sowohl der Korpuslinguistik als auch der qualitativen Inhaltsanalyse (z.B. Schreier 2012), wobei die Kategorien der Annotation mit eindeutigen Identifikationskriterien so zuverlässig wie möglich gemacht werden sollten. Die Verbesserung der Zuverlässigkeit ist ein ständiges Ziel von Arbeiten dieser Art, da mehrere Aspekte der Analyse (insbesondere die rhetorische Ebene) spezielles Wissen erfordern und explizit interpretativ sind. Nach dem Erstellen der Annotation werden Korrelationen zwischen den einzelnen Aspekten gesucht, um Gruppierungen in multimodale Genres, Variationen von Genres über die Zeit und Beziehungen zu den involvierten Distributionsmedien (Druck, Web usw.) vorzuschlagen.

Diese Charakterisierung mit Hilfe semiotischer Modi geht über frühere semiotische Darstellungen hinaus, insbesondere im Hinblick darauf, wie die Analyse genau ablaufen soll und wie verschiedene Modi miteinander in Beziehung gesetzt werden können. Ein frühes Ergebnis dieser empirischen Arbeit war der Entwurf dreier neuer semiotischer Modi, die in seitenbasierten Medien aktiv zu sein scheinen. Nach Bateman (2008: 175) sind diese Modi: (i) Textfluss (engl. *text flow*), (ii) (statischer) Bildfluss (engl. *image flow*) und Seitenfluss (engl. *page flow*). Kurzgefasst lassen sich diese drei Modi wie folgt definieren: Textfluss ist die übliche Ansicht von linguistischem Text als ein einziger eindimensionaler Strang sich entfaltender linguistischer Strukturen. Es werden keine weiteren Eigenschaften der materiellen Form genutzt, außer denen, die sich der eindimensionalen Zeichenprozession unterordnen lassen. Typografie und Textfarbe tragen also immer noch zu diesem semiotischen Modus bei. Der Textfluss erlaubt auch bildhaftes oder diagrammatisches Material, das jedoch an bestimmten Stellen in der eindimensionalen Entwicklung verankert ist: Es werden keine räumlichen Beziehungen zwischen dem Bildmaterial und dem Textmaterial über diese Verankerung hinaus genutzt. Der Bildfluss ist ähnlich, da er sich auf eine eindimensionale Entwicklung stützt, aber über bildliche Darstellungen operiert. Schließlich ist Seitenfluss ein semiotischer Modus zweiter Ordnung, der dazu dient, Inhalte anderer Modi nach den zweidimensionalen Möglichkeiten einer Seite oder einer Seiten-ähnlichen Oberfläche zu organisieren. Jeder Modus kann in der Analyse durch sein visuelles Erscheinungsbild und das räumliche Layout der in Frage stehenden Seiten erkannt werden, wenn auch nicht immer eindeutig.

Jeder der drei Modi bringt seine eigenen organisatorischen Eigenschaften mit sich, einschließlich unterschiedlicher materieller, formaler

und diskursiver Anforderungen. Diese neu identifizierten Modi ergänzen bereits etablierte semiotische Modi, wie Diagramme, bildliche Darstellungen usw., worüber bereits eine ausführliche Literatur besteht (vgl. Bertin 1983; Sachs-Hombach 2003; Farias 2005). Im Wesentlichen stellen diese Modi, wie alle semiotischen Modi, Bündel von Ressourcen bereit, die empirisch untersuchbare Hypothesen darüber vorstrukturieren, wie das Material auf einer ‚Seite‘ zu rezipieren ist. In diesem Sinne diskutiert dann Hiippala (2015) einen vollständig annotierten Korpus von 58 Doppelseiten aus englischsprachigen Touristenbroschüren, die von der Stadt Helsinki, Finnland, in der Zeit von 1967 bis 2008 veröffentlicht wurden. Die mehrschichtige Annotation macht deutlich, wie sich das Design in diesem Genre über die Zeit weiterentwickelt hat. Besonders auffällig ist die wechselnde Nutzung und Verteilung von Diagramminformationen, Anzeigen und Layout. Die Ergebnisse wurden sowohl durch traditionellere Vergleiche von quantitativen Daten, die aus dem annotierten Korpus extrahiert wurden, als auch durch experimentelle Visualisierungstechniken, erzielt. In Abb. 2 wird zum Beispiel ein Überblick über die Verteilung von semiotischen Modi aus dem Korpus gezeigt. Hier lassen sich viel genauere Betrachtungen über Designentwicklungen anstellen, als wenn nur oberflächlich über Text bzw. Bild gesprochen wird. Aus der Analyse lässt sich zum Beispiel deutlich zeigen, dass obwohl viel bildliche Darstellungen im Design vorkommen, trotzdem ausgesprochen wenig aus den Möglichkeiten des zweidimensionalen Layout-Raums gemacht wird. Dies wird aus den viel seltener auftretenden Balken zu Seitenfluss (‚Page-flow‘) ersichtlich.



**Abb. 2:** Zeitliche Verteilung der semiotischen Modi in Hiippalas Helsinki-Korpus (vgl. Hiippala 2015: 183; mit Genehmigung des Autors).

Das Grundprinzip, das semiotische Aktivität als Bündelung von semiotischen Ressourcen zur Interpretationslenkung auffasst, wird auch in einer Reihe von weiteren Arbeiten angewandt. Eine ähnliche korpusbasierte Studie ist zum Beispiel von Thomas durchgeführt worden; in seiner begleitenden Diskussion zur Methodologie weist er auf die Notwendigkeit hin, Analysen auf empirische Ergebnisse zu stützen, um interpretatorische Zirkelschlüsse zu vermeiden (Thomas 2014). Außerdem gibt es eine Reihe von Arbeiten zu statischen visuellen Narrativen wie Comics und graphischen Romanen (engl. *graphic novels*). In Bateman u.a. (2017a) wird das wohl bis heute ausgefeilteste Annotationsschema für den semiotischen Modus von Seitenkompositionen in Comics eingeführt, das jetzt auch für empirische Analysen mit Hilfe von Eyetracking-Verfahren eingesetzt wird (Bateman u.a. 2018). In Tseng u.a. (2018), Tseng und Bateman (2018) und Bateman und Veloso (2013) wird u.a. der von Tseng ausgearbeitete diskurssemantische Ansatz der *multimodalen Kohäsion* auf empirisch untermauerte und medienvergleichende Untersuchungen von Comics, Film und Literatur angewendet. In diesen Arbeiten findet man immer wieder die Vorteile einer semiotischen Position, in der Querverbindungen über feinteilig definierte semiotische Modi als Grundmechanismus verwendet werden, anstatt Bild, Text u.ä. getrennt voneinander zu behandeln. Analysen dieser Art dienen nicht nur als Unterstützung von empirischen Arbeiten, sondern zeigen deutlich den Weg zu tieferen interpretierenden Analysen z.B. des Narrativs, auch medien- und modiübergreifend. Dies soll nun an einem zweiten Beispiel gezeigt werden, in dem das gleiche Analyseinstrumentarium für das Medium Film eingesetzt wird.

### 2.2.2 Beispiel 2: Film als multimodales Artefakt

Wie bei dem ersten Beispiel ist hier die Entwicklung eines Mehr-Ebenen-Annotationsschemas, das die verschiedenen Beiträge der am Film beteiligten semiotischen Modi adäquat abbildet, zentral. Um dann zu Interpretationen und Bedeutungszuordnungen für Film zu gelangen, müssen auch ‚höhere‘, abstraktere Darstellungsebenen für die Beschreibung herangezogen werden. Im Falle von Film erstreckt sich die verformbare Materie über Ton, Bild und Bewegungen, die alle zusammen den Rohstoff für die bekannten technischen Details von Film darstellen; die Ressourcen der Diskursebene des semiotischen Modus des Film mobilisieren dann diese technischen Formen parallel. In Anlehnung an den oben zitierten Aufsatz von Boehm könnte man diese Diskursebene auch als den ‚Logos‘ des Films beschreiben. Dabei ist anzumerken, dass dieser Logos dann dem semiotischen Modus angehört – es handelt sich deswegen nicht um einen ‚Logos der Sprache‘ oder einen ‚Logos des Bilds‘, sondern um etwas Eigenes und Zusätzliches. In der Multimodalitätsforschung ist der semiotische Modus des Films bereits aus mehreren Perspektiven beleuchtet worden (vgl. van Leeuwen 1985; O’Halloran 2004; Bateman und Schmidt 2012; Tseng 2013; Wildfeuer 2014).

Als wichtigem Kulturartefakt wird dem Film auch in den DH zunehmend Aufmerksamkeit geschenkt, wobei wieder die gleichen Probleme und Grenzen bezüglich einer mangelnden semiotischen Verankerung der angewandten Ansätze zum Vorschein kommen. Heftberger (2015) analysiert zum Beispiel die formalen Aspekte von klassischen Filmen wie Dziga Vertovs *Der Mann mit der Kamera* (SU, 1929). Hier wird die den DH inhärente Fokussierung auf ‚Daten‘ und ihre Eigenschaften besonders deutlich. Heftberger schreibt:

Für eine formale Untersuchung eignet sich Vertovs Werk deshalb besonders gut, weil der Regisseur seine politischen Botschaften und künstlerischen Ideen in formalen Verfahren wie Einstellungslänge, Einstellungsgröße, Bildkomposition oder Bewegungsintensität konzipierte (Heftberger 2015).

Die meisten Filme operieren aber anders und drücken trotzdem politische und ideologische Positionen aus: wie dies erklärt werden kann, bleibt bei dem formalen Ansatz unklar. Ähnlich gehen Zeppelzauer, Mitrović und Breiteneder (2011) vor. Sie entwickeln automatische Erkennungskomponenten für das Auftreten von visuellen, akustischen und filmtechnischen (z.B. Schnitte zwischen Einstellungen) Elementen. Das gleichzeitige Auftreten bestimmter Elemente wird dann als Signal für besonders relevante Filmsegmente betrachtet. Obwohl Zeppelzauer u.a. diesem gleichzeitigen Auftreten einen besonderen Status zuschreiben, sind die behandelten Montagesequenzen eigentlich nur ein Sonderfall der allgemeinen Operation der filmischen Diskursebene – konkret handelt es sich wieder um die Diskursfunktion der oben erwähnten multimodalen Kohäsion, die in Tseng (2013) für Film ausführlich vorgestellt wird. Unser Beispiel wird jetzt zeigen, wie die Anwendung von Diskursmechanismen es möglich macht, den synchronisierten Einsatz von Materie über Sinneskanäle hinweg als bedeutungskonstituierend zu analysieren im allgemeinen organisatorischen Rahmen der semiotischen Modi.

Multimodale Kohäsion im Film beschreibt filmische Präsentationsstrategien, mit denen Figuren, Orte und Ereignisse im Film audiovisuell eingeführt und wiederkehrend aufgenommen werden. Weil ein sogenanntes kohäsives Element dann über Ton, Bild, Musik, Bewegung, Schnitt usw. getragen werden kann, bietet die kohäsive Analyse ein klares Instrument für die Zusammenführung von ‚niedrig‘-rangigen Merkmalen zu Interpretationsketten höheren Rangs. Hierbei ist die Kernannahme der multimodalen Linguistik, dass es möglich sei, eine ‚Abstraktionsleiter‘ von quantifizierten Daten hin zu geisteswissenschaftlich relevanten abstrakten Interpretationen zu ersteigen. Durch dieses Verfahren wird das in der Filmwissenschaft lang bekannte Problem gelöst, dass es einfach nicht möglich ist, filmtechnischen Merkmalen feste Bedeutungen zuzuschreiben (Branigan 1984: 29). Eine Interpretation ist immer mehr oder weniger kontextbezogen und dies muss im Interpretationsmechanismus berücksichtigt werden. Genau das leistet die Diskursebene (Ebene (iii)) der angewandten semio-

tischen Modi (der formale Hintergrund wird genauer in Bateman und Wildfeuer 2014 beschrieben). Die Diskursebene grenzt genau ein, welche Bestandteile des Kontexts in welcher Hinsicht für die interpretatorische Entscheidung relevant sind. Diese Vorgehensweise liefert eine allgemeine Methodik zur Analyse von Filmen, weil im Gegensatz zu projektspezifischen technischen Details, die in den DH-Ansätzen Verwendung finden, filmische Kohäsion immer aktiv ist: es gibt keinen Film, in dem Kohäsion keine Rolle spielt. Daher wäre eine festere Verankerung dieses Verständnisses der semiotischen Modi von Film in den Analysewerkzeugen der DH sicherlich zu empfehlen.

Um die oben erwähnte progressive Erhöhung des Abstraktionsniveaus der Annotationsebenen zu illustrieren, betrachten wir kurz das immer wiederkehrende Problem der Genre-Klassifizierung für Film und Video. Genreerwartungen spielen gemeinhin für die Interpretation von Filmen durch den Zuschauer eine zentrale Rolle, aber die Natur von ‚Genre‘ als solches ist noch umstritten (vgl. Altman 1984; Neale 2000). Um mögliche Unterschiede zwischen eng verwandten Genres wie Kriegsfilmen und Westernfilmen zu untersuchen, wurde in einem in Bateman u.a. (2016) geschilderten Projekt eine Kombination von multimodaler Kohäsions- und Ereignisanalyse mit Hilfe automatischer Merkmalsextraktion durchgeführt. Heutzutage gehört eine ganze Palette von solchen automatischen Verfahren zu den üblichen Werkzeugen der computergestützten Filmanalyse, auch (sogar meistens) außerhalb der DH. Dabei werden mehrere Dimensionen für audiovisuelle Vergleiche angeboten, z.B. Farbpaletten, Helligkeit, Gesichtserkennung, Geräuschpegelunterschiede, Bewegungsintensität sowie zunehmend explizit filmtechnische Eigenschaften wie Einstellungsgrößen, Kamerafahrten usw. (vgl. Canini u.a. 2011; Benini u.a. 2016; Suchan und Bhatt 2016; Wu und Christie 2016). Ohne eine Diskursebene, in der diese Merkmale zwecks Interpretation sinnvoll zusammengeführt werden, sind die Ergebnisse, die allein aus den Daten gewonnen werden, zwangsläufig begrenzt. Mit einer zusätzlichen Diskursebene wie der filmischen Kohäsion kann man folgendermaßen weiterkommen.

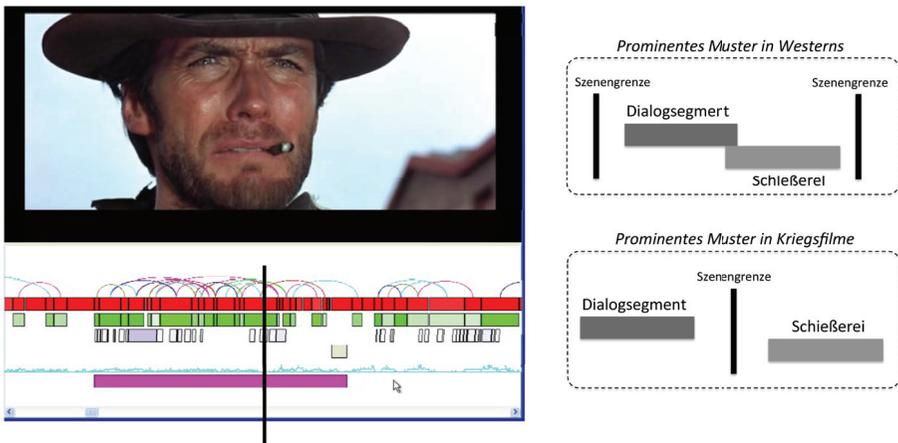
Zuerst wurden automatisch erkannte technische Merkmale sowie von Hand angefertigte Annotationen abstrakterer Art mit Hilfe eines im Projekt implementierten ‚Videoplayers‘ visualisiert. Die daraus resultierende gemeinsame Darstellung von Annotationsschichten mit unterschiedlichem Status – z.B. automatische Analyse und manuelle Kodierung – wird zweitens durch eine weitere automatisch erzeugte Annotationsebene ergänzt, die aus den weniger abstrakten Merkmalen abgeleitet wird. Dies erlaubt eine Vorsortierung der filmischen Daten nach bestimmten abstrakteren filmischen Mustern, die einerseits trotzdem automatisch detektierbar bleiben und andererseits bereits auf noch abstraktere komplexe Muster hinweisen können. Das Darstellen und Browsen von Ergebnissen in dem erweiterten Videoplayer zeigt direkt, wie eine datengetriebene, aber außerdem noch *f u n k t i o n a l* interpretierte Analyse vorgenommen werden kann. Solcherart interpretierte Analysen können dann für die Überprüfung von Forschungs-

hypothesen verwendet werden – auf Basis von Korrelationen zwischen weniger abstrakten und abstrakteren Klassifikationen des Datenkorpus.

Als konkretes Beispiel betrachten Bateman u.a. (2016) eine der ‚generisch‘ typischsten Arten von Ereignissen, die im Western zu beobachten sind: das ‚shoot out‘. Ereignisse dieser Klasse kommen oft an einem narrativen Höhepunkt oder an der narrativen Auflösung vor und zeigen eine Schießerei zwischen den Hauptprotagonisten des Films. Schießereien verschiedenster Art kommen aber auch in Kriegsfilmern vor. Daher wurde untersucht, ob Unterschiede in dem kohäsiven Aufbau dieser auf der Oberfläche ähnlich scheinenden Szenen zwischen den zwei Genres zu finden wären. Wenn dies der Fall wäre, hätten wir einen aus den Daten motivierten Grund, einen Genreunterschied anzunehmen. Dafür wurde eine weitere Annotationsebene definiert, die aus der folgenden Kombination von niedrigeren (d.h. datennahen) Merkmalen ermittelt wurde:

(Aussprache VOR Schießerei) UND Gesichter UND Schuss-Gegenschuss

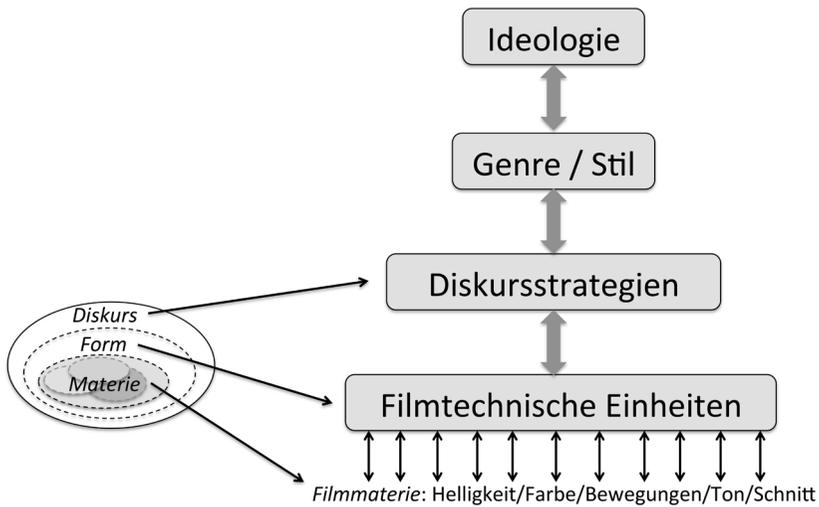
Eingesetzt für Western lieferte dieses Muster wie erwartet eine Menge ‚shoot outs‘. Visuell ist dies in Abb. 3 zu sehen: Alle Annotationsschichten, egal ob per Hand oder automatisch erzeugt, und unabhängig davon, auf welcher Abstraktionsebene sie liegen, sind im Videoplayer als zusätzliche Analysebalken ersichtlich. Dies liefert ein Verfahren zum Isolieren von Filmsegmenten, die genau jene Bedingungen aufweisen, die zu untersuchen sind.



**Abb. 3:** Ein Screenshot des Videoplayers mit einer Szene aus *Für eine Handvoll Dollar* (Sergio Leone, 1964, S/I/BRD) mit automatisch erkannten Merkmalen. Die ersten Balken zeigen (von oben nach unten): durch verbundene Bögen visuelle Kontinuität, Einstellungen, Dialoge, gesprochene Sprache, Pistolenschüsse und Schallpegel. Der letzte Balken zeigt das Ergebnis der Anwendung des ‚shoot-out‘-Musters. In den zwei rechten Bildern wird diagrammatisch dargestellt, wie sich die Schießereien im Western und Kriegsfilmern bezüglich dieses Musters unterscheiden (Bateman u.a. 2016: 149).

Wenn jedoch das gleiche Muster auf die Kriegsfilme im Projektkorpus angewendet wurde, gab es fast keine erkannten Fälle des Musters. Ein Vergleich zwischen der relativen Verteilung dieser Klasse von Ereignissen in den beiden Filmsammlungen könnte somit einen *qualitativen* Unterschied zwischen narrativen Konfigurationen im Kriegs- und im Western-Film widerspiegeln. Während in Kriegsfilmen transaktionale Aktionen wie Schieß- und Kampfhandlungen von identifizierbaren Charakteren (Soldaten) gegen allgemeine (und meistens nicht identifizierbare) Feinde ausgeführt werden, sind transaktionale Handlungen im Western Interaktionen zwischen hervorstechenden Charakteren, die ihre verbalen Interaktionen fortsetzen oder direkt an das shoot-out anschließen. Damit ist ein mögliches Differenzierungsmerkmal identifiziert worden, das ‚bottom-up‘ feststellbar und unabhängig von bestehenden Genre-Kategorien anwendbar ist. Bestehende Genre-Kategorien funktionieren dann lediglich als prätheoretische Gruppierungen, die der Suche nach belastbareren Kategorien dienen.

Dabei wird deutlich, wie das Abstraktionsniveau der geleisteten Interpretationen in dieser Art von Untersuchung sich allmählich erhöht von automatisch erkannten technischen Merkmalen bis hin zu Fragen der Narrativkonstruktion und des unterschiedlichen Einsatzes von etablierten filmischen Konstruktionen im Allgemeinen. Der Grundgedanke des Ansatzes kann dann wie folgt zusammengefasst werden: Die Verwendung eines geschichteten, mehrstufigen Verfahrens liefert ein sensitiveres Instrument dafür, signifikante Unterschiede zwischen Genres festzustellen. Genreanalysen dieser Art können sogar durchgeführt werden, ohne bereits existierende (und fast immer problematische) Genre-Labels anzunehmen. Anstatt mit solchen Vorannahmen zu arbeiten, läuft dann eine Genreidentifizierung über Kombinationen von audiovisuellen ‚high-level‘- und ‚low-level‘-Merkmalen ab. So wäre es möglich, Filme zu gruppieren, die sich systematisch dadurch voneinander unterscheiden, dass divergierende kommunikative Strategien für ähnliche narrative Zwecke eingesetzt werden. Zum Beispiel wäre die ‚shoot out‘-Strategie im Genre des Westerns ein Ausdruck für ein individuelles, auch psychisches Kräftemessen, aber im Genre des Kriegsfilms bleibt sie eher anonym. Selbstverständlich und wie auch in den DH üblich benötigt ein solches Verfahren Information aus möglichst vielen Filmen. Im Kontrast zu jetzigen DH-Ansätzen funktioniert dieser Ansatz jedoch methodologisch durch die Identifizierung von Annotationsebenen auf einem jeweils sehr unterschiedlichen Abstraktionsniveau. Darüber hinaus bedeutet die Verwendung von semiotischen Modi anstelle von spezifischeren und manchmal sogar projektbezogenen Werkzeugen, dass Ergebnisse von einzelnen Analysen aufeinander bezogen werden können. Schließlich befassen wir uns hier eigentlich nicht mit dem ‚Narrativ‘ als solchem, sondern mit allen Genres. Die Informationen, die semiotische Modi auf ihrer abstraktesten Ebene der Beschreibung liefern, betreffen *Diskurs* im Allgemeinen und nicht nur den narrativen Diskurs. Dementsprechend sind die für Film relevante Abstraktionsebenen, oder Strata, in Abb. 4 abgebildet.



**Abb. 4:** Strata der filmischen Bedeutungskonstitution und ihre Beziehung zu semiotischen Modi.

### 2.2.3 Zwischenfazit

In diesem Abschnitt wurde anhand von Beispielen erläutert, wie sich die analytischen Ziele der DH und die der multimodalen Linguistik in mancher Hinsicht ähnlich sind. Im Gegensatz zu den meisten bisherigen DH-Ergebnissen ist allerdings die Forschung der multimodalen Linguistik fest auf Interpretation und Interpretationsprozesse sowie deren Formalisierung, auch mit maschinellen Methoden, ausgerichtet. Dies muss sicherlich auch zu den Zielen der DH gezählt werden, sonst könnten sie kaum Beiträge zu den ‚humanities‘ liefern. Ein intensiverer Austausch mit der Semiotik ist daher geboten. Dabei können für die DH theoretische Prinzipien sowie methodologische Zugänge zur Analyse von komplexen Artefakten und Darbietungen gewonnen werden. Algorithmen, die auf solchen oben beschriebenen mehrschichtigen Repräsentationen fußen, könnten außerdem ihre Schlussfolgerungen transparenter machen und dabei die kritische Funktionalität der Erklärungen für die gezogenen Schlüsse unterstützen. Genauere Ausführungen zu allen hier beschriebenen Analysen sind aus der angegebenen Literatur zu entnehmen.

## 3. Multimodale Semiotik

Die Bezeichnung ‚multimodale Linguistik‘ ist damals in den 2000er Jahren mit der Strategie gewählt worden, die Betonung auf einige überwiegend aus der Linguistik herangezogene Grundprinzipien zu legen. Diese Grundprinzipien sind: (i) einen stratifizierten Blick (nach Hjelmslev 1974

[1943]) auf semiotische Systeme zu nehmen, (ii) einen allgemeinen semantisch-motivierten Begriff von ‚Diskurs‘ (nach Martin 1992; Asher und Lascarides 1994) als eines der angenommenen Strata einzuführen, und (iii) einen durchgängig empirischen Zugang zu Daten zu verlangen. Dank dieser Orientierungspfeiler wurde es möglich, beliebige Artefakte bzw. Darbietungen angesichts ihrer Multimodalität empirisch und datenorientiert zu untersuchen. Arbeiten dieser Art sind dann nicht mehr der ‚Linguistik‘ im traditionellen Sinne zuzuordnen. Ihnen liegt nunmehr die häufig verwendete Idee des ‚erweiterten Textbegriffs‘ zu Grunde und sie zeigen, wie dieser Begriff theoretisch sowie methodologisch produktiv zu verstehen ist.

In vielerlei Hinsicht hat die frühere Benennung des Ansatzes als multimodale Linguistik dann bereits ihre beabsichtigte Funktion erfüllt. Was aber eigentlich unternommen und dargestellt wurde, lässt sich nun präziser als ‚multimodale Semiotik‘ beschreiben. Hierbei behalten wir die aus der Linguistik gewonnene Orientierung auf Empirie und Modellbildung bei und fügen eine Auseinandersetzung mit einigen weiteren semiotischen Grundprinzipien hinzu. Dies ist auch für die DH potenziell von großem Nutzen. Wenn immer komplexere multimodale Gefüge ins Visier genommen werden, entstehen neue Herausforderungen, beziehungsweise müssen alte Herausforderungen wieder neu auf die Forschungsagenda gesetzt werden. Ohne semiotische Prinzipien bleibt eine Reihe von zentralen Konzepten und Modellierungsentscheidungen in den DH ohne ausreichende Grundlage. Hier gibt es viele relevante Diskussionspunkte, aber für den Zweck dieses Aufsatzes werde ich mich abschließend lediglich auf zwei ausgewählte kritische Bereiche der DH mit Bezug auf deren mögliches semiotisches Fundament konzentrieren: Modellierung und Visualisierung.

Modellierungen sollten zur Kernaufgabe der DH zählen – wie zum Beispiel Eide (2015) schreibt, auch mit Verweis auf McCarty (2005: 24):

In order to use the computer for advanced textual work, models have to be established. [...] Modelling is a core method in digital humanities, with close links to modelling in other disciplines, including the social and natural sciences and computer science (Eide 2015: 4).

Die Frage: ‚Was sind denn Modelle, egal aus welcher Disziplin?‘ führt uns zwangsläufig zur Semiotik. Modelle stellen eine inhärente *semiotische* Fragestellung dar, weil Modelle (möglicherweise abstrakte) Artefakte sind, die sich für ihre Interpretierenden in bestimmten Aspekten und durch spezifische Eigenschaften auf irgendetwas Anderes beziehen – um die Zeichendefinition von Charles Sanders Peirce heranzuziehen (z.B. Peirce 1998 [1893–1913]: 478). Modelle sind dann Zeichen *par excellence* und können nicht nicht semiotisch sein – eine Position, die jetzt auch in den DH an Gewicht gewinnt (vgl. Ciula und Eide 2017) und die im vorliegenden Beitrag weiter entwickelt wird.

Darüber hinaus gehören Modelle nach Peirce einer spezifischen Art von Zeichen (bzw. Zeichenwirkung) an. Der Sinn eines Modells ist, dass es Aus-

sagen über den modellierten Bereich erlaubt, die aus Eigenschaften des Modells selbst ableitbar sind. Wenn ein Modell diese Funktion nicht leistet (oder in dem Grad, in dem es diese Funktion nicht leistet), ist das Modell nicht (oder zu diesem Grad nicht) ausreichend. Dieser Typ von Zeichenwirkung wird bei Peirce als ‚Ikonizität‘ definiert. Ikonizität besteht genau dann, wenn die ‚internen‘ Eigenschaften eines Zeichens selbst auf die effektive Wirkung des Zeichens hinweisen. Der Vorschlag von Johanna Drucker zu ‚non-representational‘ grafischen Umgebungen, in dem die Begriffe „Bild als Etwas“ und „Bild von Etwas“ voneinander differenziert werden, weist genau in diese Richtung (Drucker 2017: 123). Damit kommen wir unmittelbar zum zweiten Thema dieses Abschnitts: der Visualisierung. Im Visualisierungsbereich ist schon lange die Rede von Ikonizität, weil Ikonizität häufig (obgleich fälschlicherweise) automatisch mit dem Bildlichen gleichgesetzt wird. Es ist daher angebracht, sich die grundlegenden semiotischen Begrifflichkeiten noch einmal zu vergegenwärtigen. Diese Angelegenheit ist auch viel mehr als nur ein rein theoretisches Unternehmen: Mit dem expliziten Bezug zur Ikonizität und der Rolle, der Ikonizität in dem Gesamtprozess der Semiose bei Peirce zugesprochen wird, werden wir in die Lage versetzt, eine genauere theoretische und formelle Betrachtung der Wirkung von Visualisierungen in Argumenten und Analysen zu liefern.

In beiden Bereichen, dem der Modellierung und dem der Visualisierung, wird die im vorigen Abschnitt gegebene Definition von semiotischem Modus stets mitwirken. Allerdings wird dafür eine genauere Verbindung zwischen diesem Ansatz der Multimodalität und dem Peirce’schen Zeichenmodell angewandt. Die semiotischen Grundlagen der früheren multimodalen Linguistik waren recht gemischt. Die Richtungen, die sich aus der Arbeit von Kress and van Leeuwen u.ä. entwickelt haben (z.B. Kress und van Leeuwen 2006 [1996]), sind auf Hjelmslev und die weitere funktionale Ausarbeitung von dessen Ansatz durch Halliday gegründet und orientieren sich nicht an Peirce. Die Richtungen, die sich stärker an der Interaktions- oder Textlinguistik orientieren (z.B. Mittelberg 2008; Fricke 2012), setzen mehr auf Peirce, insbesondere bzgl. Ikonizität und Indexikalität. Richtungen, die eher ethnomethodologisch oder interaktionsbasiert arbeiten, sind wieder anders ausgerichtet (vgl. Deppermann 2013; Haddington, Mondada und Nevile 2013; Norris 2016). Hier wird nun vorgeschlagen, dass es jetzt aus der Perspektive der Semiotik und der multimodalen Linguistik an der Zeit wäre, etwas prinzipieller und zielorientierter mit dem semiotischen Fundament von Multimodalität im Allgemeinen umzugehen und das ist genau das, was hier unter ‚multimodaler Semiotik‘ verstanden wird.

Für tiefere Diskussionen von Modellen und Visualisierungen ist eine genauere Betrachtung von Peirces Theorie der ‚Ikonizität‘ zentral. Die bisher detaillierteste Auseinandersetzung in den DH mit Peirce und Ikonizität ist wahrscheinlich die Diskussion von Ciula und Eide (2017). Ciula und Eide bauen auf Kraleman und Lattmanns (2013) früherem Versuch, Modelle mit semiotischen Mitteln zu beschreiben, sowie auf Elleströms (2013) Verständnis von Ikonizität auf. Dabei wollen sie Modellierungstätigkeiten in

den DH semiotisch besser verorten, um solche Modellierungen passgenauer an die Erfordernisse von geisteswissenschaftlichen Fragestellungen anpassen zu können – ein Ziel, das dem in diesem Aufsatz verfolgten recht ähnlich ist. Ciula und Eide stimmen mit Kralemann und Lattmann darin überein, dass Modelle im Allgemeinen als ikonische Zeichen angesehen werden müssen. Gemäß Peirce sind ikonische Zeichen nie nur als visuell aufzufassen. Sie stellen die Grundlage für Schlussfolgern jeder Art dar. Beide, Ciula und Eide und Kralemann und Lattmann, verwenden dann auch die von Peirce vorgeschlagene weitere Unterteilung von ikonischen Zeichen in drei Subtypen: ‚Image‘, ‚Diagramme‘ und ‚Metapher‘. Obwohl Peirces ‚Image‘ oft in deutschen Arbeiten als ‚Bild‘ übersetzt wird, ist es wichtig, den Begriff *keineswegs* auf das Visuelle einzugrenzen; Bauer und Ernst (2010: 43) reden dann vom ‚Bild-Ikon‘, um den Unterschied zum einfachen ‚Bild‘ zu betonen. In diesem Aufsatz werde ich Peirces Terminus unverändert weiter verwenden. In der Peirce-Literatur sind unterschiedliche Auslegungen seiner Dreiteilung zu finden. Ciula und Eide folgen der Annahme von Elleström, dass die Teilung drei Bereiche auf einem kognitiven Komplexitätskontinuum wiedergibt. ‚Image‘ ist dann die einfachste Art und benötigt daher die geringste kognitive Leistung, um verarbeitet zu werden, und ‚Metapher‘ ist die komplexeste (Ciula und Eide 2017: 5).

Peirces Definition der Dreiteilung kann aber auch kategorial verstanden werden, indem angenommen wird, dass alle drei Subtypen sich prinzipiell voneinander unterscheiden. Diese Position hat in der Tat mehrere Vorteile für genauere Analysen. Ikonen im Allgemeinen sind von Peirce so konzipiert, dass sie Schlussfolgerungen und ‚Experimente‘, die auf Basis der Eigenschaften des Zeichens selbst durchgeführt werden können, unterstützen. Das heißt, wenn zum Beispiel das Bild eines Baumes vorliegt, kann man auf Grund dieses ikonischen Zeichens über Proportionen und Abstände der Teile des Baumes zueinander Schlussfolgerungen ziehen. Dabei verwendet man direkt die Merkmale des Zeichens, um etwas über das Objekt des Zeichens zu erfahren. Die drei Subtypen eines ikonischen Zeichens können dann so verstanden werden, dass sie je eine bestimmte Art von Schlussfolgerung unterstützen.

Bei jedem Subtyp sind auch die Informationsquellen für das Schlussfolgern verschieden. Bei Images sind die perzeptiven Merkmale des Zeichens selbst in all ihrer Vielfalt die Informationsquelle; bei Diagrammen werden ausgewählte und abstrahierte Beziehungen zwischen den perzeptiven Merkmalen hervorgehoben und nur diese Beziehungen stehen dann für weitere Schlüsse zur Verfügung; und bei Metaphern wird eine weitere diagrammartige Darstellung als eine ‚parallele‘ Struktur hinzugefügt, um eine neue Sicht auf die ursprüngliche Repräsentation bereitzustellen (Peirce 1998 [1893–1913]: 273). Alle diese Kategorien sind in Peirces pragmatisches Verständnis von Zeichennutzung einzuordnen, demzufolge Zeichen für konkrete Zwecke verwendet werden müssen, um überhaupt ihre Zeichennatur zu bekommen. Das bedeutet, dass – genau wie bei Zeichen im Allgemeinen – nichts ein Modell (oder ein Ikon usw.) ist, wenn es nicht als

solches von jemandem verwendet wird. Eine tiefergehende Diskussion der kategorialen Trennung der Peirce'schen Zeichenarten und deren Relevanz für die Multimodalität ist in Bateman (2018) zu finden.

Kralemann und Lattmann (2013: 3410) halten Diagramme für die wichtigste Art von Modellen in der Wissenschaft. Dies wird von Beobachtungen wie der folgenden von Bauer und Ernst untermauert: „Das Diagramm-Ikon entwirft in der Darstellung eine Hypothese über den Gegenstand [...] Es ist diese Hypothese, mit der dann praktisch umgegangen und gearbeitet werden kann“ (Bauer und Ernst 2010: 44; Hervorhebung im Original). Oder wie Krämer es formuliert: „Wir schauen [Diagramme] nicht nur an, sondern machen etwas mit ihnen“ (Krämer 2014: 14). Hier wird klar, dass ein Diagramm nicht durch Vereinfachung oder ‚Weglassen‘ von Eigenschaften entsteht, wie zum Beispiel Elleström (2013) vorschlägt, sondern eine positive, ‚entwerfende‘ Wirkung für die Bedeutungskonstitution hat, die die Anwendung des Diagramms für weiteres Schlussfolgern bzgl. des Objekts der Darstellung ermöglicht. In den DH sind *Visualisierungen* fast ausschließlich Zeichen dieser Art. Ihre Repräsentationen sind aus den Daten errechnet und geben daher „relationships already existing in the objects being modeled“ wieder, wie Kralemann und Lattmann (2013) es beschreiben. Dies ist eine Art von Image, die auch bei Drucker (2014: ii) deutlich von „generative, diagrammatic, dynamic“ Images, die Wissen tatsächlich *produzieren*, differenziert wird. Das bedeutet aber keineswegs, dass aus solchen Repräsentationen nichts Neues gewonnen werden kann. Alle Autoren betonen, dass ein solcher Zeichengebrauch in einen (kommunikativen) Kontext oder Wissenschaftsdiskurs eingebettet werden muss, um zu wissen, wie zum Beispiel ein Diagramm (bzw. Modell) zu lesen ist. Diagramme können dann komplex und ästhetisch anspruchsvoll sein und sind, laut Peirce, unabdingbar, um zu neuen Erkenntnissen zu gelangen (vgl. auch Frank in diesem Heft).

Trotz dieser unumstritten zentralen Rolle von Diagrammen für den Wissenszuwachs bleibt es aber dabei, dass die in einem Diagramm ans Licht gebrachten Relationen aus den Daten stammen und *nur aus den Daten*. Sonst wären sie keine Diagramme im Peirce'schen Sinne. Es ist dann besonders auffallend, dass gerade diese Art von Informationsdarstellung einen so zentralen Platz in den *Digital Humanities* einnimmt. Obwohl dies für die Naturwissenschaften (wenn überhaupt) vielleicht überzeugend ist, sehen wir in der oft betonten Objektivität der Visualisierung auf Basis von Daten, wenn auch Massendaten, das fast beiläufige Ausschließen von hermeneutischen Interpretationsschritten, die die DH von anderen Umgangsweisen mit dem Digitalen unterscheiden sollte. Diese Überbetonung einer Datendarstellung ohne Interpretation ist natürlich selbst in den DH in die Kritik geraten. Dass eine Interpretation jeder Visualisierung nötig ist, wird nicht bestritten. Was aber hier als problematisch angesehen wird, ist, dass die Modelle, die Paradigmen, das Selbstverständnis des Feldes usw. so wenig darüber zu sagen haben, wie die Praxen der DH und ihre Visualisierungen in hermeneutische Diskussionen und Argumentationen einzubringen sind. Clement merkt treffend dazu an:

most critiques of DH – both from the social sciences and the humanities – do not point out a lack of accuracy, variability, or other limitation of method. Rather, most critiques of DH point to a decoupling of method from the theoretical perspectives that would ordinarily help situate the kind of intellectual effort being engaged (Clement 2016: 158).

In ähnlicher Weise kritisiert Hall (2013: 794) den Datenvisualisierungsansatz von Manovich u.ä. wegen seiner Marginalisierung gerade der intellektuellen Tätigkeit, die für die Geisteswissenschaften, egal ob digital oder nicht, definitorisch ist (vgl. aber weitere semiotische Diskussion von Dondero 2017). Eine fehlende Auseinandersetzung mit interpretatorischen Verfahren stuft Hall als besonders kritisch ein:

For if we do not explicitly do theory – because we either think we have left it behind or relegated it to some as yet unspecified point in the future – we do not end up *not* doing theory. Every methodology contains theory (Hall 2013: 798; Hervorhebung im Original).

Die Einbettung von Repräsentationen und Modellen in Peirces Semiotik, die hier vorgeschlagen wird, führt uns da weiter. Obwohl Diagramme für Wissen und Erkenntnisse unabdingbar sind, muss für Interpretationen, die nicht nur lediglich Umformatierungen von Daten bleiben wollen, der dritte Subtyp von Ikonen herangezogen werden: die Metapher. Das aktive Zusammenbringen von zwei Domänen, das die Wirkungsweise von Metaphern ausmacht, ist stets ein kreativer, aktiver und zweckorientierter Versuch, Ordnung zu schaffen. Wenn ein Objekt mit Hilfe einer Metapher betrachtet wird, sieht man es so, als ob die Strukturen und Eigenschaften der zweiten Domäne (in gewisser Hinsicht) auch für die Eigenschaften und das Verhalten des ursprünglichen Objekts gälten. Daher ‚weiß‘ man jetzt mehr über das Objekt. Dies ist ein ‚mehr‘, das nicht allein aus den Daten ableitbar ist. Hierbei gibt es keine kausale Beziehung, lediglich einen abduktiven Sprung, der Beobachtungen erklärbar zu machen scheint – und abduktive Sprünge sind genau die Mechanismen, die von der Diskurs-Schicht eines semiotischen Modus beigesteuert werden.

Ciula und Eide (2017) versuchen ihre Argumentation zu einem gewissen Grad in dieser Richtung zu entfalten. Sie schlagen vor, dass es in den DH drei Arten von Modellen gibt, die sich entlang ihres angenommenen Image-Diagramm-Metapher-Kontinuums arrangieren lassen. Aber diese Reduktion auf eine einzige Dimension der Komplexität anstelle der hier favorisierten kategorialen Dreiteilung macht präzise Analysen schwierig, sogar bei ihren eigenen multimodalen Analysen. Erstens sind keine klaren Entscheidungskriterien für die Zuordnung von Modellen zu den Subtypen von ikonischen Zeichen vorhanden, weil die Diskussion auf einer eher informellen Ebene von relativen Komplexitätsabstufungen bleibt; und zweitens haben die Beispiele selbst eine reichhaltige multimodale Binensstruktur, die dann teilweise unterschiedliche Ikonentypen-Zuordnungen erfor-

dern würde, weil unterschiedliche semiotische Modi für die Analyse herangezogen werden müssten. Als einfachstes Beispiel, d.h., ein Beispiel, das ein ‚Image-like‘ Modell illustrieren soll, betrachten Ciula und Eide ein an sich sehr interessantes Werkzeug aus der digitalen Paläografie, das erlaubt, Schriftformen aus unterschiedlichen mittelalterlichen Manuskripten mit Hilfe einer Benutzerschnittstelle so zu transformieren, dass sie visuell und dynamisch mit anderen Instanzen von Schriftformen geographisch und zeitlich verglichen werden können. Das ganze ‚Modell‘ wird dann als ‚Image-like‘ eingestuft, weil der/die Benutzer\*in des Werkzeugs tatsächlich eine Schriftform sieht, die anderen Schriftformen visuell ähnelt.

Wie eine genauere Betrachtung der involvierten semiotischen Modi deutlich gemacht hätte, sind aber die ganze Gebrauchssituation und das Werkzeug selbst viel komplexer als das, was ein ‚Image‘-Modell wiedergeben kann. Die interne Darstellung von Schriftformmerkmalen ist (nicht visuell) diagrammatisch, da ihr eine ausgewählte Menge von formalisierten Schrifteigenschaften zu Grunde liegt; die externe Darstellung ist auch diagrammatisch (aber in diesem Fall visuell), weil es diese abstrahierten Formen visuell wiedergibt. Dazu kommt der vom Werkzeug unterstützte visuelle Effekt einer kontinuierlichen Transformation zwischen Formen, die noch komplexer ist, da es eine weitere dynamische Materie (und daher weitere semiotische Modi) bereitstellt. Dass die vom System angebotene Transformation einer Schriftform vielleicht etwas mit der zeitlichen Entwicklung oder der geographischen Verbreitung dieser Form zu tun hat, ist aber auf jeden Fall eine *Diskursinterpretation*. Die interessantesten Aspekte des gegebenen Beispiels resultieren daher aus *metaphorischen* Modellierungsprozessen, die in der Lage sind, *mehr* Information durch Diskursinferenzen hineinzuimportieren, als die, die aus den Daten allein kommen könnte. Beschreibungen von Systemen dieser Art sind erst möglich, wenn, wie im vorigen Abschnitt eingeführt, die mitwirkenden semiotischen Modi mitgedacht werden, um die zur Verfügung stehenden Diskursinferenzen zu explizieren.

Ähnliche Vorbehalte gelten für die anderen Beispiele aus Ciula und Eide: alle könnten genauer dargestellt werden, wenn erst die zutreffenden semiotischen Modi auseinandergehalten und dann die jeweiligen Quellen von Informationsschlüssen spezifiziert würden. Eine wie die hier vorgeschlagene semiotische Einbettung würde diese Situation explizit machen und darüber hinaus eine weitere und präzisere Selbstreflexion bezüglich dessen, was genau gemacht worden ist, erfordern. Eine solche semiotisch untermauerte Selbstreflexion wäre für die DH im Allgemeinen von großem Vorteil, weil Modelle in den DH mehr als Diagramme sein sollten. Ein DH-Modell sollte sogar eher *nicht* ein Diagramm sein, denn dies ist die Rolle, die bereits von datengetriebenen Visualisierungen erfüllt wird. Ein DH-Modell sollte eine Peirce'sche Metapher sein, mit deren Hilfe *diskursiv* durch die verwendeten semiotischen Modi neue Eigenschaften zu den modellierten hinzugefügt werden können.

#### 4. Fazit und Ausblick

Entgegen ihrem hohen Anspruch sind die DH meines Erachtens (noch) nicht inter- und transdisziplinär genug, weil die dafür notwendigen Verbindungsglieder zwischen Methoden und Theorien weitgehend fehlen. Es gelingt den DH selten, eine Rolle als Brückenfach für breit angelegte interdisziplinäre Arbeiten zu spielen, da versteckte Spezialisierungen entlang herkömmlicher Disziplinengrenzen immer noch zu spüren sind. Solche Grenzen verhindern die Kommunikation oder dienen als (wenn auch oft implizite) Abgrenzungen. Es bleibt offen, wie effektive und sich gegenseitig bereichernde Zusammenführungen aus verschiedenen projektartigen Einzelergebnissen zu konzipieren wären. Nur wenn wir komplexe Zusammenhänge behandeln, analytische Tiefe erreichen und Interpretationen in Studien und Methoden der DH durchführen, wie von Schnapp and Presner (2000) anvisiert, sind Erkenntnisgewinne zu erwarten.

In diesem Aufsatz wurde argumentiert, dass eine Semiotik, die sich auf Peirces Ansatz und auch auf jüngere Studien zu diversen Formen von Multimodalität gründet, viel zur Brückenfunktion und dem Selbstverständnis der DH beitragen könnte. Dies wurde durch Beispiele aus den Bereichen Druck- und Filmmedien gezeigt sowie mit einer eher theoretischen Verankerung der für die DH zentralen Grundbegriffe Visualisierung und Modellierung in einer Semiotik der Multimodalität begründet. Ohne ein solches semiotisches Fundament bleiben Arbeiten häufig unvermittelt neben Ergebnissen aus anderen und auch aus übergreifenden Untersuchungsfeldern stehen. Nur mit Hilfe einer angemessenen Betrachtung der multimodalen Semiotik ist es möglich, über diverse Medien und Modi hinweg gemeinsame Analysen durchzuführen. Darüber hinaus ist ein besonderes Gewicht auf multimodale Methoden gelegt worden, die zu einer tieferen hermeneutischen Interpretation führen können. Denn dies ist sicherlich als eine der Kernaufgaben der DH anzusehen. Mehr-Ebenen-Annotationsschemata, die über unterschiedlichen Abstraktionsniveaus operieren, bilden dabei eine Methode, die sukzessive Interpretationsschritte unterstützt und explizit macht.

Der potenzielle Wert solcher Verfahren für die DH ist gut aus der sehr angemessenen Kritik an nicht-interpretativen DH-Ansätzen von Ramsay (2003) zu ersehen. Ramsay zieht eine Parallele zwischen in Algorithmen versteckter Information einerseits und Sätzen aus Wittgensteins ‚Abstraktionsleiter‘ andererseits, die trotz möglicherweise inkorrektur Zwischenansagen immer noch zum Verstehen führen kann, weil die Leiter selbst weggeworfen werden darf, nachdem jemand „auf ihr hinaufgestiegen ist“ (Wittgenstein 1921: 6.54). Ramsay schreibt dann:

The computational substrate upon which algorithmic criticism rests, however, demands that one pay attention to the hidden details of pattern formation. Algorithmic criticism might indeed be conceived as an activity that seeks to scrutinize the discarded ladder (Ramsay 2003: 171).

Genau in diesem Sinne machen explizite Annotationsebenen die internen Schritte des Interpretationsverfahrens nicht nur sichtbar, sondern darüber hinaus für weitere Schlussfolgerungen verfügbar. Wenn Interpretationen durch digitale Methoden unterstützt werden und die Spuren der offengelegten ‚Abstraktionsleiter‘ sichtbar bleiben, könnten wir uns auf den Weg zu einer echten ‚maschinellen Hermeneutik‘ (vgl. Mohr, Wagner-Pacifici und Breiger 2015) machen, die ein theoretisch sowie philosophisch überzeugendes Feld der DH ausmachen könnte.

## Anmerkung

- \* Dieser Artikel wurde durch die Kommentare einer anonymen Rezensentin und der Herausgeber dieser Sonderausgabe sowie Renate Henschel (Bremen), Julia Nantke (Wuppertal) und Janina Wildfeuer (Bremen) wesentlich verbessert. Danke an alle.

## Literatur

- Altman, Rick (1984). A Semantic/Syntactic Approach to Film Genre. *Cinema Journal* 23, 3, 6–18.
- Asher, Nicholas und Alex Lascarides (1994). Intentions and Information in Discourse. In: James Pustejovsky (ed.). *32nd. Annual Meeting of the Association for Computational Linguistics*. San Francisco: Morgan Kaufmann, 34–41. URL: <http://citeseer.ist.psu.edu/98759.html>.
- Avarado, Rafael und Paul Humphreys (2017). Big data, thick mediation, and representational opacity. *New Literary History* 48, 4, 729–749.
- Bánski, Piotr (2010). Why TEI stand-off annotation doesn't quite work: and why you might want to use it nevertheless. In: *Proceedings of Balisage: The Markup Conference*. DOI:10.4242/BalisageVol5.Banski01.
- Bateman, John A. (2007). Towards a grande paradigmatique of film: Christian Metz reloaded. *Semiotica* 167, 1–4, 13–64.
- Bateman, John A. (2008). *Multimodality and Genre: a foundation for the systematic analysis of multimodal documents*. Basingstoke: Palgrave Macmillan.
- Bateman, John A. (2011). The Decomposability of Semiotic Modes. In: Kay L. O'Halloran und Bradley A. Smith (eds.). *Multimodal studies: Multiple approaches and domains*. London: Routledge, 17–38.
- Bateman, John A. (2016). Methodological and theoretical issues for the empirical investigation of multimodality. In: Nina-Maria Klug und Hartmut Stöckl (eds.). *Handbuch Sprache im multimodalen Kontext*. Berlin: Mouton de Gruyter, 36–74.
- Bateman, John A. (2017). Triangulating transmediality: a multimodal semiotic framework relating media, modes and genres. *Discourse, Context & Media* 20, Supplement C, 160–174. DOI: [doi.org/10.1016/j.dcm.2017.06.009](https://doi.org/10.1016/j.dcm.2017.06.009).
- Bateman, John A. (2018). Peircean Semiotics and Multimodality: Towards a New Synthesis. *Multimodal Communication aop*. DOI: <https://doi.org/10.1515/mc-2017-0021>.

- Bateman, John A., Annika Beckmann und Rocío Varela (2018). From Empirical Studies to Visual Narrative Organization: Exploring Page Composition. In: Alexander Dunst, Jochen Laubrock und Janina Wildfeuer (eds.). *Empirical Comics Research. Digital, Multimodal, and Cognitive Methods*. London und New York: Routledge, 127–153.
- Bateman, John A., Judy L. Delin und Renate Henschel (2004). Multimodality and empiricism: preparing for a corpus-based approach to the study of multimodal meaning-making. In: Eija Ventola, Cassily Charles und Martin Kaltenbacher (eds.). *Perspectives on Multimodality*. Amsterdam: John Benjamins, 65–87.
- Bateman, John A. und Karl-Heinrich Schmidt (2012). *Multimodal Film Analysis: How Films Mean*. London: Routledge.
- Bateman, John A., Chiao-I Tseng, Ognyan Seizov, Arne Jacobs, Andree Lüdtkke, Marion G. Müller und Otthein Herzog (2016). Towards next-generation visual archives: image, film and discourse. *Visual Studies* 31, 2, 131–154. DOI: <http://dx.doi.org/10.1080/1472586X.2016.1173892>.
- Bateman, John A. und Francisco O. D. Veloso (2013). The Semiotic Resources of Comics in Movie Adaptation: Ang Lee's *Hulk* (2003) as a Case Study. *Studies in Comics* 4, 1, 137–159.
- Bateman, John A., Francisco O.D. Veloso, Janina Wildfeuer, Felix HiuLaam Cheung und Nancy Songdan Guo (2017a). An open multilevel classification scheme for the visual layout of comics and graphic novels: motivation and design. *Journal of Digital Scholarship in the Humanities* 32, 3, 476–510. DOI: <http://dx.doi.org/10.1093/llc/fqw024>.
- Bateman, John A. und Janina Wildfeuer (2014). A multimodal discourse theory of visual narrative. *Journal of Pragmatics* 74, 180–218. <http://dx.doi.org/10.1016/j.pragma.2014.10.001>.
- Bateman, John A., Janina Wildfeuer und Tuomo Hiippala (2017b). *Multimodality – Foundations, Research and Analysis. A Problem-Oriented Introduction*. Berlin: Mouton de Gruyter.
- Bauer, Matthias und Christoph Ernst (2010). *Diagrammatik: Einführung in ein kultur- und medienwissenschaftliches Forschungsfeld*. Bielefeld: Transcript.
- Benini, Sergio, Michele Svanera, Nicola Adami, Riccardo Leonardi und András Bálint Kovács (2016). Shot scale distribution in art films. *Multimedia Tools and Applications* 75, 23, 16499–16527. DOI: <https://doi.org/10.1007/s11042-016-3339-9>.
- Berry, David M. (2011). The computational turn: Thinking about the digital humanities. *Culture Machine* 12. URL: <https://www.culturemachine.net/index.php/cm/article/viewArticle/440>
- Berry, David M. (ed.) (2012). *Understanding digital humanities*. Houndmills: Palgrave Macmillan.
- Bertin, Jacques (1983). *Semiology of graphics*. Madison, Wisconsin: University of Wisconsin Press. Translated *Sémiologie graphique* (1967) by William J. Berg.
- Boehm, Gottfried (2004). Jenseits der Sprache? Anmerkungen zur Logik der Bilder. In: Christa Maar und Hubert Burda (eds.). *Iconic Turn. Die neue Macht der Bilder*. Köln: DuMont, 28–43.
- Branigan, Edward (1984). *Point of view in the cinema: a theory of narration and subjectivity in classical film*. Berlin: Mouton.

- Bubenhöfer, Noah und Joachim Scharloth (2015). Maschinelle Textanalyse im Zeichen von Big Data und Data-Driven Turn – Überblick und Desiderata. *Zeitschrift für germanistische Linguistik* 43, 1, 1–26.
- Bucher, Hans-Jürgen (2011). Multimodales Verstehen oder Rezeption als Interaktion. Theoretische und empirische Grundlagen einer systematischen Analyse der Multimodalität. In: Hans-Joachim Diekmannshenke, Michael Klemm und Hartmut Stöckl (eds.). *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. Berlin: Erich Schmidt, 123–156.
- Burdick, Anne, Johanna Drucker, Peter Lunenfeld, Todd Presner und Jeffrey Schnapp (eds.) (2012). *Digital\_humanities*. Cambridge, MA und London: MIT Press.
- Canini, Luca, Sergio Benini und Riccardo Leonardi (2011). Affective analysis on patterns of shot types in movies. In: *Proceedings of 7th International Symposium on Image and Signal Processing and Analysis (ISPA)*. Dubrovnik: IEEE, 253–258.
- Ciula, Arianna und Øyvind Eide (2017). Modelling in digital humanities: signs in context. *Journal of Digital Scholarship in the Humanities* 32 (suppl. 1), i33–i46. DOI: 10.1093/llc/fqw045.
- Clement, Tanya E. (2016). Where is Methodology in Digital Humanities? In: Matthew K. Gold und Lauren F. Klein (eds.). *Debates in the digital humanities*. Minneapolis: University of Minnesota Press, 153–175.
- Cohn, Neil (2013). *The Visual Language of Comics: Introduction to the structure and cognition of sequential images*. London: Bloomsbury.
- Deppermann, Arnulf (2013). Multimodal interaction from a conversation analytic perspective. *Journal of Pragmatics* 46, 1, 1–7.
- Dondero, Maria Giulia (2017). The semiotics of design in media visualization: Mereology and observation strategies. *Information Design Journal* 23, 2, 208–218.
- Drucker, Johanna (2014). *Graphesis: Visual Forms of Knowledge Production*. Cambridge, Massachusetts und London: Harvard University Press.
- Drucker, Johanna (2017). Digital Humanities als epistemische Praxis. *Zeitschrift für Medienwissenschaft* 16, 1, 114–124. Im Gespräch mit Annika Haas.
- Eco, Umberto (1981). Der Einfluß Roman Jakobsons auf die Entwicklung der Semiotik. In: Martin Krampen, Klaus Oehler, Roland Posner und Thure von Uexküll (eds.). *Die Welt als Zeichen: Klassiker der modernen Semiotik*. Berlin: Wolf Jobst Siedler, 173–204.
- Eide, Øyvind (2015). *Media Boundaries and Conceptual Modelling*. Basingstoke: Palgrave Macmillan.
- Elleström, Lars (2013). Spatiotemporal aspects of iconicity. In: Lars Elleström, Olga Fischer und Christina Ljungberg (eds.). *Iconic investigations*. Amsterdam: John Benjamins, 95–120.
- Elleström, Lars (2014). Media transformation: the transfer of media characteristics among media. Basingstoke: Palgrave Macmillan.
- Evans, Leighton und Sian Rees (2012). An interpretation of digital humanities. In: David M. Berry (ed.). *Understanding digital humanities*. Houndmills: Palgrave Macmillan, 21–41.
- Farias, Priscila L. (2005). Images, diagrams and metaphors: a contribution from semiotics to information design. In: Carla G. Spinillo und Solange G. Coutinho (eds.).

- Selected Readings of the Information Design International Conference 2003*. Recife, Brasil: SBDI – Sociedade Brasileira de Design da Informação, 71–78.
- Forceville, Charles J. (1999). Educating the eye? Kress and Van Leeuwen's *Reading Images: The Grammar of Visual Design* (1996). *Language and Literature* 8, 2, 163–178. DOI: <https://doi.org/10.1177/096394709900800204>.
- Forceville, Charles J. (2007). Book Review: *Multimodal Transcription and Text Analysis: A Multimedia Toolkit and Coursebook* by Anthony Baldry and Paul J. Thibault. *Journal of Pragmatics* 39, 6, 1235–1238.
- Frank, Ingo (in diesem Heft). Diagrammatische Denkwerkzeuge in den Digital Humanities – Ansatz zur zeichentheoretischen Grundlegung. *Zeitschrift für Semiotik* 39, 1–2, 51–81.
- Fricke, Ellen (2012). *Grammatik multimodal: Wie Wörter und Gesten zusammenwirken*. Berlin und New York: Mouton de Gruyter.
- Gibbons, Alison (2011). *Multimodality, Cognition, and Experimental Literature*. Routledge Studies in Multimodality. London: Routledge.
- Gitelman, Lisa (ed.) (2013). *“raw data” is an oxymoron*. Cambridge, MA: MIT Press.
- Gold, Matthew K. und Lauren F. Klein (eds.) (2016). *Debates in the digital humanities*. Minneapolis: University of Minnesota Press.
- Grant, August E. und Jeffrey S. Wilkinson (eds.) (2009). *Understanding media convergence: the state of the field*. New York und Oxford: Oxford University Press.
- Haddington, Pentti, Lorenza Mondada und Maurice Nevile (eds.) (2013). *Interaction and Mobility: Language and the Body in Motion*. Berlin: De Gruyter.
- Hall, Gary (2013). Toward a Postdigital Humanities: Cultural Analytics and the Computational Turn to Data-Driven Scholarship. *American Literature* 85, 4, 781–809.
- Hayles, N. Katherine (2003). Translating media: why we should rethink textuality. *The Yale Journal of Criticism* 16, 2, 263–290.
- Heftberger, Adelheid (2015). „Die Verschmelzung von Wissenschaft und Filmchronik“ – Das Potential der reduktionslosen Visualisierung von *Das elfte Jahr* und *Der Mann mit der Kamera* von Dziga Vertov. *Chronos, Geschichte und Informatik – Histoire et Informatique* 18–19, 229–263. DOI: <http://dx.doi.org/10.17613/M6F22M>.
- Hiippala, Tuomo (2015). *The structure of multimodal documents: An empirical approach*. London: Routledge.
- Hiippala, Tuomo (2016). Semi-automated annotation of page-based documents within the Genre and Multimodality framework. In: *Proceedings of the 10th sighthum workshop on language technology for cultural heritage, social sciences, and humanities*. Berlin: Association for Computational Linguistics, 84–89.
- Hjelmslev, Louis (1974 [1943]). *Prolegomena zu einer Sprachtheorie*. München: Hueber.
- Holly, Werner (2009). Der Wort-Bild-Reißverschluss. Über die performative Dynamik der audiovisuellen Transkriptivität. In: Helmuth Feilke und Angelika Linke (eds.). *Oberfläche und Performanz*. Tübingen: Niemeyer, 93–110.
- Jakobson, Roman (1971). *Roman Jakobson: Selected Writings*, Bd. II: Word and Language chap. Language in relation to other communicative systems. The Hague: Mouton, 570–579.
- Jenkins, Henry (2008). *Convergence Culture: Where Old and New Media Collide*. New York: NYU Press.

- Jewitt, Carey und Gunther Kress (2003). *Multimodal literacy*. Frankfurt a.M. und New York: Peter Lang.
- Kapuścińska, Anna (2017). *Grenzphänomene zwischen Text und Bild am Beispiel multimedialer Nachrichtensendungen*. Berlin: Erich Schmidt.
- Kendon, Adam (1980). Gesture and speech: two aspects of the process of utterance. In: Mary R. Key (ed.). *Nonverbal communication and language*. The Hague: Mouton, 207–227.
- Kloepfer, Rolf (1977). Komplementarität von Sprache und Bild. Am Beispiel von Comic, Karikatur und Reklame. In: Roland Posner und Hans-Peter Reinecke (eds.). *Zeichenprozesse. Semiotische Forschung in den Einzelwissenschaften*. Wiesbaden: Athenäum, 129–145.
- Kralemann, Björn und Claas Lattmann (2013). Models as icons: modeling models in the semiotic framework of Peirce's theory of signs. *Synthese* 190, 16, 3397–3420.
- Krämer, Sybille (2003). Erfüllen Medien eine Konstitutionsleistung? Thesen über die Rolle medientheoretischer Erwägungen beim Philosophieren. In: Stefan Münker, Alexander Roesler und Mike Sandbothe (eds.). *Medienphilosophie: Beiträge zur Klärung eines Begriffs*. Frankfurt a.M.: Fischer, 78–90.
- Krämer, Sybille (2008). *Medium, Bote, Übertragung. Kleine Metaphysik der Medialität*. Frankfurt a.M.: Suhrkamp.
- Krämer, Sybille (2014). Zur Grammatik der Diagrammatik. Eine Annäherung an die Grundlagen des Diagrammgebrauches. *Zeitschrift für Literaturwissenschaft und Linguistik* 44, 176, 11–30.
- Kress, Gunther (2010). *Multimodality: a social semiotic approach to contemporary communication*. London: Routledge.
- Kress, Gunther (2014). What is mode? In: Carey Jewitt (ed.). *The Routledge Handbook of multimodal analysis*. 2. Auflage London: Routledge, 60–75.
- Kress, Gunther und Theo van Leeuwen (1990). *Reading Images. Sociocultural Aspects of Language and Education*. Geelong, Victoria, Australia: Deakin University Press.
- Kress, Gunther und Theo van Leeuwen (2006 [1996]). *Reading Images: The Grammar of Visual Design*. London und New York: Routledge.
- Kwastek, Katja (2015). Vom Bild zum Bild – Digital Humanities jenseits des Textes. In: Constanze Baum und Thomas Stäcker (eds.). *Grenzen und Möglichkeiten der Digital Humanities*, Sonderband 1 der *Zeitschrift für digitale Geisteswissenschaften*. DOI: [https://doi.org/10.17175/sb001\\_002](https://doi.org/10.17175/sb001_002).
- Lee, Jennifer C. und Santosh Khadka (eds.) (2018). *Designing and implementing multimodal curricula and programs*. London und New York: Routledge.
- Liu, Alan (2013). The Meaning of the Digital Humanities. *PMLA* 128, 2, 409–423.
- Manovich, Lev (2012). How to compare one million images. In: David M. Berry (ed.). *Understanding digital humanities*. Basingstoke: Palgrave Macmillan, 249–278.
- Manovich, Lev (2018). 100 Billion Data Rows per Second: Media Analytics in the Early 21st Century. *International Journal of Communication* 12, 473–488.
- Martin, James R. (1992). *English text: systems and structure*. Amsterdam: John Benjamins.
- Martinec, Radan und Anthony Salway (2005). A system for image-text relations in new (and old) media. *Visual Communication* 4, 3, 337–371.
- McCarty, Willard (2005). *Humanities computing*. Basingstoke: Palgrave.

- McNeill, David (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago, Illinois: University of Chicago Press.
- Meister, Jan Christoph (2007). Computational Narratology oder: Kann man das Erzählen berechenbar machen? In: Corinna Müller und Irina Scheidgen (eds.). *Mediale Ordnungen. Erzählen, Archivieren, Beschreiben*. Marburg: Schüren Verlag, 19–39.
- Mittelberg, Irene (2006). *Metaphor and metonymy in language and gesture: Discourse evidence for multimodal models of grammar*. Ann Arbor, MI: UMI. Cornell University Dissertation.
- Mittelberg, Irene (2008). Peircean semiotics meets conceptual metaphor. Iconic modes in gestural representations of grammar. In: Alan Cienki und Cornelia Müller (eds.). *Metaphor and Gesture*. Amsterdam und Philadelphia: John Benjamins, 115–154.
- Mohr, John W., Robin Wagner-Pacifici und Ronald L. Breiger (2015). Towards a computational hermeneutics. *Big Data & Society* 2, 2. DOI: <https://doi.org/10.1177/2053951715613809>.
- Mondada, Lorenza (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics* 65, 137–156.
- Moretti, Franco (2013). *Distant reading*. London und New York: Verso.
- Müller, Cornelia (1998). *Redebegleitende Gesten: Kulturgeschichte – Theorie – Sprachvergleich*. Berlin: Berliner Wissenschafts-Verlag.
- Nantke, Julia (in diesem Heft). Annäherungen an eine digitale Semiotik: Zwischen computergestützter Semiotik und Semiotik als Metadisziplin der Digitalen Literaturwissenschaften. *Zeitschrift für Semiotik* 39, 1–2, 83–108.
- Neale, Steve (2000). *Genre and Hollywood*. London: Routledge.
- New London Group (2000). A pedagogy of Multiliteracies designing social futures. In: Mary Kalantzis und Bill Cope (eds.). *Multiliteracies: Literacy Learning and the Design of Social Futures*. London: Routledge, chap. 1, 9–38.
- Norris, Sigrid (2016). Multimodal Interaction – Language and Modal Configurations. In: Nina-Maria Klug und Hartmut Stöckl (eds.). *Handbuch Sprache im multimodalen Kontext*. Berlin: Mouton de Gruyter, 121–144.
- Nöth, Winfried (2000). *Handbuch der Semiotik*. 2. vollständig neu bearbeitete und erweiterte Auflage. Stuttgart und Weimar: J.B. Metzler.
- Ochs, Elinor (1979). Transcription as Theory. In: Elinor Ochs und Bambi B. Schieffelin (eds.). *Developmental Pragmatics*. New York: Academic Press, 43–72.
- O'Halloran, Kay L. (1999). Interdependence, Interaction and Metaphor in Multisemiotic Texts. *Social Semiotics* 9, 3, 317–354.
- O'Halloran, Kay L. (2004). Visual semiosis in film. In: Kay L. O'Halloran (ed.). *Multimodal discourse analysis: systemic functional perspectives*. London: Continuum, 109–130.
- O'Toole, Michael (2011 [1994]). *The language of displayed art*. Abingdon, Oxon: Routledge.
- Peirce, Charles Sanders (1998 [1893–1913]). *The Essential Peirce – Volume 2. Selected Philosophical Writings (1893–1913)*. Bloomington: Indiana University Press.
- Posner, Roland (1986). Zur Systematik der Beschreibung verbaler und nonverbaler Kommunikation: Semiotik als Propädeutik der Medienanalyse. In: Hans-Georg Bosshardt (ed.). *Perspektiven auf Sprache: Interdisziplinäre Beiträge zum Gedenken an Hans Hörmann*. Berlin und New York: De Gruyter, 267–313.

- Ramsay, Stephen (2003). Towards an algorithmic criticism. *Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing* 18, 2, 167–174. DOI: <https://doi.org/10.1093/lc/18.2.167>.
- Ravelli, Louise (2006). *Museum Texts: Communication Frameworks*. London und New York: Routledge.
- Ravelli, Louise J. und Robert J. McMurtrie (2016). *Multimodality in the Built Environment: Spatial Discourse Analysis*. New York und London: Routledge.
- Renear, Allen (1997). Out of praxis: three (meta)theories of textuality. In: Kathryn Sutherland (ed.). *Electronic text: investigations in method and theory*. Oxford: Clarendon Press, 107–126.
- Sachs-Hombach, Klaus (2003). *Das Bild als kommunikatives Medium: Elemente einer allgemeinen Bildwissenschaft*. Univ. Habilitationsschrift. Magdeburg. Köln: Halem.
- Sahle, Patrick (2015). Digital Humanities? Gibt's doch gar nicht! In: Constanze Baum und Thomas Stäcker (eds.). *Grenzen und Möglichkeiten der Digital Humanities*. Sonderband 1 der *Zeitschrift für digitale Geisteswissenschaften*. DOI: [https://doi.org/10.17175/sb001\\_004](https://doi.org/10.17175/sb001_004).
- Schnapp, Jeffrey und Todd Presner (2009). The Digital Humanities Manifesto 2.0. Online. URL: [http://www.humanitiesblast.com/manifesto/Manifesto\\_V2.pdf](http://www.humanitiesblast.com/manifesto/Manifesto_V2.pdf).
- Schreier, Margrit (2012). *Qualitative content analysis in practice*. London: Sage.
- Spillner, Bernd (1982). Stilanalyse semiotisch komplexer Texte. Zum Verhältnis von sprachlicher und bildlicher Information in Werbeanzeigen. *Kodikas/Code. Ars Semeiotica* 4/5, 1, 91–106.
- Stetter, Christian (2005). *System und Performanz. Symboltheoretische Grundlagen von Medientheorie und Sprachwissenschaft*. Weilerswist: Velbrück Wissenschaft.
- Stöckl, Hartmut (1992). Der 'picture relation type' – Ein praktischer Analysemodus zur Beschreibung der vielfältigen Einbettungs- und Verknüpfungsbeziehungen von Bild und Text. *Papiere zur Linguistik* 46, 11, 49–61.
- Stöckl, Hartmut (2004). In between modes: language and image in printed media. In: Eija Ventola, Cassily Charles und Martin Kaltenbacher (eds.). *Perspectives on Multimodality*. Amsterdam: John Benjamins, 9–30.
- Stöckl, Hartmut (2006). Zeichen, Text und Sinn – Theorie und Praxis der multimodalen Textanalyse. In: Eva Martha Eckkrammer und Gudrun Held (eds.). *Textsemiotik: Studien zu multimodalen Texten*. Frankfurt a.M.: Peter Lang, 11–36.
- Suchan, Jakob und Mehul Bhatt (2016). The geometry of a scene: On deep semantics for visual perception driven cognitive film studies. In: *IEEE Winter Conference on Applications of Computer Vision (WACV 2016)*. Lake Placid, NY: IEEE, 1–9. DOI: <https://doi.org/10.1109/WACV.2016.7477712>.
- Svensson, Patrik (2009). Humanities computing as digital humanities. *Digital Humanities Quarterly* 3, 3. URL: <http://digitalhumanities.org/dhq/vol/3/3/000065/000065.html>.
- Thomas, Martin (2014). Evidence and circularity in multimodal discourse analysis. *Visual Communication* 13, 2, 163–189. DOI: <https://doi.org/10.1177/1470357213516725>.
- Tseng, Chiao-I (2013). *Cohesion in Film: Tracking Film Elements*. Basingstoke: Palgrave Macmillan.
- Tseng, Chiao-I und John A. Bateman (2018). Cohesion in Comics and Graphic Novels: an empirical comparative approach to transmedia adaptation in *City of Glass*. *Adaptation* 11, 2, 122–143. DOI: <https://doi.org/10.1093/adaptation/apx027>.

- Tseng, Chiao-I, Jochen Laubrock und Jana Pflaeging (2018). Character Developments in Comics and Graphic Novels: A Systematic Analytical Scheme. In: Alexander Dunst, Jochen Laubrock und Janina Wildfeuer (eds.). *Empirical Comics Research. Digital, Multimodal, and Cognitive Methods*. London und New York: Routledge, 154–175.
- van Leeuwen, Theo (1985). Rhythmic structure of the film text. In: Teun A. van Dijk (eds.). *Discourse and communication: new approaches to the analysis of mass media discourse and communication*. Berlin: De Gruyter, 216–232.
- van Leeuwen, Theo (1991). Conjunctive structure in documentary film and television. *Continuum: journal of media and cultural studies* 5, 1, 76–114.
- van Leeuwen, Theo (1999). *Speech, Music, Sound*. London: MacMillan.
- Vanhoutte, Edward und Ron van den Branden (2010). Text Encoding Initiative (TEI). In: Marcia Bates und Mary Niles Maack (eds.). *Encyclopedia of Library and Information Sciences (ELIS)* (third edition). New York: Taylor & Francis, Kap. 512, 5172–5181. DOI: <https://doi.org/10.1081/E-ELIS3-120043748>.
- Warnke, Martin und Lisa Dieckmann (2016). Prometheus meets Meta-Image: implementations of Aby Warburg's methodical approach in the digital era. *Visual Studies* 31, 2, 109–120. DOI: <https://doi.org/10.1080/1472586X.2016.1173890>.
- Wharton, Tim (2009). *Pragmatics and Non-Verbal Communication*. Cambridge: Cambridge University Press.
- Wildfeuer, Janina (2013). Trompeten, Fanfaren und orangefarbene Tage. Zur Intersemiose in *Die fabelhafte Welt der Amélie*. In: Lars C. Grabbe, Patrick Rupert-Kruse und Norbert M. Schmitz (eds.). *Multimodale Bilder. Zur synkretistischen Struktur des Filmischen*. Darmstadt: Böhner-Verlag, 81–101.
- Wildfeuer, Janina (2014). *Film Discourse Interpretation. Towards a New Paradigm for Multimodal Film Analysis*. London und New York: Routledge.
- Winkler, Hartmut (2008). Zeichenmaschinen: oder warum die semiotische Dimension für eine Definition der Medien unerlässlich ist. In: Stefan Münker und Alexander Roesler (eds.). *Was ist ein Medium?* Frankfurt a.M.: Suhrkamp, 211–222.
- Wittgenstein, Ludwig (1921). Logisch-Philosophische Abhandlung. *Annalen der Naturphilosophie* 14, 185–262.
- Wu, Hui-Yin und Marc Christie (2016). Analysing Cinematography with Embedded Constrained Patterns. In: Marc Christie, Quentin Galvane, Arnav Jhala und Rémi Ronfard (eds.). *Eurographics Workshop on Intelligent Cinematography and Editing*. Lisbon: The Eurographics Association. DOI: <https://doi.org/10.2312/wiced.20161098>.
- Zeppelzauer, Matthias, Dalibor Mitrović und Christian Breiteneder (2011). Cross-modal analysis of audio-visual film montage. In: *Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN)*, 1–6. Maui, HI, USA: IEEE. DOI: <https://doi.org/10.1109/ICCCN.2011.6005782>.
- Zhang, Qingchen, Laurence T. Yang, Zhikui Chen und Peng Li (2018). A survey on deep learning for big data. *Information Fusion* 42, 146–157. DOI: <https://doi.org/10.1016/j.inffus.2017.10.006>.

*Prof. Dr. phil. John A. Bateman*  
*FB10 – Sprach- und Literaturwissenschaften*  
*Universität Bremen*  
*Bibliothekstr. 1*  
*D-28334 Bremen*  
*E-Mail: bateman@uni-bremen.de*