

Janina Wildfeuer; John A. Bateman

### Theoretische und methodologische Perspektiven des Multimodalitätskonzepts aus linguistischer Sicht

2018

<https://doi.org/10.25969/mediarep/16402>

Veröffentlichungsversion / published version  
Zeitschriftenartikel / journal article

#### Empfohlene Zitierung / Suggested Citation:

Wildfeuer, Janina; Bateman, John A.: Theoretische und methodologische Perspektiven des Multimodalitätskonzepts aus linguistischer Sicht. In: *IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft*. Heft 28, Jg. 14 (2018), Nr. 2, S. 5–46. DOI: <https://doi.org/10.25969/mediarep/16402>.

#### Erstmalig hier erschienen / Initial publication here:

[http://www.gib.uni-tuebingen.de/index.php?option=com\\_content&view=article&id=111&Itemid=157&menuitem=miArchive&showIssue=91](http://www.gib.uni-tuebingen.de/index.php?option=com_content&view=article&id=111&Itemid=157&menuitem=miArchive&showIssue=91)

#### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use:

This document is made available under a Deposit License (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual, and limited right for using this document. This document is solely intended for your personal, non-commercial use. All copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute, or otherwise use the document in public.

By using this particular document, you accept the conditions of use stated above.

Janina Wildfeuer/John A. Bateman

## **Theoretische und methodologische Perspektiven des Multimodalitätskonzepts aus linguistischer Sicht**

### **Abstract**

The article deals with fundamental and broadly discussed issues within the paradigm of multimodal analysis concerning how the intersemiotic interplay of semiotic resources constructs meaning. Taking a specifically linguistic focus on theories and methods within this paradigm, the article orientates particularly towards discourse analytical and text linguistic accounts which aim at a theoretically and methodologically fine-grained account of the notions of semiotic mode and medium. The central idea is that re-constructing semiotic modes as theoretically tightly interwoven bundles of material, form and dynamic discourse semantics provides a suitable foundation for fine-grained empirical analysis, which in turn allows a comprehensive examination and evaluation of the theoretical constructs proposed.

Der Beitrag beschäftigt sich mit grundlegenden, das Multimodalitätsparadigma seit seinen Anfängen begleitenden Fragen nach der Bedeutungskonstruktion im Zusammenspiel unterschiedlichster Zeichenressourcen und Modalitäten. Er wirft dabei einen spezifisch linguistischen Fokus auf Theorien und Methoden innerhalb dieses Paradigmas und orientiert sich insbesondere an diskursanalytischen und textlinguistischen Fragestellungen, die eine theoretische Ausdifferenzierung und methodologische Präzisierung der zentralen Begriffe Zeichenmodalität und Medium anstreben. Der Grundgedanke dabei ist,

dass Zeichenmodalitäten als eng verflochtene Bündel von Material, Form und dynamischer Diskurssemantik eine geeignete Grundlage für detaillierte empirische Analysen bieten, die letztlich eine umfassende Prüfung und Evaluation der theoretischen Ergebnisse und Forschung über das einzelne Artefakt hinaus ermöglicht.

## Einleitung

Die unter dem Schlagwort ›Multimodalität‹ verorteten Analysen unterschiedlichster Kommunikationsformen im Hinblick auf ihre semiotische Komplexität stellen derzeit eine der einflussreichsten Paradigmen der Geisteswissenschaften dar. Vor allem im Anschluss an die bahnbrechenden Arbeiten von Gunther Kress und Theo van Leeuwen in den 1990er und frühen 2000er Jahren (KRESS/VAN LEEUWEN 1996, 2001; vgl. auch VAN LEEUWEN 2005 und KRESS 2010) entwickeln sich in zahlreichen Disziplinen unterschiedlichste Arbeiten theoretischer, methodologischer sowie analytischer Natur, die alle das Ziel verfolgen, multimodale, d.h. aus mehreren Zeichen(-ressourcen) bestehende Artefakte im Hinblick auf ihre Bedeutungskonstruktion zu untersuchen (vgl. BUCHER 2007: 49).

Die Vielzahl an Zugängen zu diesen Artefakten macht sich unter anderem in der Zahl der Sammelbände und Handbücher bemerkbar, die in den letzten Jahren erschienen sind und weiterhin erscheinen (vgl. zum Beispiel VENTOLA/CHARLES/KALTENBACHER 2004; ECKKRAMMER/HELD 2006; MACHIN 2007; ROYCE/BOWCHER 2007; JEWITT 2009; O'HALLORAN/SMITH 2011; MÜLLER et al. 2013, 2014; JEWITT 2014; MAIORANI/CHRISTIE 2014; NORRIS/MAIER 2014; WILDFEUER 2015; NORRIS 2015; KLUG/STÖCKL 2016; JEWITT 2016; SEIZOV/WILDFEUER 2017; BATEMAN et al. 2017a), spiegelt zugleich aber die enorme Diversität und in großem Maße auch Uneinigkeit innerhalb dieser Ansätze wider. Verschiedenste Ansichten über die Definitionen von Modalität und Medialität sowie über methodologische Herangehensweisen stehen konkurrierend gegenüber (vgl. SCHNEIDER/STÖCKL 2011: 24-25) und besonders im Vergleich einzelner philologischer Ansätze (beispielsweise im Rahmen germanistischer Linguistik einerseits und einer breiteren anglophonen Ausrichtung andererseits; siehe unten) wird die unterschiedliche Fokussierung deutlich.

Mit einem spezifisch linguistischen Fokus und den damit einhergehenden diskursanalytischen und textlinguistischen Fragestellungen will sich der vorliegende Beitrag beschäftigen und dabei die dem Multimodalitätsparadigma seit seinen Anfängen begleitenden, grundlegenden Fragen adressieren: Welche Bedeutungen entstehen im Zusammenspiel unterschiedlichster Zeichenressourcen und Modalitäten und wie arbeiten die vielfältigen Mechanismen im großen Prozess der Bedeutungskonstruktion zusammen? Hierbei gilt es nicht allein nach Bedeutungsrelationen zwischen verbalen Texten und Bildern zu fragen, wie es die Sprachwissenschaft seit den 1970er Jahren oft

verfolgt (vgl. BARDIN 1975; MASSEY 1980; SPILLNER 1982; TITZMANN 1990; HARTMANN 2002; NÖTH 2004 u.v.a.), sondern diejenigen zeitgenössischen Ansätze vorzustellen, die es erlauben, neben einer theoretischen Ausdifferenzierung und methodologischen Präzisierung einiger zentraler Begriffe auch die notwendige empirische Analyse voranzutreiben, die letztlich eine umfassende Prüfung und Evaluation der theoretischen Ergebnisse und Forschung über das einzelne Artefakt hinaus ermöglicht.

Der explizit linguistische Fokus ergibt sich im Rückgriff auf solche Ansätze, die sich in vielen Fällen zunächst mit der Analyse verbaler Texte bzw. Diskurse auseinandersetzen und ihre Mechanismen zur Bedeutungskonstruktion dann auf die Interaktion unterschiedlichster Modalitäten übertragen, um nach ihrer Funktion und Rolle innerhalb dieses Prozesses zu fragen. Während oft eine Unterscheidung zwischen semiotischen Herangehensweisen einerseits und text- und diskurslinguistischen Ansätzen andererseits getroffen wird (vgl. z. B. KLUG/STÖCKL 2015), wollen wir im Folgenden diese Trennung gerade nicht aufrecht erhalten, sondern eine effektive Verbindung angewandter Semiotik- und Text- bzw. Diskurstheorien vorstellen, die zugleich ein notwendiges Resultat aus der Beschäftigung mit Medienartefakten und somit dem Gebrauch unterschiedlichster semiotischer Ressourcen darstellt. Als besonders vielversprechend und für die Analyse von komplexen semiotischen Artefakten besonders geeignet sehen wir hier vor allem aktuelle diskurssemantische Arbeiten in Kombination mit durchaus traditionellen semiotischen Theorien, deren Vorteile wir im Laufe der Diskussion aufzeigen wollen.

Hierfür werden wir zunächst einen kurzen Überblick über die oft unter dem Schlagwort ›multimodale Linguistik‹ – so diskutabel der Begriff in seinen terminologischen Widersprüchlichkeiten auch sein mag – zusammengefassten Ansätze liefern, der zum einen die im internationalen Kontext längst verankerten Multimodalitätstheorien linguistischer Natur kurz vorstellt, zum anderen die in der deutschen Wissenschaftslandschaft sich erst langsam etablierenden Ansätze aufzeigt. In einem weiteren wichtigen Schritt gilt es dann, terminologische und definitorische Unterscheidungen zu treffen, die eine Präzisierung des mitunter wichtigsten Konzeptes im Multimodalitätsparadigma, dem Konzept des *mode* oder der Modalität, mit sich bringen und auch eine Abgrenzung zum Medienbegriff leisten können. Hier fehlt es vor allem an jenen Ausdifferenzierungen, die die bestehende große Lücke zwischen vorhandenen theoretischen und empirischen Ansätzen füllen und zugleich Schwächen in der methodologischen Herangehensweise ausgleichen können. Eine erste beispielhafte Anwendung und Übertragung dieser Präzisierungen werden wir mit Blick auf ein innerhalb der Comicforschung bereits viel zitiertes Beispiel vornehmen.

## 1. Multimodalität und Linguistik – in Deutschland und anderswo

Die Entwicklung des Multimodalitätsparadigmas geht nach van Leeuwen (2011: 541) bereits auf psychologische Überlegungen in den 1920er Jahren über wechselseitige Effekte unserer sensorischen Wahrnehmung zurück. Ein zunächst unabhängiges, linguistisches und kommunikationstheoretisches Interesse entstand zeitgleich durch Entwicklungen und Neuerungen im Kontext öffentlicher Kommunikation, wie z. B. durch den Aufstieg von Film und Fernsehen sowie Veränderungen im Schrift- und Druckstil hin zu mehr Illustrationen und dem Gebrauch von Bildern. Dass die Linguistik hiervon Notiz nahm, ist vor allem eine Folge innerdisziplinärer Entwicklungen hin zu einer Linguistik des Textes oder Diskurses (vgl. HARRIS 1959; SPITZMÜLLER/WARNKE 2011), die alle Bestandteile im Zusammenspiel betrachtet und dabei sehr schnell feststellt, dass Sprache nie alleine vorkommt und immer in Kombination zu Mimik und Gestik oder Layout und Interpunktion verstanden werden muss.

Neben den vier von van Leeuwen beschriebenen linguistischen Ausrichtungen (u.a. auch die frühe Prager und Pariser Schule, vgl. VAN LEEUWEN 2011), die sich seit den 1920er Jahren entwickelten und neben Sprache auch andere Modalitäten in den Fokus ihrer Untersuchung stellten, darf auch das bereits von Peirce verfolgte Interesse an nicht-sprachlichen Zeichensystemen nicht unerwähnt bleiben. Seine Überlegungen zur Differenzierung unterschiedlicher sensorischer Systeme und Eindrücke und deren Analyse in Verbindung mit sprachlichen Einheiten (vgl. PEIRCE 1931-1958: §§2.275-2.308) stellen einen wichtigen Ausgangspunkt für semiotische und linguistische Ansätze im Multimodalitätsparadigma dar. Dass dieser allerdings nicht überall und in gleichem Maße Geltung findet, ist ein interessanter, vor allem aber höchst diskutabler Unterschied in der Entwicklung der verschiedenen Ansätze: Zwar auch im internationalen Kontext immer noch als »very new and [...] very creative endeavour« (VAN LEEUWEN 2014: 22) bezeichnet, hat sich das Thema »Multimodalität« in den letzten 25 Jahren doch bereits eindeutig als Forschungsrichtung der anglophonen Sprach- und Kommunikationswissenschaft etabliert. Neben zahlreichen internationalen Konferenzen, den bereits genannten Sammelbänden, Buchserien (z. B. die Routledge Studies in Multimodality) sowie einer expliziten Ausrichtung bzw. Neueinrichtung unterschiedlicher Zeitschriften (z. B. Visual Communication sowie Multimodal Communication oder das noch junge Journal of Multimodal Communication Studies, das an der Warschauer Universität angesiedelt ist), zeugen auch viele Bachelor- und Masterprogramme sowie die Gründung einzelner Forschungszentren (das MODE Research Centre am Institute of Education am University College London, das Multimodal Analysis Lab an der Universität in Singapur, das Multimodal Research Centre an der Auckland University of Technology) von großem Interesse und differenzierter akademischer Auseinandersetzung mit der Thematik.

In Deutschland sieht die Situation tatsächlich anders aus: Zwar lassen sich in den letzten fünf bis zehn Jahren vermehrt Interesse an und

Aufmerksamkeit für multimodale Analysen erkennen und auch die Zahl der Workshops und Konferenzbeiträge steigt, jedoch gibt es äußerst selten spezifische Studiengänge oder Professuren zum Thema. Zum einen mag die Vorherrschaft der Disziplin der Medienwissenschaft als eigene Forschungsrichtung und Studiengang für die Analyse medialer Artefakte (oft auch institutionell in einem anderen Fachbereich oder Institut angesiedelt) gegenüber anderen geisteswissenschaftlichen Disziplinen ein Grund dafür sein, warum Entwicklungen der multimodalen Analyse im deutschsprachigen Raum nur langsam Akzeptanz finden. Ein weiterer Grund ist sicherlich der seit langem und in großem Maße immer noch anhaltende Imperialismus, der den linguistisch-semiotischen Theorien nach Saussure und Barthes zugeschrieben wird und in dem Sprache immer noch als Grundlage jeder Bedeutungskonstruktion angesehen wird. Der Erfolg dieser Theorien im Hinblick auf ihre Anwendbarkeit auf (audio-)visuelle Medien bleibt allerdings aus, was vor allem der theoretischen Abstraktheit und fehlenden Möglichkeiten der praktischen und empirischen Überprüfbarkeit geschuldet ist.<sup>1</sup>

Neuere Ansätze lassen sich nun auch im Rahmen der noch jungen Disziplin der Medienlinguistik (vgl. PERRIN 2011; LUGINBÜHL 2014; ANDROUTSOPOULOS 2014) finden, in der vermehrt eine Ausweitung verbaler Analysen unter Einbezug aller beteiligten semiotischen Einheiten stattfindet. Allerdings spielt Sprache in diesen Analysen immer noch die zentrale Rolle und Artefakte, die gänzlich oder fast ohne verbalen Anteil auskommen, finden nur selten die volle Aufmerksamkeit. Auch der Begriff der Multimodalität wird vornehmlich auf das Zeichensystem Sprache angewandt und folglich als eine allgemeine Theorie sprachlicher Multimodalität verstanden. Ellen Fricke beispielsweise als Vertreterin eines innerhalb Deutschlands viel zitierten Multimodalitätsansatzes trägt in ihrer Beschreibung einer generellen Multimodalität von Sprache den potenziell unterschiedlichen Manifestationen in Schrift und gesprochener Sprache Rechnung. Sprachliche Multimodalität gründet sich ihrer Meinung nach auf einer »[...] strukturellen und funktionalen Integration in ein und denselben Matrixcode oder alternativ [auf] der Manifestation eines Kodes in zwei unterschiedlichen Medien« (FRICKE 2012: 46). Fricke's Grundgedanke ist damit ein sprachzentrierter, der menschliche Kommunikation nicht monomodal, also lediglich auf gesprochene oder geschriebene Sprache beschränkt, sondern auch (und insbesondere) Gestik und Mimik in die Interpretation mit einbezieht, allerdings stets Sprache als zugrunde liegende semiotische Entität annimmt. Für das Ziel, Gesten als Teil sprachlicher Kommunikation zu analysieren, ist dies eine gerechtfertigte Festlegung. Ähnliche, vornehmlich germanistische Arbeiten mit Fokus auf Sprache in Verbindung mit anderen semiotischen Einheiten liegen von Cornelia Müller und Kollegen vor (vgl. MÜLLER 2009; MÜLLER/CIENKI 2009; MÜLLER et al. 2013, 2014).

---

<sup>1</sup> Das viel diskutierte Film-als-Sprache-Paradigma innerhalb der Filmsemiotik ist hier nur ein Beispiel von vielen, das die größtenteils gescheiterten Bemühungen um eine Übertragung grammatischer Mechanismen auf nicht-sprachliche Texte aufzeigt (vgl. NICHOLS 1975; STAM 1989; THOMPSON/BOWEN 2009).

Ein Nachteil dieser Ansätze ist, dass andersgeartete Medien, wie zum Beispiel Film oder auch Comics, nicht mit den von Fricke postulierten Eigenschaften von multimodaler Sprache gleichgesetzt werden können, da sie keine übergeordnete Modalität darstellen, die sich in unterschiedlichen Realisierungen zeigt. Stattdessen verbinden sie mehr als einen Kode sowie mehr als eine Sinnesmodalität und sind im Gegensatz zur Sprache deswegen nicht multimedial, sondern ebenso multimodal (vgl. BATEMAN/SCHMIDT 2011; WILDFEUER 2014a; BATEMAN/WILDFEUER 2014a; BATEMAN et al. 2017a; vgl. auch die Ausführungen zu *mode* und *Medium* weiter unten). Ihre multimodale Eigenschaft haben diese Artefakte aufgrund der Integration unterschiedlichster semiotischer Ressourcen, die gemeinsam seine Bedeutung konstruieren und deswegen auch gemeinsam, in ihrem intersemiotischen Zusammenspiel, betrachtet werden müssen. Solche Definitionen von Filmen oder Comics sowie ebenso Bilderbüchern, Photographien, Online-Diskursen u.v.m. als multimodale Texte folgen der Annahme, dass nicht mehr allein Sprache die Bedeutung eines kommunikativen Artefaktes bedingt, sondern dass diese immer Teil eines größeren, multimodalen Ensembles ist, das Bedeutung trägt.

Diese Annahme ist Grundvoraussetzung für das im internationalen Kontext fest etablierte und von Kress und van Leeuwen (2006, 2001) begründete Konzept von Multimodalität, das davon ausgeht, dass »language is no longer the central semiotic mode« (KRESS/VAN LEEUWEN 2001: 2). Auch andere Modalitäten in einem Text können demnach die Funktionen von Sprache übernehmen, ohne selbst Sprache zu sein. Dieses Konzept von Multimodalität sieht grundsätzlich alle Formen von Text – und nicht Sprache an sich – als aus mehreren semiotischen Ressourcen bestehende Artefakte und stellt deswegen eine globalere Perspektive gegenüber der von Fricke dar. Sie nimmt konkrete kommunikative Textformen in den Blick, die Bilder, Gestik, Mimik, Musik, Kameraspiel, etc. einsetzen, und ermöglicht so die Analyse von einer potenziell unendlichen Anzahl von Artefakten, die diese verschiedenen Modalitäten vereinen. Multimodalität wird dabei kategorial verwendet, um als konstitutive Eigenschaft von Texten in den Vordergrund zu rücken. Bucher sieht dies

als einen Wechsel in der Betrachtungsweise auf alle Formen der Kommunikation. Das bedeutet, dass jede Kommunikationsanalyse multimodal ausgerichtet sein muss und zeigen sollte, wie sich Sinn und Bedeutung eines Kommunikationsbeitrags aus den unterschiedlichen Modi ergeben (BUCHER/SCHUMACHER 2011: 11).

Diesen Wechsel der Betrachtungsweise verfolgen – so kann zumindest jüngst festgestellt werden – nun auch immer mehr Ansätze innerhalb der deutschen Wissenschaftslandschaft. Noch zögerlich, aber mit steigender Relevanz geht es nicht mehr darum, einzelne Modalitäten mit Sprache zu vergleichen und aufgrund des Sinneskanals, über den sie wahrgenommen werden, zu beschreiben. Stattdessen stehen Fragen nach dem Zusammenhang und der narrativen Struktur sowie nach dem Beitrag einzelner Modalitäten für die Kohärenz des Textes oder Diskurses im Vordergrund. Auch bereits etablierte diskurstheoretische, rhetorische oder stilistische Diskussionen finden immer öfter

Erwähnung und Anwendung (vgl. z. B. DIEKMANNSHENKE 2011; FRAAS et al. 2013; MAIER 2014).

Dabei stehen neben den rein sprachlich basierten und traditionell linguistischen Definitionen von Multimodalität, wie sie Fricke und Müller verfolgen, auch die allgemein kommunikationstheoretische Sicht in der Nachfolge Kress' und van Leeuwens (2001) sowie eine plurikodale (MITCHELL 1995; POSNER 1986) oder handlungstheoretische Auffassung (BUCHER 2011; SACHS-HOMBACH 2011) nebeneinander. Nichtsdestotrotz verfolgen viele dieser Ansätze vor allem einzelne Fallstudien innerhalb eines spezifischen Mediums sowie individuelle Textanalysen, ohne dass der jeweilige methodologische Ansatz eine präzise Ausdifferenzierung erfährt oder gar eine empirische Evaluation gewährleistet ist. Eine solche *ad hoc*-Qualität der Analysen wird nicht nur für sprachliche (vgl. HALLIDAY 1994: xvi), sondern in besonderem Maße auch für multimodale Ansätze (vgl. FORCEVILLE 2007) kritisch diskutiert und Forderungen nach einer lückenlosen empirischen Aufarbeitung werden lauter (vgl. BUCHER 2011). Ziel der folgenden Ausführungen wird es deswegen sein, die linguistischen Perspektiven auf Multimodalität über die Analysen einzelner Texte hinaus zu bündeln und für die notwendigen empirischen Diskussionen neu zu formulieren, um eine theoretisch wertvolle Basis für linguistische Multimodalitätsforschung allgemein bereitzustellen.

## 2. *modes* und Modalitäten

Obwohl ein, wenn nicht sogar der zentrale Begriff des Multimodalitätsparadigmas, wird *mode* zur Bezeichnung einer semiotischen Entität durchaus nicht einheitlich verstanden und definiert, ja sogar oft unterschiedlichst benannt, wie Klug/Stöckl (2015: 244-245) aufzeigen: z. B. als *mode of communication*, als *textual* oder *semiotic resource* oder auch als *representational mode*. Zudem steht der Begriff immer auch im Spannungsfeld zu Termini wie *Zeichen*, *Kode*, *Sinneswahrnehmung* oder auch *Medium* und vereint somit »mediale, kodebezogene und sozial-kulturelle Faktoren« (KLUG/STÖCKL 2015: 245) gleichermaßen.

Kritik an der Herangehensweise der meisten Definitionsansätze äußert sich deswegen vor allem im Hinblick auf fehlende eindeutige Unterscheidungen und Abgrenzungen sowie eine einheitliche und strenge Definition, die auch Verbindung mit und Anwendung in der tatsächlichen praktischen Analyse findet (vgl. FORCEVILLE 2010; KLUG/STÖCKL 2015). Während für sprachliche Analysen ausgefeilte Modelle zur Charakterisierung der jeweiligen Phänomene auf und für alle Ebenen sprachlicher Beschreibung vorliegen (vgl. z. B. für sprachliche Texte und Diskurse SPITZMÜLLER/WARNKE 2011), driften Ansätze außerhalb des Sprachsystems schnell ins Ungenaue ab und werden lückenhaft. Charakterisierungen anderer Zeichenressourcen gleichen dann einer disziplinär jeweils unterschiedlichen Checkliste, die abgearbeitet und dort markiert wird, wo

einzelne Punkte Anwendung finden, ohne dass allerdings eine vollständige Auflistung aller notwendigen und hinreichenden Kriterien aufgestellt wird.

Ziel der folgenden Ausführungen ist es deswegen, eine umfangreiche und für unterschiedlichste Arten anwendbare Definition des *semiotic mode* zu liefern, dessen Terminus wir dem allgemeinen Gebrauch folgend (vgl. auch KLUG/STÖCKL 2015) im Deutschen mit *Zeichenmodalität* übersetzen. Gleichzeitig werden wir weiterhin den im anglophonen Kontext bereits vielfach bestimmten Begriff *mode* verwenden.<sup>2</sup> Für eine genaue Bestimmung des Terminus und seiner Anwendungsbereiche werden wir zunächst eine ontologische Analyse der konstituierenden Faktoren der Zeichenmodalität vornehmen. Ontologisch ist diese Analyse im Sinne der notwendigen und hinreichenden Kriterien, die für eine semiotische Ressource vorliegen müssen, um als Zeichenmodalität im Sinne eines Interpretationsprozesses verstanden werden zu können. Weil diese Begriffsbestimmung ein semiotisches Unterfangen ist, werden wir immer auch Parallelen bzw. Unterschiede zu traditionellen semiotischen Herangehensweisen aufzeigen und schließlich Vorschläge für darauf aufbauende methodologische Vorgehensweisen machen. In einem weiteren Schritt werden wir dann unterschiedliche semiotische Ebenen der Beschreibung und Charakterisierung einer Zeichenmodalität vorstellen, die wir schließlich in Form eines Stratifizierungsmodells (vgl. Kap. 2.3) darstellen.

## 2.1 Ontologische Grundlagen für Zeichenmodalitäten

Für jede grundlegende Diskussion semiotischer *modes* oder Modalitäten ist es hilfreich und wichtig, den Status eines *modes* als Ergebnis einer interpretativen Praxis hervorzuheben, die durch eine Gemeinschaft von Nutzern konstruiert und aufrechterhalten wird (vgl. GOODMAN 1969; BATEMAN 2016). Die folgende Beobachtung von Kress und Kollegen hebt die daraus entstehenden Konsequenzen besonders detailliert hervor:

[...] the question of whether X is a mode or not is a question specific to a particular community. As laypersons we may regard visual image to be a mode, while a professional photographer will say that photography has rules and practices, elements and materiality quite different from that of painting and that the two are distinct modes (KRESS et al. 2000: 43).

Auch wenn sich die in diesen Interpretationsprozess involvierte Nutzergemeinschaft wie im Falle von Sprache als durchaus groß erweisen kann, muss die

---

<sup>2</sup> Neben *Modalität* findet sich im Deutschen in einigen Arbeiten auch die Verwendung *Modus*. Beide Begriffe sind linguistisch bereits vorgeprägt und bezeichnen im traditionellen Sinn eine semantische Eigenschaft bzw. eine grammatische Kategorie: Die Modalität eines Satzes beinhaltet die semantische Information, welche Einstellung des Produzenten zum Inhalt des Satzes vorliegt. Modus dagegen ist eine Konjugationskategorie des Verbs, die genau diese Eigenschaft z. B. im Indikativ oder Konjunktiv ausdrücken kann.

Für das Multimodalitätsparadigma wählen wir neben dem englischen Begriff *mode* den deutschen Begriff der *Zeichenmodalität*, um kategorial von einer Einheit sprechen zu können, die wiederum semantische und andere Eigenschaften tragen kann. Eine kompakte und hilfreiche Übersicht zu den unterschiedlichen Begrifflichkeiten geben Klug/Stöckl (2015).

Frage danach und nach ihren Verwendungsweisen zuallererst empirisch gestellt werden: Liegt ein systematischer Gebrauch bestimmter expressiver Ressourcen<sup>3</sup> innerhalb eines spezifischen Kontextes vor und können wir diese Ressourcen identifizieren und charakterisieren? Erst in der Folge führt diese erste empirische Überprüfung zu grundlegenden Fragen wie z. B. nach den systematischen sozio-kulturellen Praktiken innerhalb dieser Gesellschaft.

Ein weiterer wichtiger Ausgangspunkt für die vorzunehmende definitivische Präzisierung ist der Rückgriff auf die Eigenschaft der *Materialität*, die im Zuge der Bedeutungskonstruktion durch die Nutzer verwendet wird. Das für diese Konstruktion nutzbare Material ist vielfältig, muss jedoch zumindest soweit hinreichend kontrollierbar sein, dass stets eine zielgerichtete, zweckmäßige (und nicht willkürliche) Artikulation daraus hervorgebracht wird. Nur so kann jedes Material auch als Träger semiotisch aufgeladener Differenzierungen fungieren. Diese Art von Kontrolle kann nicht nur durch physikalische Prozesse der Nutzer (wie das Produzieren von Tönen durch Mund- und Lippenbewegungen oder die Ausführung bestimmter Gesten oder Körperbewegungen) erfolgen, sondern auch mithilfe anderer physikalisch-technischer Gegebenheiten innerhalb der Nutzergemeinschaft vollzogen werden (z. B. durch den Gebrauch eines Whiteboards, einer bestimmten Buchdrucktechnik oder eines Bildschirms oder Projektors zur Darstellung von PowerPoint- oder Keynote-Präsentationen).

Das für den Prozess der Bedeutungskonstruktion nutzbare Material dient faktisch und im Sinne seiner physikalischen Verfügbarkeit und perceptiblen Wahrnehmbarkeit als Vorlage für die zu erfolgenden semiotischen Entscheidungen und wird deswegen von uns auch als *Canvas* (Leinwand) bezeichnet (vgl. BATEMAN et al. 2017a: Kap. 3.2). Aus der physikalischen Materialität und den korrespondierenden technologischen Möglichkeiten der Artikulation entsteht ein virtueller Canvas, in dem das Material zur Nutzung zur Verfügung steht. Ein Vergleich zu Hjelmslevs (1961) Konzept der geformten Materie liegt hier nahe, wenn auch die Hjelmslevsche duale Unterscheidung von Ausdrucks- und Inhaltsebene viel stärker an semiotische Entscheidungen gebunden ist als der von uns intendierte Gebrauch eines virtuellen Canvas. Bereits vor der semiotischen Anwendung eröffnet dieser eine Menge von Anregungen und Angeboten (*affordances*), die in einem weiteren Schritt dann für die Bedeutungskonstruktion genutzt werden können.

Stöckl (2015) betont besonders im Hinblick auf nicht-verbale Zeichenmodalitäten, dass der Fokus auf dem Material einige der von Peirce beschriebenen, aber weniger oft verwendeten semiotischen Kategorien hervorbringt, die für ein Verständnis darüber, wie Zeichenmodalitäten miteinander operieren, hilfreich sein können: es handelt sich dabei insbesondere um das Qualizeichen (mit perceptuellen Eigenschaften) und das Sinzeichen (als Instanz, in dem perceptuelle Eigenschaften manifestiert werden können). Beide

---

<sup>3</sup> Der Begriff der (semiotischen) Ressource wird hier als vortheoretischer Terminus verwendet, dessen genauere Bestimmung erst noch zu erfolgen hat (siehe unten). Auch dieser Begriff wird im Multimodalitätsparadigma vielfach eingesetzt und unterschiedlich verstanden.

Zeichentypen sind Peirce zufolge immer ein wichtiger Teil der Produktions- und Interpretationsprozesse von bedeutungskonstruierenden Praktiken, wenn sie auch – in ihrer Form als Material – nicht einzelnen Sinneskanälen zugewiesen werden können (vgl. PEIRCE 1931-1958, §§2.275-2.308; NÖTH 2016). Eine solche Zuweisung findet in vielen Arbeiten zum Thema der Multimodalität tatsächlich häufig statt und führt in der Folge zum Beispiel häufig dazu, dass eine klare Trennung zwischen visuellen und auditiven Einheiten vorgenommen wird, ohne dass dafür eine nachvollziehbare Argumentation vorliegt.

Für unsere Zwecke soll diese Trennung nicht gelten: wir weisen individuellen Sinneskanälen oder auch biophysikalischen Unterscheidungen dieser Kanäle keine definitorische Rolle für die ontologische Bestimmung einer Zeichenmodalität zu. Besonders aktuelle Arbeiten im Bereich der Perzeptionsforschung und Neurokognition unterstützen unsere Annahme, indem sie herausarbeiten, dass bereits in einem sehr frühen Prozessstadium starke Kopplungen und Interaktionen zwischen einzelnen Sinneskanälen zu beobachten sind (vgl. CLARK 2011; SEELEY 2012; KLUSS et al. 2012). Es ist demnach sowohl aus einer theoretischen als auch praktischen Perspektive höchst problematisch, Grenzen zwischen den unterschiedlichen Sinnen oder gar eine sensorische Abschottung aufrecht zu erhalten und sie für eine Bestimmung von Zeichenmodalitäten hinzuzuziehen. Die Annahme der Materialität als grundlegende Stufe zur Interpretation einer Zeichenmodalität erlaubt Kombinationen von Sinneskanälen und -erfahrungen und somit auch die Interpretation multisensorischer Modalitäten. Auch hier ist es vor allem eine empirische Untersuchungsfrage zu prüfen, welche Dimensionen von Materialität für die jeweiligen Zeichenmodalitäten eine Rolle spielen.

## 2.2. Die Formung materieller Artikulationen

Auch wenn häufig Verbindungen zwischen sensorischen und semiotischen Modalitäten gezogen werden, ist es doch ebenso unumstritten, dass mehr als nur die Identifikation eines Sinneskanals zur Bestimmung der Zeichenmodalität benötigt wird. Es reicht allerdings auch nicht aus, lediglich Artikulationen eines spezifischen Materials anzunehmen; stattdessen müssen diese Artikulationen (in Form von Qualizeichen) als Instanzen wiederkehrender Muster (Sinzeichen) innerhalb einer Nutzergemeinschaft (als Legizeichen) erkennbar sein. Eine Menge von unterscheidbaren Kennzeichen mit bestimmten kontextsensitiven Bedeutungen (z.B. Ampelanlagen oder auf den Boden gemalte Straßenkreide als Wegweiser) kann dann ein *Zeichenrepertoire* ausmachen.

Anstelle der häufig zu findenden Metaphern des ›Codebuchs‹ oder ›Zeichenkatalogs‹, das individuellen Zeichen eine gewisse Prominenz zuweist, ziehen wir hier den Begriff des *Zeichenvehikels* in Betracht, der insbesondere physikalisch greifbare Spuren bezeichnet, die lediglich mithilfe einer Menge von minimalen Differenzierungen charakterisiert werden. Diese Differenzierungen im Material müssen erkennbar mit unterschiedlichen semiotischen Ereignissen

korrelieren, die die Nutzergemeinschaft zu kennzeichnen sucht. Wie wir später darstellen werden, kann diese Menge von Differenzierungen als eine weitere semiotische Ebene (Stratum) gesehen werden, die den Bereich der Materialität beinhaltet bzw. umfasst (vgl. Kapitel 2.3).

Diese ›negative‹ Definition der Bedeutungszeichen geht eher auf Saussure (1959) als auf Peirce zurück und wurde durch Hjelmslev (1961) in seiner ›algebraischen Semiotik‹ weiterentwickelt. Spuren im Material entsprechen in diesem Sinne nicht direkt den Referenten eines Zeichens, sondern es sind vielmehr erkennbare Differenzierungen zwischen diesen Spuren, die dann auch eine Unterscheidung der semiotischen Kategorien ermöglichen. In der Folge können wir jeden nicht-materiellen semiotischen Beitrag auf Grundlage der syntagmatischen und paradigmatischen Organisationsachsen charakterisieren, d.h. mit der paradigmatischen Ebene der Auswahl und Selektion sowie der syntagmatischen Ebene des Ausdrucks und der Rekodierung paradigmatischer Elemente in bestimmten strukturellen Konfigurationen näher bezeichnen. Eben jene strukturelle Komplexität ist damit immer auch ein intrinsischer Bestandteil unseres Modells und zwingend eine Möglichkeit der näheren Bestimmung der Zeichenmodalität.

Halliday (1978: 128-129) zufolge können paradigmatische Differenzierungen außerdem in Hierarchien von mehr oder weniger spezifischen, aber trotzdem immer abstrakten semiotischen Entscheidungen (*semiotic choices*) organisiert werden. Der Zweck einer strukturellen Konfiguration ist dann, Spuren in den Differenzierungen der materiellen Form zu hinterlassen, während die paradigmatische Beschreibung eine Organisationsstruktur für den Raum der semiotischen Entscheidungen, der mit jeder semiotischen Ressource verfügbar ist, anbietet. Wenn wir dies zurückprojizieren auf die Peircesche Kategorisierung der Zeichentypen, bedeutet dies, dass wir auch die paradigmatische Organisationsebene als eine Charakterisierung der Organisation von Legezichen ansehen können (vgl. BATEMAN 2014a: 261-263; BATEMAN 2018).

Eine weitere Konsequenz aus den oben angeführten Überlegungen von Kress und Kollegen ist, dass nicht alle von den RezipientInnen genutzten Zeichenmodalitäten gleichermaßen in ihren paradigmatischen und syntagmatischen Ebenen artikuliert werden. Es ist dann hilfreich, eine Topologie dieser Organisationsebenen anzuwenden, die von Kress/van Leeuwen (2001: 113) mit einer Unterscheidung zwischen lexikalisch-organisierten und grammatisch-organisierten semiotischen Ressourcen aufgestellt wurde. Erstere bestehen aus einer Menge von Zeichen mit wenig zusätzlicher Organisation – die voneinander unterscheidbaren Zeichen können demzufolge einfach aufgelistet werden. Letztere allerdings stellen ihre unterscheidbaren Zeichen innerhalb eines produktiven Systems von Bedeutungspotenzial bereit, das wiederum die Möglichkeit unterbreitet, einfachere Zeichen zu komplexeren zusammensetzen und dabei strukturelle Mechanismen ähnlich denen einer Grammatik zu

verwenden.<sup>4</sup> Lexikalisch-organisierte semiotische Ressourcen verfügen somit lediglich über eine oberflächliche paradigmatische Organisation, während grammatisch-organisierte Ressourcen eine erhebliche Tiefe in der paradigmatischen Struktur nachweisen können, die durch entsprechend komplexe syntagmatische Strukturen unterstützt werden und folglich auch komplexe semiotische Entscheidungen sichtbar machen.

### 2.3. Der Gebrauch materieller Artikulationen im Diskurs

Der letzte grundlegende Bestandteil für unsere Definition der Zeichenmodalität wird durch eine weitere semiotische Ebene in unserem Modell zur Verfügung gestellt: die Ebene der Diskursemantik. Aufgabe dieser Ebene für und innerhalb einer Zeichenmodalität ist es, die unterschiedlichen Implementierungen und Bereitstellungen des semiotisch aufgeladenen Materials mit seinem Gebrauchskontext und den jeweiligen kommunikativen Absichten in Beziehung zu setzen. Das bedeutet: die Diskursemantik einer Zeichenmodalität stellt die interpretativen Mechanismen bereit, die erstens eine direkte Verbindung der Elemente der Zeichenmodalität mit dem jeweiligen Kontext vornehmen sowie zweitens auch die *beabsichtigte* Menge an Interpretationen eingrenzen. Diese Interpretationen können in ihrer Begrenztheit dann stark variieren, je nachdem wie spezifisch oder abstrakt ihre Anleitungen zur Interpretation ausfallen.

Viele traditionelle Modelle oder Zugänge zu Multimodalität gehen von einer viel direkteren Beziehung zwischen Zeichen (die aus einem bestimmten Material geformt wurden) und den Bedeutungen für diese Zeichen aus. Dies mag auch für sehr einfache Zeichenmodalitäten an der unteren Grenze dessen, was wir hier als *semiotic mode* beschreiben, der Fall sein, wenn es sich um eher triviale interpretative Notwendigkeiten handelt, die die Bedeutungen dieser unterscheidbaren Zeichen unabhängig von ihrem Gebrauchskontext bestimmen. Ein rotes Ampellicht zum Beispiel wird innerhalb einer bestimmten Kultur immer als Zeichen für »Stopp« gesehen werden und es bedarf hier keiner weiteren Rückgriffe auf die Entstehungsgeschichte roter und grüner Zeichen im konkreten Kontext der aktuellen Straßenverkehrslage, die einst für diese Bedeutungszuweisung verantwortlich war. Auch wenn es durchaus möglich ist, dass eine andere Nutzergemeinschaft in einem anderen Kontext diesem roten Ampelzeichen einmal eine andere Bedeutung zuweist, sind diese Alternativen im Allgemeinen kontextunabhängig und somit unabhängig von jedem anderen Gebrauch roter und grüner Zeichen. Es handelt sich hierbei um einen kode-basierten Blick auf Zeichen, der in unserem Modell dem entspricht, was wir als einfachste und begrenzte Form der Zeichenmodalität sehen und in

---

<sup>4</sup> Von einer Metaperspektive auf unsere Diskussion können wir unsere Ausführungen in diesem Kapitel folglich durchaus als Versuch beschreiben, von einer lexikalischen Definition der Zeichenmodalität abzurücken und uns stattdessen einer »grammatischen« zuzuwenden, die analytische Schlussfolgerungen über diese Zeichenmodalitäten und ihren Gebrauch zulassen.

der die materiell festgelegten Differenzierungen ein ›Lexikon‹ generieren, das relativ feste Bedeutungszuweisungen vornimmt.

Dieser kode-basierte Blick gilt in vielen Ansätzen immer noch als ein generelles Modell für Zeichensysteme, obwohl die Trennung von Zeichen von ihrem Gebrauch und Kontext stets Anlass für Probleme und Missverständnisse ist. Ein besonderes großes Problem innerhalb dieser Diskussionen ist zum Beispiel die Unterscheidung von Kode/Kodierungen und Inferenzen (vgl. u.a. TSENG/BATEMAN 2012; BATEMAN/WILDFEUER 2014a). Diejenigen Ansätze, die (korrekterweise) Inferenzmechanismen als relevant für die Interpretation von Bedeutung sehen, schließen leider (inkorrektweise) den Begriff des Kodes aus ihren Überlegungen aus. Dies wiederum führt oft zu starken Bedenken an den Anwendungsmöglichkeiten dieser semiotischen Modelle auf eine Vielzahl nicht-verbaler Medien. Unsere hier vorgestellte Definition dagegen betont die indirekte Beziehung zwischen materiellen Spuren und Bedeutungszuweisungen, die immer auch das Konzept der Inferenzziehung und vielmehr noch unterschiedliche Typen von Inferenzen auf unterschiedlichen Ebenen semiotischer Abstraktion in die Betrachtung miteinbezieht. Fälle, in denen materiell manifestierte Differenzierungen ein Lexikon mit relativ festen Bedeutungen konstruieren und kompositorisch operieren (als Kode ohne Inferenzziehung), spielen dagegen nur eine untergeordnete Rolle, weil sie keine ausreichende Grundlage für komplexe Prozesse der Bedeutungskonstruktion bieten.

Das Vorhandensein einer Diskurssemantik ist demzufolge nicht nur ein besonderes Kennzeichen, sondern auch das Gütesiegel unserer Versuche der ordnungsgemäßen Bestimmung einer Zeichenmodalität. Ohne diese Ebene der Diskurssemantik kann eine Zeichenmodalität lediglich in sehr spezifischen Kontexten und nur sehr eingeschränkt wirken. In diesem Fall können wir von so genannten ›semiotischen Proto-Modalitäten‹ sprechen – eine Beschreibung, die den Ausführungen Hallidays (1978: 121) über frühkindliche Sprechphasen und ihrer Bezeichnung als ›protolinguistisch‹ folgt. Demgegenüber erlaubt die zusätzliche Stratifizierung, wie wir sie für Zeichenmodalitäten vornehmen und die durch die zusätzliche Ebene der Diskurssemantik und den damit vermittelten Richtlinien zur Interpretation ermöglicht wird, eine Generalisierung semiotischer Konfigurationen über unterschiedliche Kontexte hinweg.

In früheren Ansätzen wird die Ebene der Diskurssemantik oft sehr allgemein einer Ebene der Pragmatik oder Kommunikation im Kontext zugeschrieben. Dies führt in der Folge dazu, dass den jeweils verfügbaren Diskurssemantiken nicht jene genaue Prüfung widerfährt, die für individuelle Zeichenmodalitäten und ihre Kombinationen aber gerade notwendig sind. Auch Kontextabhängigkeit und feine Bedeutungsunterschiede bleiben offene Probleme, für die innerhalb pragmatischer Herangehensweisen keine adäquate Lösung zu finden ist. Dies wiederum entspricht den bekannten Schwierigkeiten innerhalb der Text- und Diskursanalyse, die jeweiligen Mechanismen der Diskurskonstruktion in Form einer Diskurssemantik zu rekonstruieren und eindeutige Interpretationswege nachzuweisen, die sonst einem allgemeinen Pragmatik-Ansatz überlassen und nicht aufzuspüren wären (vgl. KAMP 1981; WILDFEUER

2014b). Das Konzept der Diskurssemantik macht es wiederum möglich, bestimmte Klassen von Inferenzen zu isolieren, die insbesondere als Vermittler zwischen kompositionell konstruierten semantischen Informationen und abstrakterem, kontextuellem Wissen arbeiten (vgl. ASHER/LASCARIDES 2003; WILDFEUER 2014b).

Diskurs, wie wir ihn hier verstehen, arbeitet demzufolge auf einer lokalen, individuellen und vor allem textorientierten Ebene (vgl. MARTIN 1992) und steht komplementär zum breiter angelegten Diskurskonzept nach Foucault, das größere Kontexte und kulturelle Strukturen beinhaltet.<sup>5</sup> Diskurssemantik als eine Ebene der Beschreibung von Zeichenmodalitäten ermöglicht es, die materielle Ebene des Textes als Einzelobjekt in dem entsprechenden Kontext und mit Rückgriff auf den jeweiligen Gebrauch zu beschreiben, ohne dafür eine Beschränkung auf eine textuelle oder diskursive Beschreibungsebene vornehmen zu müssen, wie dies oft in linguistischen (und insbesondere germanistischen) Ansätzen der Fall zu sein scheint. Insbesondere für die Betrachtung nicht-verbaler Texte bzw. Artefakte hat dies grundlegende Folgen, z.B. im Hinblick auf die lang anhaltende Diskussion um die Frage, ob Bilder als autonome, kommunikative Artefakte gelten und auch ohne den Zusatz sprachlicher Informationen interpretiert werden können (vgl. BARTHES 1964: 10-11). Um in dieser Weise zu funktionieren, müsste es sich Peirces traditionellem Ansatz zufolge bei Bildern um solche Zeichentypen handeln, die gänzlich ohne andere Informationen und auf sich selbst gestellt wirken, um zum Beispiel ikonischer (oder indexikalischer oder symbolischer) Art sind. Um allerdings kommunikativ wirksam zu sein, müssten sie in Propositionen oder Argumente eingebettet sein, die helfen, ihre kommunikative Rolle näher zu bestimmen.

Unserem Modell folgend ist die Frage nach der Autonomie dagegen falsch gestellt: Jede Zeichenmodalität kann autonom sein, ja das Kriterium der Autonomie ist in der Ebene der Diskurssemantik sogar intrinsisch verankert. Diese Diskurssemantik liefert zugleich das Maß an Informationen, das zur Interpretation dieser Autonomie notwendig ist. Jede Zeichenmodalität kann also per definitionem die gesamte Bandbreite an semiotischen Entscheidungen enthalten, wie sie beispielsweise für die Unterscheidung von ikonischen, indexikalischen und symbolischen Zeichen konzipiert wurden (vgl. NÖTH 2016). Die Entscheidung darüber, ob etwas als Index, Ikon oder Symbol interpretiert wird, ist dann gänzlich eine diskursive Entscheidung, die aufgrund der durch die jeweilige Diskurssemantik gelieferten Mechanismen ermöglicht wird. Aufgrund dieser Eigenschaften sehen wir unsere Ausführungen als gänzlich übereinkommend mit der Peirceschen Auffassung von aktiver *Semiose* als dem Prozess des Zeichenlesens und -interpretierens. Die explizite Bezugnahme auf die Ebene der Diskurssemantik in jeder Interpretation einer Zeichenmodalität fügt dieser relativ abstrakten Auffassung entsprechend formalisierte Prinzipien

---

<sup>5</sup> Für eine ausführlichere Auseinandersetzung mit den Begriffen Text und Diskurs sowie den englischen Äquivalenten sei auf die Ausführungen in Wildfeuer (2018) verwiesen.

hinzu, die in der Folge auch für eine empirische Überprüfung operationalisiert werden können.

Mit Blick auf die formalisierten Prinzipien bewegen wir uns im Kontext methodologischer Präzisierung, die wir in Kapitel 3 detailliert und mit Blick auf Operationen von *Textualität* vornehmen werden. Textualität liegt vor, wenn Artefakte und Performanzen mehr oder weniger explizite Hinweise (*cues*) für ihre eigene Interpretation aufweisen, die entweder auf einer sehr allgemeinen Ebene von Genre oder Textsorte oder aber spezifischer auf der Ebene der Übergänge zwischen Sätzen oder Einheiten des Textes wirken (vgl. KESSELHEIM 2011; BATEMAN et al. 2013; STÖCKL 2016). Diese Auffassung von Textualität und die mit ihr verfügbaren Analysemittel (siehe unten) stellen eine feinere und fokussierte methodologische Herangehensweise an die Analyse dieser Artefakte dar, die mit dem einfachen Blick auf Bilder oder Wörter nicht vorliegt.

### 2.3 Zur Definition der Zeichenmodalität

Unser grundlegender Ansatz zur Definition der Zeichenmodalität umfasst nun die wichtige diskursive Komponente, die es ermöglicht, die Modalität nicht getrennt von ihrer Materialität zu beschreiben. Zur Verbindung dieser und aller anderen Komponenten bedienen wir uns der im Bereich der systemisch-funktionalen Soziosemiotik entwickelten und wiederum den Hjelmslevschen Prinzipien folgenden Darstellungsweise als *stratifizierte Systeme*. Dazu muss zunächst jedes materielle Substrat als essentielle Komponente einer jeden Zeichenmodalität festgelegt und so als erstes Stratum der *semiotic mode* beschrieben werden. Das Material selbst kann dabei über unterschiedliche Sinneskanäle hinweg operieren. Auf einer zweiten, vermittelnden Ebene stellt ein weiteres Stratum mehr (grammatisch-) oder weniger (lexikalisch-organisierte) *kompositionell* arbeitende und funktional motivierte Möglichkeiten der formalen Differenzierung bereit. Auf einer dritten Ebene, die die beiden anderen Ebenen der semiotischen Abstraktion umgibt, befindet sich das abstraktere Stratum der (lokalen) Diskurssemantik, das abduktiv auf den Beschreibungen der unteren Ebenen operiert.<sup>6</sup>

Das stratifizierte Modell der Zeichenmodalität als Ganzes stellen wir graphisch in Abbildung 1 dar. Es zeigt – von unten beginnend – zunächst die Ebene der Materialität und als zweites die Ebene der Form, die mithilfe der paradigmatischen und syntagmatischen Achsen die technischen Eigenschaften der Modalität organisiert. Beide Ebenen umfassend folgt dann das Stratum der Diskurssemantik, das eine Kontextualisierung der beiden anderen Ebenen vornimmt und diese so im Kontext interpretierbar macht. Dieses Modell, das sowohl aus funktionalen als auch formalen Ansätzen der Diskursanalyse

---

<sup>6</sup> Im Gegensatz zur Verwendung des Begriffs ›Stratum‹ bei Kress/van Leeuwen (2001: 4) behalten wir hier die semiotische Bedeutung der Strata als dichte und formal miteinander verwobene Beschreibungen auf unterschiedlichen, hierarchisch geordneten Ebenen der Abstraktion bei.

entwickelt wurde, kann nun auf alle Zeichenmodalitäten übertragen werden, ganz gleich welche spezifische Materialität sie besitzen.

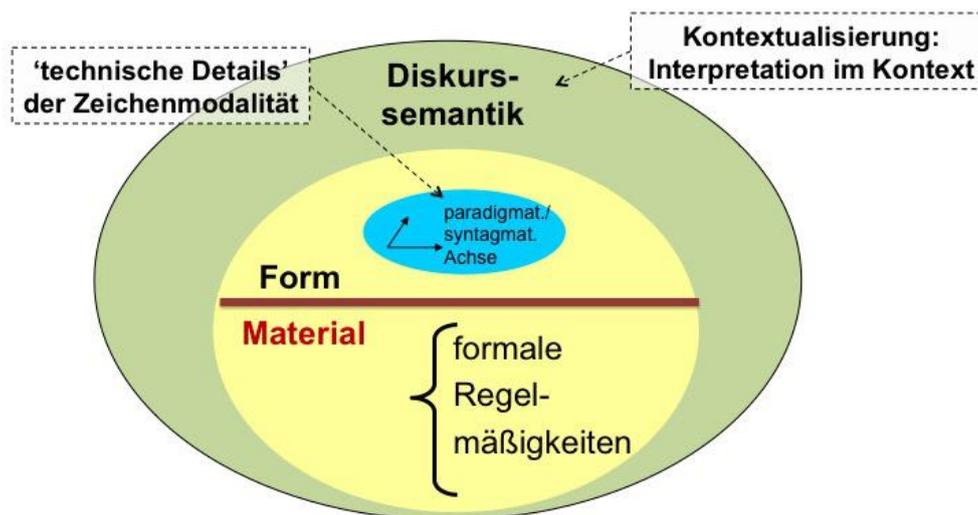


Abb. 1:  
Abstrakte Definition der Zeichenmodalität (vgl. BATEMAN 2016)

Aufgrund der drei semiotischen Strata umfassen alle semiotischen Modalitäten notwendigerweise perzeptuelle Eigenschaften, eine ›lexikogrammatische‹ Organisation sowie die entsprechenden Diskursmechanismen. Diese Differenzierungen sind nicht allein terminologischer Art; stattdessen können unterschiedliche Facetten von Multimodalität aufgrund ihrer unterschiedlichen Mechanismen viel besser auseinandergehalten werden. Beispielsweise sind die unteren beiden Strata semiotisch durch das Prinzip der *Realisierung* (oder auch des Ausdrucks) miteinander verbunden, was bedeutet, dass die jeweiligen Muster der Zeichenmodalität in den materiellen Eigenschaften realisiert werden und so formale Eigenschaften die jeweilige Modellierung stützen. Die Beziehung zwischen dem zweiten und dem dritten Stratum der Diskurssemantik besteht dagegen aufgrund von anfechtbaren Interpretationsregeln, deren Hypothesenbildung immer eine Überprüfung bedarf. Mithilfe einer umfassenden Beschreibung all dieser Relationen ist es möglich, die interagierenden Komponenten der Zeichenmodalität voneinander zu unterscheiden und adäquat darzustellen.

Ein weiterer Vorteil dieser Beschreibung ist, dass die tatsächlichen materiellen Einheiten, die für das Erkennen und Interpretieren einer Zeichenmodalität eine Rolle spielen, somit bereits ein Ergebnis diskursiver Interpretation

sein können und gerade nicht nur als Input für diese Interpretation gelten (vgl. BATEMAN/WILDFEUER 2014a: 190-192). Diese erst durch die Ebene der Diskurssemantik möglich gewordene Darstellung, die den aktiven Interpretationsprozess deutlich macht, verändert wesentlich unser Verhalten in der Analyse multimodaler Artefakte, indem sie erstmals ermöglicht, die einzelnen Beiträge der Zeichenmodalitäten in ihrer Kombination aufzuzeigen. Zwar ist die Analyse dieser Kombinationen und ihrer jeweiligen Bestandteile im intersemiotischen Zusammenspiel ein seit langem anvisiertes Ziel der Multimodalitätsforschung, scheiterte bisher jedoch immer an fehlender definitorischer Klarheit über die tatsächlichen Eigenschaften semiotischer Modalitäten.

In unseren Arbeiten zur Comicanalyse beispielsweise (vgl. BATEMAN/WILDFEUER 2014b: 376-378) zeigen wir auf, wie die präzise Beschreibung einer Zeichenmodalität in Verbindung mit ihrer diskurssemantischen Analyse feinste Unterschiede im konventionellen Gebrauch von Zeichen aufzeigen und die jeweils notwendigen Interpretationsschritte herausarbeiten kann. In einem Beispiel von Scott McCloud (1994: 128), in dem er einer gekrümmten Linie in einem Panel zuweist, aus einer Pfeife heraus aufsteigenden Rauch anzuzeigen, während die fast gleiche Linie in einem anderen Panel den Geruch von Müll darstellt, ist das konventionelle Dekodieren dieser Zeichenbedeutung allein nicht möglich. Stattdessen ist es notwendig, die entsprechenden Diskursinterpretationen im jeweiligen Kontext zu finden, die eine sinnvolle und damit auch maximal kohärente Verbindung der Zeichen mit den jeweiligen Bedeutungen erlauben (vgl. MÜLLER 2012). Dies geschieht aufgrund von Hypothesen, die die gekrümmten Linien durch ihre Platzierung im Panel und in Relation zu den anderen visuellen Informationen darin erzeugen. Nur durch die sinnvolle Anbindung der Linien an die anderen Informationen mithilfe einer abduktiv gefundenen Diskursrelation ist es möglich, die jeweilige Bedeutung in ihrem spezifischen Kontext zu bestimmen und somit den Unterschied zwischen aufsteigendem Rauch und dem Geruch von Müll zu erkennen. Eine wichtige Rolle innerhalb dieses Interpretationsprozesses spielt auch die Dynamizität der Bedeutungskonstruktion, die in der Analyse berücksichtigt werden muss. Indem sich der Kontext um die Linien herum im Laufe des Diskurses verändert, verändert sich auch die Bedeutungszuweisung und es ist nicht mehr möglich, die gleiche Hypothese zur Interpretation aufzustellen, die für das erste Panel plausibel war. Stattdessen muss eine neue Hypothese gefunden werden, die durch den weiteren, dynamischen Verlauf des Diskurses und dem damit einhergehenden Kontext bestätigt wird. Diese Differenzierung der diskurssemantischen Operationen kann nur mithilfe des vorgestellten Modells und eben jener Komponente als zentraler Bestandteil der Zeichenmodalität vorgenommen werden.

## 2.4 Zeichenmodalität und Medium

Die nun vorliegende präzise Definition einer Zeichenmodalität erlaubt uns in einem weiteren Schritt auch, eine Abgrenzung dieser Einheit gegenüber dem (teilweise ebenso schwer fassbaren und hier keineswegs in aller Länge diskutierbaren) Begriff des Mediums vorzunehmen.

Tatsächlich treten Zeichenmodalitäten nie alleine auf und nie außerhalb eines spezifischen Gebrauchskontexts; ebenso wenig treten sie ohne bestimmten Nutzen bzw. kommunikativen Zweck auf. Sie sind damit immer eingebunden in eine größere Kommunikationssituation, die bestimmte Bedingungen und Einschränkungen vornimmt, die auf das Artefakt und die damit einhergehenden Zeichenmodalitäten einwirken. Eine spezifische Form dieser Kommunikationssituation ist zum Beispiel das *Medium*, in dem das jeweilige Artefakt ausgedrückt wird. Die Frage nach der Medienspezifität nimmt damit eine wichtige Stellung im Rahmen multimodaler Analysen ein, wird doch eine Analyse der unterschiedlichen Modalitäten und ihrer Kombination in einem spezifischen Kontext, d.h. auch in der jeweiligen Kommunikationssituation, immer auch Eigenschaften und Merkmale für das jeweilige Medium hervorbringen.

Auf der anderen Seite nimmt auch das jeweilige Medium Einfluss auf den Gebrauch unterschiedlicher Modalitäten und ihre Kombinationsmöglichkeiten, indem es verschiedenste Materialien bereitstellt, die innerhalb einer Nutzergemeinschaft und in sozial und historisch geprägten Kontexten eingesetzt werden können. Wir sehen ein Medium deswegen als einen historisch bedingten Ort des Einsatzes und der Mobilisierung von Zeichenmodalitäten zur Erfüllung verschiedener kommunikativer Zwecke (vgl. BATEMAN 2017). Ein Buch als Medium zum Beispiel ermöglicht die Verwendung der Zeichenmodalitäten der geschriebenen Sprache, der Typographie oder des Seitenlayouts. Einige weitere Modalitäten sind ebenso prototypisch für das Medium Buch, andere sind es nicht und möglicherweise aufgrund ihrer Diskurssemantik auch gar nicht geeignet dafür.

Es gilt dafür erneut festzuhalten, dass Zeichenmodalitäten nicht an einzelne sensorische Kanäle gebunden sind und es folglich nicht zielfördernd ist, beispielsweise von *dem* visuellen oder auditiven *mode* zu sprechen, auch wenn dies oft getan wird (vgl. u.a. KRESS/VAN LEEUWEN 1998). Stattdessen kann allerdings durchaus sinnvoll von visuellen Medien gesprochen werden, die ihr visuell wahrnehmbares Material in entsprechender Formung und nach Bereitstellung der notwendigen Diskursmechanismen in Form von Zeichenmodalitäten äußern können. Wenn also die Rede von einem visuellen Medium wie dem Comic ist, bezieht sich diese Charakterisierung auf das dem Medium inhärente und zur Bedeutungskonstruktion genutzte Material, das visuell wahrnehmbar ist. Im Falle von Comics sind das neben dem Druck geschriebener Sprache auch alle bild- und zeichentechnischen Eigenschaften wie Farbe, Linienstärke oder auch Layout. Ob diese jeweils auch als Zeichenmodalität interpretiert werden, weil sie in den Prozess der Bedeutungskonstruktion involviert sind, gilt es erst mithilfe des skizzierten Modells und einer diskurssemantischen Analyse zu

klären. Neben diesen materiellen Eigenschaften des Mediums können auch seine Übertragungsart oder die Art und Weise der Darstellung ausschlaggebend für die Bezeichnung des Mediums sein.

Medien stellen daher grundsätzlich Möglichkeiten der Mobilisierung und Realisierung semiotischer Modalitäten dar und entsprechen somit dem von Hartmut Winkler (2008) beschriebenen Konzept des Biotops für Semiose:

Medien sind gesellschaftliche Maschinen, die ein Biotop für die Semiose, für die Artikulation und für die Herausbildung von Zeichen bereitstellen (WINKLER 2008: 213).

Diese Auffassung ähnelt stark den in vielen funktionallinguistischen Ansätzen verwendeten Formulierungen des *mode* oder *channel of communication* (vgl. z. B. MARTIN 1985: 17) und steht auch in engem Zusammenhang mit dem Vorschlag Marie-Laure Ryans, dass ein Medium in seine Übertragungskomponenten einerseits und seine semiotischen Komponenten andererseits aufgeteilt wird (vgl. RYAN 2004). Diese Umorientierung und Neufassung des Medienbegriffs im Hinblick auf seine unterschiedlichen Bestandteile, zu denen dann auch die jeweils verwendeten Zeichenmodalitäten gehören, ist auch innerhalb des Multimodalitätsparadigmas als eine wichtige Veränderung festgehalten worden:

[...] we have come from a period in which there had been a stable constellation of the mode of writing with the medium of the book. That had led to a kind of naturalization in which to talk about the mode seemed like talking about the medium (KRESS 2004: 446).

Es muss damit ein erklärtes Ziel multimodaler Analyse sein, gegen die von Kress angesprochene Naturalisierung zu wirken und die Unterschiede zwischen Medium und *mode* sowie ihre Wechselwirkungen aufzuarbeiten. Dies ist vor allem eine empirische Aufgabe, die es mithilfe des von uns skizzierten theoretischen Ansatzes zu klären gilt. Denn trotz der umfangreichen Einzelanalysen unterschiedlichster Medientexte können wir noch keine verlässlichen Aussagen darüber treffen, ob beispielsweise Filme oder Comics tatsächlich eigene *semiotic modes* darstellen, auch wenn bereits einige Gesetzmäßigkeiten und spezifische Funktionsweisen Argumente für den Status einer Zeichenmodalität liefern (vgl. auch die Beispielsanalyse in Kapitel 3). Es gilt hier weiterhin, nach der zugehörigen und geeigneten Diskurssemantik für die jeweiligen Artefakte zu suchen, die die beteiligten Zeichenmodalitäten charakterisieren und ihre jeweiligen Muster der Bedeutungskonstruktion erklären kann. Auf Grundlage der nun vorliegenden theoretischen Ausführungen werden wir dazu im Folgenden auch die notwendige methodologische Präzisierung vornehmen und eine linguistisch-empirische Analyse multimodaler Artefakte vorschlagen.

### 3. Multimodalität und Diskurs

Unsere Ausführungen zur Definition der Zeichenmodalität haben wiederholt die zentrale Rolle der Diskurssemantik für eine präzise Bestimmung und Analyse im multimodalen Kontext betont. Im Folgenden wollen wir aufzeigen, wie

diese Bestimmung in konkreten Gebrauchskontexten auch methodologisch und mit Blick auf empirisch evaluierbare Ergebnisse auszuführen ist.

Viele semiotisch-orientierte Multimodalitätsansätze werfen dafür ihren Blick auf die semiotischen Ressourcen, die die jeweiligen Modalitäten verwenden und anbieten. Kress/van Leeuwen (2006) bieten beispielsweise ein Analyseinstrumentarium an, das gerne und häufig von anderen Ansätzen übernommen wird, allerdings in seiner statischen Beschreibung und Klassifizierung der Ressourcen ebenso oft zu Problemen führt. Stöckls (2004) sehr detaillierte Auflistung unterschiedlicher Modalitäten und ihrer Wechselwirkungen liefert eine ähnlich präzise Darstellung, betont aber zugleich, dass die mit dieser Auflistung nicht beschreibbare Dynamizität der Bedeutungskonstruktion eine ebenso wichtige Rolle spielt. Wie wir in unserer Interpretation des McCloud-Beispiels gezeigt haben, kann eine solche dynamische Interpretation aufgrund eines sich ändernden Kontextes und variierenden Interpretationsbedingungen mithilfe einer Liste von vorher festgelegten Kriterien nicht überzeugend erklärt werden. Besonders die Kombination von unterschiedlichen Ressourcen in der tatsächlichen semiotischen Arbeit erfordert eine genauere Analyse des spezifischen Kontextes und der darin geltenden Bedingungen. Auch Bucher (2011) betont, dass eine kontextunabhängige Charakterisierung individueller Modalitäten gerade nicht hilfreich für eine Analyse ist, in der der spezifische multimodale Kontext die Interpretation bedingt und die Rezeption mithilfe von expliziten Hinweisen leitet.

Unser Ziel ist es demnach, linguistisch-semiotische Ansätze vorzustellen, mit denen eine dynamische Interpretation der Zusammenwirkung unterschiedlichster Zeichenmodalitäten innerhalb eines spezifischen Kontextes möglich ist. Dazu bedienen wir uns zeitgenössischer Arbeiten aus der Textlinguistik und Diskursanalyse, die mit Blick auf verbale Texte sowohl unsere Annahme einer Diskurssemantik als wichtiger Bestandteil der Zeichenmodalität teilen als auch die Dynamizität der Interpretation durch Veränderungen im Kontext und der Informationsvergabe hervorheben.

### **3.1 Linguistische Diskursanalyse**

Entwicklungen innerhalb der Linguistik, die bereits in den 1950er und 1960er Jahren ihren Ursprung haben (vgl. HARRIS 1952) und bis heute anhalten, haben unter anderem einen Fokus auf die Einheit des Diskurses als zentrale Größe semantischer Untersuchungen gelegt (vgl. KAMP 2007: 8-9). Dieser Fokus zeugt von einem Ansatz zur Analyse von Bedeutungskonstruktion, der über die Grenzen traditioneller Semantik deutlich hinausgeht und diskurspezifischen Eigenschaften in der Interpretation wichtige Funktionen zuweist.

Diskurs als Größe oberhalb der Satzebene wird als dynamische Konstruktion gesehen, deren oft (und in multimodalen Artefakten in besondere Maße) nicht explizit genannten Informationen ebenso dynamisch und auf Basis von Welt- und Kontextwissen inferiert werden. Bedeutung wird demnach

durch Hypothesenbildung im Hinblick auf die Informationen ermittelt, die im und durch den Kontext neu hinzugefügt werden und so die Interpretation verändern können. Hier spielen grammatische oder syntaktische Informationen weniger eine Rolle, stattdessen sollen die logischen Inferenzprozesse der RezipientInnen, die das Update-Potential einer neuen Information erkennen und mit den bisherigen Inhalten verbinden, möglichst genau und mithilfe formaler Methoden nachvollzogen werden. Dies ist für die Beschreibung nicht-verbaler Artefakte besonders hilfreich, da grammatische Strukturen in ihnen nur schwer nachgewiesen werden können, auch wenn dies ein viel adressiertes Anliegen war und ist.<sup>7</sup> Somit zielt die Frage danach, wie Bedeutung aus (audio-)visuellen Artefakten mit Rückgriff auf den inferentiellen und damit abduktiven Rezeptionsvorgang der RezipientInnen konstruiert werden kann, auf eben jene Ansätze innerhalb der Diskursforschung, die genau diese Schritte einer dynamischen und inferentiellen Interpretation beschreiben können.

Hier bieten sich vor allem Ansätze an, die sich einer logisch und philosophisch gestützten Herangehensweise bedienen und nach den Inferenzprozessen während der Interpretation fragen. Auf Grundlage der Tradition formaler Semantik zur Darstellung sprachlicher Mechanismen, die in der Montague-Grammatik ihren Ursprung findet (vgl. MONTAGUE 1974) und im Laufe der Zeit zahlreiche Weiterentwicklungen erlebte, sehen Vertreter der dynamischen Diskurssemantik eine wichtige Verbindung zwischen Logik als Wissenschaft von den Inferenzen und Semantik im Allgemeinen (vgl. KAMP/REYLE 1993: 13). Die logische Fundierung der Analyse von (sprachlicher) Bedeutung zielt auf eine präzise und verlässliche Informationsfindung ab, die sich dafür der formalen Darstellung mithilfe von logisch-mathematischen Methoden und Interpretationsmustern bedient. Eine solche Formalisierung sowohl der Einheiten des Diskurses als auch der Interpretationsprozesse setzt eine Konzentration auf das Material voraus, das die einzelnen bedeutungstragenden Elemente charakterisiert. Dabei ist nicht das bloße Vorhandensein von unterschiedlichen semiotischen Ressourcen entscheidend, sondern vor allem das Vorliegen von Kohärenz und Struktur als ein dem Material inhärentes Kriterium, das durch die Interpretation der Zeichenmodalitäten und ihrer Zusammenhänge im Kontext Bedeutung konstruiert. Hier sind es vor allem satz- oder einheitenübergreifende Diskursrelationen zwischen Propositionen des Diskurses, die seine ganzheitliche Interpretation ermöglichen. Der Wahrheitsgehalt der Bedeutung, der im Satz meist als unterspezifiziert gesehen wird, wird innerhalb dieser Beschreibungen aufgrund spezifischer Konditionen bestimmt und den Einheiten nicht einfach zugeschrieben, wie es in traditionellen semantischen Ansätzen oft der Fall ist. Stattdessen spielen Kontextbedingungen ebenso eine Rolle wie

---

<sup>7</sup> Zum Beispiel führt die Suche nach einer Filmsyntax bis heute zu Diskussionen im Rahmen der Filmsemiotik und -wissenschaft (vgl. MÖLLER-NASS 1986; MONACO 2000). Ähnliche Auseinandersetzungen finden sich im Bereich der linguistischen und psychologisch-kognitiven Comicforschung, in der die Arbeiten von Neil Cohn zur »grammar of visual language« (2013) vielfach rezipiert werden. Aber auch im Rahmen multimodaler Analysen ist der Grammatikbegriff (immer noch) beliebt (vgl. KRESS/VAN LEEUWEN 1996; FRICKE 2012), aber gleichermaßen umstritten (vgl. auch BATEMAN/WILDFEUER 2014a).

der individuelle Wissensstand der jeweiligen RezipientInnen. Mithilfe dieser dynamisch abfragbaren Informationen können dann verlässliche Angaben über die Interpretation gemacht werden. Plausibel werden diese Angaben, so die allgemein vertretene Annahme, aufgrund von Struktur und Kohärenz im Diskurs, die zugleich Bedingungen für die Interpretation festlegen und so den Inferenzprozess des Rezipienten steuern. Wie wir bereits betont haben, ist eine solche Steuerung der Interpretation für die Frage, wie Bedeutung durch und im Diskurs entsteht, von besonderer Relevanz, da nur durch textuelle Beweise in Form von Inferenzen bzw. der Diskursstruktur allgemeingültige Aussagen über die Interpretation getätigt werden können. Genau an dieser Stelle kommt die Ebene der Diskurssemantik als Bestandteil der Zeichenmodalitäten ins Spiel. Sie ermöglicht die Analyse des geformten Materials im Hinblick auf seine Funktion im Kontext und stellt als analysefähige Mechanismen zum Beispiel die unterschiedlichen Diskursrelationen zur Verfügung, die kohärente und ganzheitliche Interpretation erlauben.

Wie wir in Wildfeuer (2014b) darstellen, sind zwei der verfügbaren formalen Ansätze für eine Analyse multimodaler Artefakte von besonderer Relevanz, da sie ihren Fokus zum einen auf die Beschreibung der Repräsentationsebene des Diskurses, zum anderen auf die Darstellung der Diskursstruktur und somit die Analyse von Relationen zwischen Einheiten im Diskurs legen. Es handelt sich dabei um die Diskursrepräsentationstheorie (DRT) nach Hans Kamp und Kollegen (vgl. KAMP 1981; KAMP/REYLE 1993) sowie die darauf aufbauende Segmentierte Diskursrepräsentationstheorie (SDRT; vgl. ASHER/LASCARIDES 2003), die beide formalisierte Beschreibungsmodelle bereitstellen, die auch für eine Übertragung auf multimodale Analysen geeignet sind. Besonders interessant ist, dass beide Theorien neben den semantischen Informationen eines Diskurses und den daraus konstruierten Repräsentationen auch pragmatische Inhalte in die Betrachtung miteinbeziehen, um alle durch die unterschiedlichen Einheiten zusammengebrachten Verhältnisse und Relationen berücksichtigen zu können. Sie bilden damit innerhalb ihrer Instrumentarien die von uns beschriebene Ebene der Diskurssemantik direkt ab.

Mit Blick auf unterschiedliche multimodale Artefakte konnten wir in einigen Arbeiten der letzten Jahre (vgl. z.B. BATEMAN/SCHMIDT 2012; WILDFEUER 2012, 2014 für Film; BATEMAN/WILDFEUER 2014a, 2014b; WILDFEUER/BATEMAN 2014 für Comic; WILDFEUER/SCHNELL/SCHULZ 2015; WILDFEUER 2017b für Online-Diskurse) bereits aufzeigen, dass sich diese Ansätze als besonders fruchtbar für eine empirische Überprüfbarkeit erweisen, wenn auch die tatsächliche empirische Arbeit auf Grundlage der vorgenommenen methodologischen Präzisierung erst noch zu erfolgen hat. Wir werden aufbauend auf den bereits erfolgten Ausdifferenzierungen im Folgenden deswegen nur einen kurzen Überblick über die methodologischen Feinheiten unseres Instrumentariums geben, diese aber zugleich mithilfe einer Beispielanalyse veranschaulichen.

Dafür wählen wir ein bereits in Krafft (1978) diskutiertes (und von Stephan Packard in einem Vortrag in Tübingen wieder aufgegriffenes<sup>8</sup>) Beispiel aus der Comicserie *Achille Talon* von Michel Greg (1967). Die kurze Sequenz *Sans Sourd Ciller* der Ausgabe *Achille Talon aggrave son cas* (GREG 1967: 22-23) erzählt, wie der Protagonist der Serie einen Spaziergang durch die Stadt macht und dabei auf unterschiedliche Personen und Ereignisse trifft. Erstere grüßen und reden mit ihm, allerdings wird ihr jeweiliger Redebeitrag in den Sprechblasen nur durch Linien oder Punkte und Fragezeichen realisiert und somit als unverständlich markiert (vgl. Panel 1 und 2 in Abbildung 2 als Beispiele). Auch die vielen und teilweise sehr lauten Geräusche der Straße und Umgebung scheinen den Protagonisten nicht zu stören. Erst zum Ende der Sequenz reagiert er auf einen Kommentar einer Figur und erklärt dann, dass die kleinen Wachs-kügelchen, die man sich in die Ohren steckt, um die idiotischen Geräusche um sich herum nicht mehr zu hören, wirklich fantastisch seien (vgl. das letzte Panel in Abbildung 2).

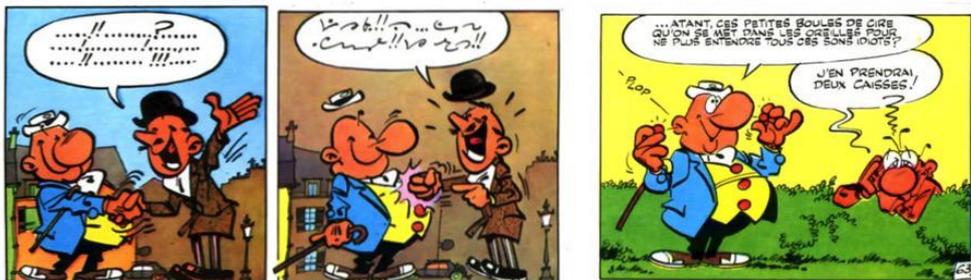


Abb. 2:  
Beispielpanels aus Michel Greg, *Sans sourd ciller* (GREG 1967: 22-23), diskutiert in Krafft (1978: 15-23)

Während Krafft (1978: 15-23) auf die kohäsiven Gestaltungsmittel innerhalb dieser Sequenz hinweist und die Identifikation des Protagonisten in den unterschiedlichen Panels mithilfe von Verweisketten erklärt, wollen wir im Folgenden vor allem auf das intersemiotische Zusammenspiel von Bild- und Textebene innerhalb eines Panels einerseits sowie die narrative Struktur der Gesamtsequenz andererseits eingehen. Ersteres ist insofern interessant, da insbesondere die schriftlichen Informationen, die die Interpunktion innerhalb der Sprechblasen liefert, zunächst unverständlich bleiben und vielfach interpretiert werden können. Dies ist vermutlich vom ›Autor/Zeichner des ›Textes‹ genau so intendiert und folglich wird diese Interpretation mithilfe vieler pragmatisch orientierter Ansätze auch als plausibel anerkannt: der Autor/Zeichner hat volle Kontrolle über das Repräsentierte, er/sie ›weiß‹, was die Charaktere sagen oder nicht sagen und LeserInnen können deswegen annehmen, dass, wenn diese Informationen wie im vorliegenden Panel zurückgehalten werden, der/die

<sup>8</sup> Wir berufen uns hier auf den Vortrag »Gestaffelte Sichtbarkeiten«, den Stephan Packard auf einem Kolloquium zum Thema Multimodalität in Tübingen im Februar 2015 gehalten hat, vgl. auch sein Beitrag in diesem Band.

Autor/in eine bewusste Entscheidung darüber getroffen hat – und die Informationen in diesem Fall für das Verständnis und die Interpretation des Panels auch nicht relevant sind. Diese Auffassung folgt dem traditionellen Konzept der Motivation solcher Artefakte (vgl. KOCH 2015), das für Effekte dieser Art immer einen spezifischen Grund sucht.

Zugleich veranschaulicht das Beispiel aber auch, dass der/die AutorIn/ZeichnerIn mit den Möglichkeiten der Zeichenmodalität Comic spielt und hier insbesondere die unterschiedlichsten Gestaltungsdetails der Textebene in Sprechblasen (vgl. auch FORCEVILLE 2011; COHN 2013) einsetzt. Die Inhalte geben in diesem Fall somit nicht einfach wieder, was in der Geschichte gesprochen wird, sondern müssen anderweitig interpretiert werden. Die Alternativen dafür sind vielfältig, neben Unverständlichkeit allgemein kann zum Beispiel auch geschlossenfolgert werden, dass die Figuren unterschiedliche Sprachen sprechen, die der Protagonist aber nicht verstehen kann. Dass sich diese und andere Interpretationen in der weiteren Rezeption des Comics tatsächlich als nicht ganz richtig herausstellen, wird erst durch die Hinzunahme der Informationen in den anderen Panels und insbesondere die vom Protagonisten selbst gegebene Erklärung zum Schluss der Sequenz deutlich (vgl. Panel 3 in Abbildung 2). Erst dann kann den Linien und Fragezeichen in den Sprechblasen nachträglich die aufgrund des Kontextes plausible Bedeutung zugewiesen werden, dass der Protagonist die Beiträge aufgrund der Wachskugeln in seinen Ohren akustisch nicht versteht. Diese Interpretation ist das Ergebnis einer dynamischen und sich im Laufe des Narrativs verändernden Hypothesenbildung über die Bedeutung des Sprechblaseninhalts, für die die narrative Struktur der Gesamtsequenz eine Erklärung gibt. Wir werden diesen dynamischen Interpretationsprozess mithilfe des im Folgenden beschriebenen Instrumentariums detailliert nachweisen.

### 3.2 Multimodale Diskursanalyse

Zur Analyse aller im multimodalen Artefakt verfügbaren Einheiten und ihres Zusammenspiels hat sich, wie wir vielfach nachweisen konnten und wie bereits in der ersten Beschreibung einer Interpretation des Comicbeispiels oben deutlich wird, eine zweigeteilte Analyse auf zwei Ebenen des multimodalen Verstehens (vgl. BUCHER 2011) als hilfreich erprobt. Dabei sollte zunächst auf einer ersten Ebene eine Identifizierung und Gruppierung aller bedeutungstragenden Elemente im Diskurs bzw. Kommunikationsraum stattfinden. Das Wissen über das entsprechende Medium (in diesem Fall der Comic) und die in ihm verfügbaren Zeichenmodalitäten ist hier wichtig, da es eine Filterung und Beschreibung der relevanten Modalitäten ermöglicht, die wiederum auf einer zweiten Ebene der Kohärenz und strukturellen Erschließungspfade sinnvoll miteinander in Beziehung gesetzt werden (vgl. BUCHER 2011: 17). Diese beiden Ebenen entsprechen zum einen und in großem Maße auch den für die Definition der Zeichenmodalität verwendeten Strata, wobei sie sich aber vor allem auf den

oberen beiden Schichten ansiedeln, die das geformte Material und die ihn umgebende Diskurssemantik betreffen. Zum anderen nehmen diese beiden Ebenen die Bereiche der semantischen und pragmatischen Informationsverarbeitung im Diskurs ein, die es für eine umfassende Analyse zu berücksichtigen gilt (siehe oben).

Unser Instrumentarium arbeitet ebenfalls auf diesen beiden Ebenen und stellt dafür in Anlehnung an die Arbeiten von Asher/Lascarides (2003) zwei Teillogiken zu Verfügung, die eine genauere Beschreibung, Analyse sowie Formalisierung der beiden Ebenen ermöglichen. Aufgrund seiner formal-logischen Konzeption bezeichnen wir unser Instrumentarium als so genannte *Logik der multimodalen Diskursinterpretation* (vgl. Abbildung 2; WILDFEUER 2014a; WILDFEUER 2014c; BATEMAN/WILDFEUER 2014a).

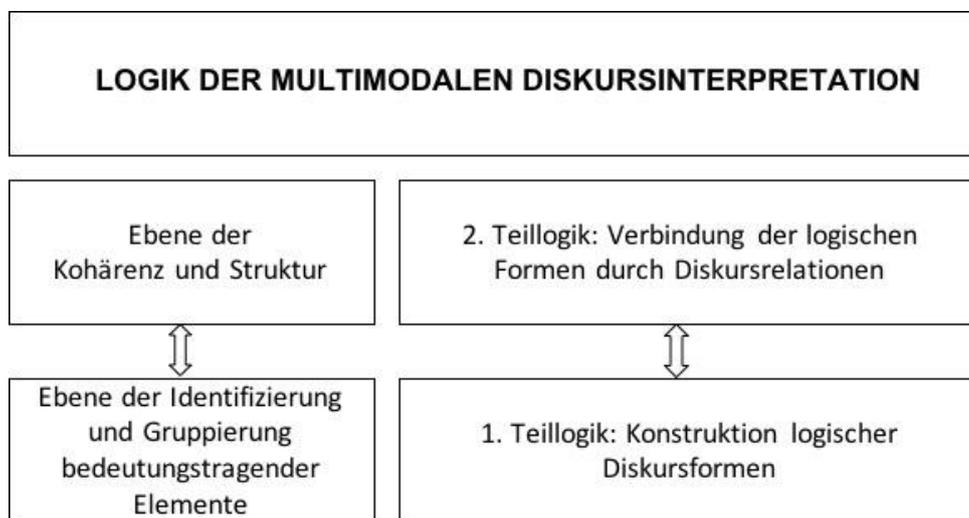


Abb. 3:  
Die zwei Ebenen der Logik der multimodalen Diskursinterpretation

### 3.2.1 Zur Analyse des semantischen Gehalts in multimodalen Artefakten

Die erste Logik innerhalb dieses Instrumentariums ist die des Informationsgehaltes (*logic of information content*, vgl. ASHER/LASCARIDES 2003), in der so genannte logische Formen des Diskurses konstruiert werden, die die interagierenden semiotischen Ressourcen identifizieren und ihren gemeinsamen semantischen Gehalt aufzeigen. Die Teillogik umfasst dafür ein Beschreibungsvokabular, das aus grundlegenden Ansätzen der multimodalen Diskursanalyse zur Beschreibung narrativer und darstellender Prozesse (vgl. KRESS/VAN LEEUWEN 1996, 2001; VAN LEEUWEN 2005) übernommen wird. So ist es möglich, die Ereignisse oder Darstellungen im Artefakt aufgrund der Intersemiose der

unterschiedlichen Elemente im tatsächlichen Material als so genannte Eventualität, also als eine mögliche Annahme, die durch die aktive Semiose des Rezipienten interpretiert, aber grundsätzlich auch durch eine anderslautende Proposition bezeichnet werden kann, zu beschreiben. Als Darstellungsmittel für diese Beschreibung wählen wir – in Anlehnung an die Arbeiten von Kamp und Kollegen (vgl. KAMP/REYLE 1993) – Diskursrepräsentationsstrukturen (DRS), die das Format einer Box haben, in der alle Diskursreferenten, d.h. alle in der Interpretation beteiligten Einheiten, aufgelistet werden. Ein typisches Beispiel für solch eine Box findet sich rechts in Abbildung 4 und stellt als Ergebnis die semantische Repräsentation des Comic-Panels links dar.



Abb. 4:  
Diskursrepräsentationsstruktur für ein Panel aus Achille Talon (GREG 1967: 22)

Die Diskursrepräsentationsstruktur enthält in der ersten Zeile eine Beschreibung des semantischen Gehaltes, von dem wir annehmen, dass er normalerweise so von RezipientInnen aufgrund der im unteren Teil der Box aufgelisteten Diskursreferenten inferiert wird: Wir erkennen im Bild die beiden Partizipanten a und b (vgl. KRESS/VAN LEEUWEN 2006)<sup>9</sup>, die sich gegenseitig die Hand geben. Aus dieser Handlung können wir schließen, dass sich die beiden begrüßen, was wir mit  $e_{\pi_i} = \text{begrüßen}$  auch formal darstellen. Es handelt sich bei dieser Beschreibung um eine *abduktive* und mithilfe aktiver Semiose getroffene Hypothese, die Welt- und Kontextwissen über die bildliche Gestaltung, die möglicherweise bekannten Charaktere und den Comic miteinbezieht. Abduktion ist nach Peirce (1931-1958) die Suche nach der bestmöglichen Interpretation aufgrund von Hypothesen, die im Diskurs durch das Hinzukommen neuer Informationen auch verworfen werden können. In der letzten Zeile der Box wird dieser Inferenzprozess noch einmal genauer beschrieben, indem diejenigen Diskursreferenten genannt werden, die für die anfechtbare

<sup>9</sup> Der repräsentierte Partizipant links im Panel kann aufgrund von spezifischem Diskurs- und Kontextwissen über die Comicserie auch als der Protagonist Achille Talon identifiziert werden.

Schlussfolgerung (ausgedrückt durch den Operator  $\vdash$ , siehe auch weiter unten) eine Rolle spielen und so im Ergebnis den semantischen Gehalt ergeben.

Diese Diskurshypothese bringt zum einen eine genauere Interpretation der Linien im Bild mit sich, die – wie wir bereits im Beispiel von McCloud oben gezeigt haben – aufgrund ihres spezifischen Kontextes auch eine spezifische Bedeutung haben. Als Elemente in der Nähe der Hände der Partizipanten ist es in diesem Fall plausibel, sie als Bewegungslinien zu interpretieren, die eine Bewegung der Hände andeuten und somit auch die Interpretation der Handlung als  $e_{ni} = \text{begrüßen}$  stützen. Die Interpretation dieser Handlung impliziert dann, dass die beiden Partizipanten miteinander reden bzw. dass der eine Partizipant zum anderen etwas sagt. Unser ›lexikogrammatiches‹ Wissen (vgl. Abb. 1 und Kap. 2.3) über Sprechblasen in Comics bestätigt diese Interpretation und lässt uns außerdem davon ausgehen, dass der zweite Partizipant hört, was der erste ihm sagt, auch wenn der/die AutorIn/ZeichnerIn uns dies möglicherweise nicht mitteilen will und deswegen die Ausrufe- und Fragezeichen verwendet (vgl. unsere Ausführungen zur Motivation oben). So ist der Inhalt der Sprechblase zwar Teil der Repräsentationsstruktur und fließt in die Interpretation von  $e_{ni} = \text{begrüßen}$  mit ein. Er scheint aber – wie in der letzten Zeile der Box deutlich wird, in der (c) nicht erwähnt wird – keine zentrale Rolle einzunehmen und nicht so relevant wie andere Informationen im Panel zu sein. Auch eine verständliche Aussage wie ›Guten Tag!‹ zum Beispiel würde die Inferenz des semantischen Gehalts als  $e_{ni} = \text{begrüßen}$  nicht groß verändern, da sie lediglich die entsprechende Formel zur Begrüßung repräsentiert.

Unsere Darstellungsweise ermöglicht demnach neben der präzisen Beschreibung der für das Verständnis der Ereignisse in diesem Panel notwendigen Inferenzprozesse auch eine genauere Sicht auf die intersemiotischen Zusammenhänge der einzelnen verfügbaren Modalitäten und Ressourcen. Es sind hier vor allem die Linien und Farben auf der visuellen Ebene, die die Charaktere und ihrer Handlungen darstellen und so zur Inferenz des semantischen Gehalts führen. In anderen Panels der Sequenz können auch weitere Details der Bildebene wie Bewegungslinien oder ein wechselnder farbiger Hintergrund als relevante Elemente des Narrativs fungieren und in Verweisketten eine Rolle spielen (vgl. KRAFFT 1978: 18). Die boxartige Darstellung mithilfe unseres Instrumentariums bietet hierfür unterschiedlichste Beschreibungsmöglichkeiten, diese Elemente als Diskursreferenten in die Auflistung mitaufzunehmen. Jedes narrative Ereignis bzw. jede Eventualität kann somit als logische Diskursform und Repräsentationsstruktur gefasst werden.

Neben dieser Auflistung von Elementen ist auch ihr intersemiotischer Zusammenhang innerhalb der Diskursrepräsentationsstruktur von Belang. Insbesondere mit Blick auf das spätere Panel spielt hier sowohl die jeweilige Form und Gestaltung der Sprechblase als auch der zu interpretierende Inhalt eine Rolle. Wie diese im kohärenten Ganzen der narrativen Struktur zu analysieren sind, werden wir mithilfe der zweiten Ebene unseres Instrumentariums verdeutlichen.

### 3.2.2 Zur Analyse der Diskursstruktur multimodaler Artefakte

Auf dieser zweiten Ebene der Kohärenz werden die logischen Formen bzw. narrativen Ereignisse zu Diskursstrukturen verbunden, indem Diskursrelationen zwischen den einzelnen Einheiten inferiert werden. Die zweite Teillogik stellt dafür ein Kompendium von Diskursrelationen für jedes Artefakt zur Verfügung, das nicht nur detaillierte Definitionen, sondern auch hinreichend zu erfüllende Bedingungen zur Interpretation dieser Relationen umfasst. Bei der zweiten Teillogik handelt es sich demnach um diejenigen Mechanismen, die wir mit Blick auf das Stratum der Diskurssemantik bereits angedeutet haben.

Diskurs- oder auch rhetorische Relationen (vgl. z. B. MANN/THOMPSON 1987) werden innerhalb diskursanalytischer Ansätze mithilfe abduktiver Prinzipien der Bedeutungskonstruktion formal bestimmt. Im verbalen Diskurs basiert dieser Interpretationsprozess auf der allgemeinen Annahme einer kohärenten Korrelation zwischen zwei oder mehr Propositionen als Resultat des Griceschen Kooperationsprinzips (vgl. GRICE 1975). Eine Analyse der Relationen aufgrund dieser Prinzipien ist deswegen immer anfechtbar, kann aber mithilfe allgemein gültiger Annahmen und fester Interpretationsregeln für den jeweiligen Kontext als normalerweise so vom Rezipienten verstanden aufgefasst werden. Die zugrunde liegende nicht-monotone Logik stellt für den Interpretationsprozess verschiedene Inferenzmuster zur Verfügung, die vor allem im Rahmen der Künstlichen Intelligenz eine wichtige Rolle spielen und der Illustration von Wissensverarbeitungsprozessen dienen (vgl. zum Beispiel RUSSEL et al. 2012). Diese Muster führen zu der bereits angesprochenen anfechtbaren Schlussfolgerung (*defeasible inference*), die auch für die Inferenz der Diskursrelationen als Mechanismus geltend gemacht wird.<sup>10</sup> Asher/Lascarides (2003) nutzen diese abduktiven Prinzipien der Interpretation von Diskursrelationen, indem sie im Rahmen ihres Instrumentariums konkrete Definitionen und Defaultaxiome bereitstellen, die für die Inferenz der jeweiligen Relation im Diskurs erfüllt sein müssen. Diese Bedingungen, die ebenfalls mit formallogischen Mitteln dargestellt werden (vgl. WILDFEUER 2014b), beschreiben Zustände, die es ermöglichen, die unterspezifizierten Informationen aufgrund ihrer pragmatischen Präferenziertheit im jeweiligen Kontext zu bestimmen.

Auch für die Analyse multimodaler Artefakte hat es sich als sinnvoll erwiesen, ein fest definiertes Set von Diskursrelationen aufzustellen, das es ermöglicht, die auf der ersten Ebene des Instrumentariums gefunden Diskursrepräsentationsstrukturen miteinander und damit zu größeren Strukturen zu verbinden. Für Film und Comics liegen diese Sets bereits dokumentiert vor (vgl. BATEMAN/SCHMIDT 2012 und WILDFEUER 2014a für Film; BATEMAN/WILDFEUER 2014 für Comic), auch für einige Analysen einzelner anderer Artefakte haben sich die

---

<sup>10</sup> Ein bekanntes Muster für diese Schlussfolgerungen ist z.B. der *Defeasible Modus Ponens* (DMP; dt. Anfechtbarer Modus Ponens). Für genauere Ausführungen zu diesen Mustern sei auf die Ausführungen in Kapitel 5 in Asher/Lascarides (2003: 179-248) verwiesen.

sieben von Asher/Lascarides (2003) aufgestellten Relationen als praktikabel erwiesen (vgl. WILDFEUER/SCHNELL/SCHULZ 2015; WILDFEUER 2018).

Wir führen im Folgenden eine Liste der sieben typischen Relationen zur Verbindung von Diskurseinheiten in narrativen Texte auf, weisen aber explizit darauf hin, dass je nach Medium und seinen spezifischen Anforderungen jeweils eine Anpassung der Defaultaxiome und Bedeutungspostulate erfolgen sollte, bei der es dann auch zu einer Erweiterung des Sets kommen kann (vgl. BATEMAN/WILDFEUER 2014a: 195; WILDFEUER/BATEMAN 2014: 6).

<b>Diskursrelation</b>	<b>Notwendige Bedingungen / Verhältnisse zwischen den Einheiten und im Kontext</b>
<i>Narration</i>	Zeitliche(-örtliche) Aufeinanderfolge: b folgt a; gemeinsames Thema
<i>Elaboration</i>	Teil-Ganzes-Beziehung; Spezifizierungsverhältnis;
<i>Explanation</i>	Zeitliche Aufeinanderfolge: a folgt b; b gibt Erklärung für a; im Diskurs verankerter Grund/Erklärung
<i>Result</i>	Zeitliche Aufeinanderfolge: b ist Ergebnis/Resultat von a; im Diskurs verankerter Grund/Erklärung
<i>Background</i>	Zeitliche Aufeinanderfolge: b folgt a; zwei unterschiedliche Zustände; Hintergrundinformationen
<i>Contrast</i>	Strukturelle Ähnlichkeit; semantische Unähnlichkeit
<i>Parallel</i>	Strukturelle Ähnlichkeit; semantische Ähnlichkeit

Tab. 1:

Übersicht über die von Asher/Lascarides beschriebenen Diskursrelationen zur Verbindungen zweier Diskurseinheiten / Aussagen ›a‹ und ›b‹.

Formallogisch werden diese Relationen dann jeweils in Form der bereits angesprochenen Bedeutungspostulate und Defaultaxiomen dargestellt. Der Einfachheit halber verzichten wir hier auf diese formalen Beschreibungen und verweisen für Details auf die von uns ausführlich diskutierten Arbeiten zu Comic und Film (vgl. BATEMAN/SCHMIDT 2012; WILDFEUER 2014a; BATEMAN/WILDFEUER 2014a; WILDFEUER/BATEMAN 2014). Alle Formeln beinhalten jeweils eine Spezifizierung des zeitlich-örtlichen Verhältnisses der Einheiten im Diskurs sowie die Angabe einer Bedingung, die erfüllt sein muss, damit der Inferenzprozess erfolgreich ist.

Im Rückgriff auf unser Beispiel ist es möglich, mithilfe dieser Beschreibungsebene die in Kapitel 3.2.1 aufgestellten Hypothesen zur Semantik des ersten Panels zu überprüfen und diskurssemantisch zu evaluieren. Nachdem für jedes Panel der Sequenz mithilfe der ersten Logik Diskursrepräsentationsstrukturen beschrieben werden konnten, können diese mithilfe der zweiten Logik und den darin verfügbaren Diskursrelationen sinnvoll miteinander verbunden werden. Die Ereignisse in den einzelnen Panels können größtenteils als zeitlich aufeinanderfolgend verstanden werden, außerdem bedingen sie sich teilweise gegenseitig bzw. teilen ein Thema: der gezeigte Spaziergang durch die Stadt führt durch bzw. über Straßen mit mehr oder weniger Verkehr und

an unterschiedlichen Häuserzügen und Plätzen mit unterschiedlich agierenden Menschen, die singen, Lärm machen etc. vorbei. Aufgrund dieser im Kontext herrschenden Bedingungen lassen sich vor allem Narration-Relationen inferieren (vgl. die Bedingungen in Tabelle 1). Auch eine Eventualität für das letzte Panel der Sequenz, das als  $e_{nj} = \text{herausnehmen}$  (der Wachskugeln aus den Ohren) zu interpretieren ist, ist so in die Struktur eingebettet und folgt den vorherigen Ereignissen zeitlich. Durch die außerdem verfügbaren verbalen Informationen in der Sprechblase, in der der Protagonist über die Wachskugeln spricht, kann sogar eine zweite Eventualität für dieses Panel interpretiert werden, die sich auf den sprachlichen Textanteil bezieht und mit  $e_{nk} = \text{erklären}$  als Sprechhandlung des Protagonisten beschrieben werden kann.

Diese Informationen nehmen, wie oben bereits ausgeführt, Einfluss auf unsere zuvor getroffene Analyse der Ausrufe- und Fragezeichen im ersten Panel. Beide Eventualitäten können nämlich eine Erklärung für die Ereignisse in den ersten Panels der Sequenz liefern, indem sie einen Grund dafür angeben, warum der Sprechblaseninhalte in den vorherigen Panels keine expliziten Informationen vermittelt. Sie leiten so die Inferenz einer Explanations-Relation an (vgl. die Bedingungen in Tabelle 1), die im Grunde alle vorherigen Ereignisse im Nachhinein erklärbar macht.

Unsere zuvor aufgestellten Hypothesen über die Rolle der Sprechblase und ihres Inhalts müssen demnach folgendermaßen korrigiert werden: Zwar sagt der eine Partizipant etwas zum anderen, allerdings – so wird durch die Erklärung deutlich – hört letzterer den Gruß des Ersteren nicht und versteht ihn somit auch nicht. Diese Interpretation ist das Ergebnis einer dynamischen und sich im Laufe des Diskurses verändernden Inferenzziehung über die Bedeutung des Sprechblaseninhalts. Hier kommt dann auch die Rolle der Kombination der unterschiedlichen Ressourcen wieder ins Spiel, wenn für die Korrektur der Interpretation die typographischen Elemente auf der verbal-textuellen Ebene zusammen mit den bildlichen Elementen interpretiert werden und hierfür dann keine Motivation des Autors/Zeichners mehr angenommen werden kann, diesen Inhalt zu verschleiern. Die Inferenz der Situation im Panel ( $e_{ni}$ ) operiert damit ganz eindeutig über die unterschiedlichen semiotischen Ressourcen hinweg und erklärt, dass die Frage- und Ausrufezeichen das Unverständnis des Partizipanten signalisieren.

Aus narratologischer Perspektive haben wir es hier mit einer neu zu interpretierenden Fokalisierung zu tun, die wir mit unserem Instrumentarium nun genauer beschreiben können. Während es sich zu Beginn der Diskursinterpretation noch um eine neutrale Null-Fokalisierung handelt, bei der der Erzähler mehr weiß als die jeweiligen Figuren über einander, ändert sich diese Interpretation zum Ende der Sequenz und nach Erläuterung der Verhältnisse zu einer internen Fokalisierung, bei der der Inhalt der Sprechblase durch die Perception der nicht-hörenden Figur gefiltert und erst dann den RezipientInnen präsentiert wird. Mit dieser Erläuterung können wir die Aussage Horstkotte und Pedris (2011: 35) stützen, dass Fokalisierung auf Diskursebene stattfindet und – so unsere Ergänzung – immer ein Ergebnis der Hypothesenfindung im

Diskurs ist. Das Element der Fokalisierung in dieser Comicsequenz ist dann nur ein weiteres Beispiel dafür, Comics durchaus als Metamedien auffassen zu können, die oft stärker als viele andere Medien mit ihren Eigenheiten und Möglichkeiten zu spielen wissen (vgl. dazu auch BATEMAN et al. 2017a: Kap. 12).

Unser Instrumentarium ermöglicht, diese Feinheiten in der Gestaltung mithilfe der Rekonstruktion der Diskursinterpretationen nachzuweisen. Neben der formalen Beschreibung der einzelnen Relationen zur Dokumentation der Inferenzprozesse kann auch die gesamte Diskursstruktur der Comicsequenz dargestellt werden. Neben der Explanation-Relation werden dafür auch die anderen Diskursrelationen aufgezeigt und schließlich die gesamte Diskursstruktur konstruiert, um so die für eine sinnvolle Interpretation notwendige Kohärenz nachzuweisen. Alle Diskursrepräsentationsstrukturen können in dieser sich dynamisch entfaltenden Struktur miteinander verbunden und so als Interpretationsprozess der RezipientInnen dargestellt werden.<sup>11</sup> Insbesondere können dabei auch die viel diskutierten *textual cues*, d.h. gezielte Hinweise im Text zur Lenkung der RezipientInnen (vgl. BATEMAN/WILDFEUER 2014a oder auch BORDWELL 1989 für Film), herausgearbeitet werden, die zum Beispiel durch eine bestimmte Diskursrelation (z.B. der Kausalität) angezeigt werden. In unserem Fall ist ein solcher Hinweis durch die Explanation-Relation gegeben, die neben der intradiegetischen Erklärung auch eine Erklärung für den in der Aufeinanderfolge der Panels konstruierten Witz gibt. Die kausale Relation klärt die LeserInnen sozusagen darüber auf, welche Details verstanden werden müssen, damit die Geschichte sinnvoll und zugleich auch unterhaltsam ist. Indem die Comicsequenz die unterschiedlichen Möglichkeiten der Zeichenmodalität und ihres Materials einsetzt und diesen Einsatz im Nachhinein sogar erläutert, macht sie ihre humoristische Argumentation sichtbar.

#### 4. Zusammenfassung und Ausblick

Eine detaillierte Analyse von Feinheiten in der Bedeutungszuweisung im sich dynamisch verändernden Kontext kann, so haben wir mit unseren Ausführungen oben gezeigt, nur mit Blick sowohl auf die materiellen Eigenschaften des multimodalen Artefakts als auch auf seine diskurssemantischen Besonderheiten erfolgen. Das von uns vorgestellte Instrumentarium folgt damit der zuvor vorgenommenen theoretischen Präzisierung des Konzeptes der Zeichenmodalität und ermöglicht insbesondere mit der Bereitstellung von Mechanismen auf der Ebene der Diskurssemantik, die unterschiedlichen Prozesse multimodaler Bedeutungskonstruktion ausführlich darzustellen.

Wie wir anhand der unterschiedlichen Analyseschritte beschreiben konnten, haben wir zum einen eine entsprechende Grundlage für die Überprüfung unserer Hypothesen in größeren Datenmengen. Diese Hypothesen

---

<sup>11</sup> In unseren formallogischen Beschreibungen benutzen wir auch für diese Darstellung Boxen, die die jeweiligen Diskursrelationen auflisten. Eine weitere graphische Repräsentation ähnlich einer Baumstruktur veranschaulicht außerdem die Zusammenhänge (vgl. z.B. WILDFEUER 2014a: 75-77).

nämlich sind nun derart präzise (und in formallogischen Beschreibungen) darstellbar, dass wir sie auf eine Vielzahl von Artefakten innerhalb unterschiedlicher Korpora anwenden können. Als ein weiteres Anwendungsgebiet unserer semantischen Analyse sehen wir außerdem den Brückenschlag zu anderen Disziplinen der Medien- und Kulturwissenschaft, die sich gleichermaßen mit diesen und weiteren multimodalen Artefakten befassen und mithilfe ihrer jeweiligen Zugänge ähnliche oder auch ganz anders geartete Interpretationen finden. Für beide Perspektiven möchten wir im Folgenden einen kurzen Ausblick vornehmen.

#### **4.1 Empirische multimodale Analyse**

Auf Grundlage unserer terminologischen und methodologischen Ausdifferenzierungen ist es nun dringend notwendig, die von uns beschriebenen Details in der Bedeutungskonstruktion in einer Menge von Beispielen empirisch zu überprüfen und somit mehr Informationen über die Eigenheiten des Mediums Comic und seine Zeichenmodalitäten zu erfahren.

Tatsächlich nimmt die Quantität an verfügbaren multimodalen Daten beständig zu und ist es aufgrund der technischen Fortschritte heute einfacher denn je, diese Daten in ihrer Komplexität aufzuzeichnen und zu speichern. Beispielsweise müssen Gesprächsdaten heute nicht mehr nur schriftlich transkribiert, sondern können per Video aufgenommen werden, so dass neben den bisher verfügbaren Informationen über Intonation, Pausen etc. auch Daten über Gestik und Mimik, Gesprächssituation, Hintergrundgeräusche etc. verfügbar sind. Auch der Zugriff auf Filmsequenzen und Comicseiten ist ohne größeren Kosten- und Materialaufwand möglich und es ist keine Seltenheit mehr, dass beträchtlichere Mengen dieser Informationen in digitalen Sammlungen abgespeichert und für die breite Öffentlichkeit verfügbar gemacht werden. Der Zugriff auf das ›rohe‹ Material für ausführliche multimodale Analysen stellt damit kein unüberwindbares Hindernis mehr dar.

Die Verarbeitung dieses Materials im Hinblick auf die vorzunehmende feingliedrige Analyse hat dagegen noch mit einigen Herausforderungen zu kämpfen, die vor allem jene Arbeitsschritte betreffen, die über eine Beschreibung der vorhandenen Einheiten hinausgehen. Zwar sind heute viele unterschiedliche Tools zur Bearbeitung audiovisueller Daten verfügbar, die eine Transkription und Annotation von Filmsequenzen oder auch Videogesprächsdaten ermöglichen (u.a. z.B. ELAN, ANVIL etc.; vgl. BATEMAN 2014b). Die direkte Verbindung dieser Transkriptionen mit den theoretischen Hypothesen unserer Analysen (z.B. in Form von Diskursrelation als Manifestation der Inferenzprozesse) stellt sich allerdings immer noch als besonders schwierig heraus. Wie wir in Bateman (2014b) hervorheben, ist dies ein allgemeines Problem multimodaler empirischer Analyse und vor allem auf die enorme Menge an Einzelheiten im Artefakt zurückzuführen:

The gap between transcription (still predominantly textual) and the material under analysis is simply too great to allow re-presentation. In the multimodal case there is almost always simply too much occurring to consider any re-presentation of the data as ›complete‹ (BATEMAN 2014b: 244).

Hilfreich zur Überbrückung dieser Lücke sind vor allem empirische Ansätze mit mehreren Annotationsebenen, die gezielte Aspekte einiger Einheiten ›kodieren‹, anstatt sie durch andere Einheiten zu ersetzen, beispielsweise mithilfe einer XML-basierten Annotation (vgl. BATEMAN 2014b: 247-251; BATEMAN et al. 2017b) oder aber auch mit anderen formalisierten Methoden wie den im Vorhinein vorgestellten. Die bereits für eine computergesteuerte Verarbeitung natürlicher Sprache entworfene (vgl. ASHER/LASCARIDES 2003) und nun auch für multimodale Daten aufgearbeitete formallogische Herangehensweise strebt nicht nur eine Modellierung menschlicher Informationsverarbeitung an, sondern ermöglicht auch, diese Modellierung anhand großer Datenmengen empirisch durchzuführen.

Für den Umgang mit multimodalen Korpora bedeutet dies eine deutlich umfangreichere Auseinandersetzung mit unterschiedlichen Ebenen der Analyse, die bereits in den theoretischen Grundlagen festgehalten ist und in der praktischen Arbeit dann Anwendung finden muss. Dazu gehören neben den technischen und materiellen Details auf einer niedrigeren Stufe (z.B. für Comics: Linienstärke, Farbabstimmung, Seitenlayout etc.; für Film: optischer Fluss der Bildsequenz, Farbabstimmung, Schnitttiefe) auch Transkriptionen aus unterschiedlicher Perspektiven auf die Daten (z.B. auf Sprache aus phonetischer oder intonatorischer Perspektive), der Einbezug von Ergebnissen experimenteller Arbeiten (z.B. aus Eye-Tracking-Studien, vgl. BATEMAN et al. 2016) sowie in besonderem Maße auch die Ebene abstrakterer Analysen und Hypothesenbildung. Für letztere bieten die vorgestellten linguistischen Arbeiten zur Multimodalität nun entsprechend detaillierte Analyseansätze mit der notwendigen theoretischen und methodologischen Präzision.

## 4.2 Interdisziplinäre multimodale Analyse

Mit Rückgriff auf empirisch evaluierbare Ergebnisse unserer Analysen sollten sich in einem weiteren Schritt dann auch verlässlichere Aussagen zu Hypothesen ergeben, die nicht (nur) aus linguistischer Perspektive an multimodale Artefakte herangetragen werden. Dies können zum Beispiel Fragen zu unterschiedlichen Medien und ihrer Spezifität, zu Genrestrukturen oder zu narratologischen Feinheiten sein – um hier nur einige typische Aspekte und Fragen einer interdisziplinären Medienwissenschaft zu nennen. Diese Hypothesen sind in vielen Fällen sehr theoretischer Natur und beziehen nur selten eine Analyse des Materials oder der grundlegenden semantischen Informationen in ihre Überlegungen mit ein. Eine Beschreibung der jeweils notwendigen Inferenzprozesse für eine Bedeutungszuweisung ist dagegen eine wichtige und

notwendige Grundlage für übergeordnete Interpretationen und stellt somit ein Verbindungselement für interdisziplinäre Zusammenarbeit dar.

Beispielsweise können unsere linguistisch-semiotischen Analysen den Einsatz bildlicher Repräsentationen zum Ausdruck emotionaler Zustände näher beleuchten und damit eine Brücke zu psychologischen oder auch neurokognitiven Ansätzen schlagen. Diese Repräsentationen werden bereits seit langem in unterschiedlichen Medien (z.B. in Film, Malerei und eben auch Comics) zur Wiedergabe interner emotionaler Zustände der Charaktere eingesetzt. Das bedeutet wiederum, dass es keine Realitätsgarantie für diese visuellen Repräsentationen gibt, sondern dass die Zuweisung von Bedeutung auf die dargestellten Einheiten und Ereignisse nie in ihrer visuellen Form vorgegeben sein kann und immer durch eine *Diskursinterpretation* im spezifischen Kontext geklärt werden muss. Dies betrifft jedes visuelle Material, ganz gleich ob es Charaktere der Diegese und ihre Handlungen, zusätzliche graphische Details wie Bewegungslinien oder Sprechblasen oder, wie in dem oben beschriebenen Falle, der Inhalt von Sprechblasen ist. Während die Interpretation dieser Zeichen zu Beginn des Interpretationsprozesses völlig offen bleibt, klärt erst das letzte Panel in der Sequenz sehr spät auf, was die Ausrufe- und Fragezeichen tatsächlich anzeigen und wie der emotionale Zustand des Protagonisten damit einzuschätzen ist.

Diese Interpretation wird durch die in der *Zeichenmodalität* Comic vorhandenen Möglichkeiten unterstützt und eine umfassende Analyse muss sich dementsprechend mit der Beschreibung dieser Möglichkeiten auseinandersetzen, um dann z.B. medien- oder genrespezifische Aussagen über die Konstruktion von Emotionen tätigen zu können. Daneben können unsere theoretischen Erkenntnisse zur kontextspezifischen Diskursinterpretation Grundlage für empirische Forschung z.B. in der Kognitionswissenschaft sein, indem sie prägnante Schlüsselstellen in den Artefakten herausarbeiten und für eine empirische Überprüfung vorbereiten (vgl. z. B. BATEMAN et al. 2018; TSENG et al. 2018).

Es lassen sich zahlreiche weitere Beispiele der interdisziplinären Zusammenarbeit zur multimodalen Analyse aufführen, in denen die von uns vorgestellte linguistisch-semiotische Präzisierung der Definition der Zeichenmodalität sowie eine methodologische Ausdifferenzierung von Nutzen sein können. Die Betrachtung aller in der Interpretation beteiligten Ebenen liefert eine so bisher nicht verfügbare, akkurate Beschreibung der Zeichenmodalität als grundlegende Einheit des Multimodalitätsparadigmas. In Verbindung mit der Betonung der für diese Interpretation notwendigen Diskursemantik wird damit eine klarere Darstellung der textuellen Funktionsweisen multimodaler Artefakte möglich, die bisherige pragmatisch-theoretische Ansätze um ihre dringend notwendige empirische Überprüfbarkeit erweitert.

## Literatur

- ANDROUTSOPOULOS, JANNIS: *Medienlinguistik. Sprachwissenschaftliche Zugänge zur Medienanalyse*. Tübingen [Narr] 2014
- ASHER, NICOLAS / ALEX LASCARIDES: *Logics of Conversation*. Cambridge [Cambridge UP] 2003
- BARDIN, LAURENCE: Le texte et l'image. In: *Communication et langages*, 26, 1964, S. 98-112
- BARTHES, ROLAND: *Elements of Semiology*. Translated by Annette Lavers and Colin Smith. London [Cape] 1964
- BATEMAN, JOHN A.: Using Multimodal Corpora for Empirical Research. In: JEWITT, CARE (Hrsg.): *The Routledge Handbook of Multimodal Analysis*. London [Routledge] 2014a, S. 238-252
- BATEMAN, JOHN A.: Dynamische Diskurssemantik als allgemeines Modell für Semiosen. Überlegungen am Beispiel des Films. In: *Zeitschrift für Semiotik*, 35(3-4), 2014b, S. 249-284
- BATEMAN, JOHN A.: Methodological and Theoretical Issues for the Empirical Investigation of Multimodality. In: KLUG, NINA-MARIA; HARTMUT STÖCKL (Hrsg.): *Handbuch Sprache im multimodalen Kontext*. Berlin [De Gruyter] 2016, S. 36-74
- BATEMAN, JOHN A.: Triangulating Transmediality: A Multimodal Semiotic Framework Relating Media, Modes and Genres. In: *Discourse, Context and Media*, 20, 2017, S. 160-174
- BATEMAN, JOHN A.; KARL-HEINRICH SCHMIDT: *Multimodal Film Analysis. How Films Mean*. London [Routledge] 2012
- BATEMAN, JOHN A.; ANNIKA BECKMANN; ROCIO INES VARELA: From Empirical Studies to Visual Narrative Organization: Exploring Page Composition. In: DUNST, ALEXANDER; JOCHEN LAUBROCK; JANINA WILDFEUER (Hrsg.): *Empirical Comics Research: Digital, Multimodal and Cognitive Methods*. London [Routledge] 2018, S. 127-153
- BATEMAN, JOHN A.; JANINA WILDFEUER; TUOMO HIIPPALA: *Multimodality. Foundations, Research, Analysis. A Problem-Oriented Introduction*. Berlin [De Gruyter] 2017a
- BATEMAN, JOHN A.; FRANCISCO O.D. VELOSO; JANINA WILDFEUER; FELIX CHEUNG; NANCY GUO: An Open Multilevel Classification Scheme for the Visual Layout of Comics and Graphic Novels: Motivation and Design. In: *Journal of Digital Scholarship in the Humanities*, 32, 2017b, S. 476-510
- BATEMAN, JOHN A.; CHIAO-I TSENG; OGNAN SEIZOV; ARNE JACOBS; ANDREAS LÜDTKE; MARION G. MÜLLER; OTTHEIN HERZOG: Towards Next-Generation Visual Archives: Image, Film and Discourse. In: *Visual Studies*, 31, 2016, S. 131-154
- BATEMAN, JOHN A.; MATTHIS KEPSEK; MARKUS KUHN: Film, Text, Kultur – Beiträge zur Textualität des Films. In: BATEMAN, JOHN A.; MATTHIS KEPSEK; MARKUS KUHN (Hrsg.): *Film, Text, Kultur: Beiträge zur Textualität des Films*. Marburg [Schüren] 2013, S. 7-29

- BATEMAN, JOHN A.; JANINA WILDFEUER: A Multimodal Theory of Visual Narrative. In: *Journal of Pragmatics*, 74, 2014a, S. 180-208
- BATEMAN, JOHN A.; JANINA WILDFEUER: Defining Units of Analysis for the Systematic Analysis of Comics: A Discourse-Based Approach. In: *Studies in Comics*, 5(2), 2014, S. 371-401
- BORDWELL, DAVID: *Making Meaning. Inference and Rhetoric in the Interpretation of Cinema*. Cambridge [Harvard UP] 1989
- BUCHER, HANS-JÜRGEN: Textdesign und Textwirkung in der massenmedialen Kommunikation. In: Roth, Kersten Sven; Jürgen Spitzmüller (Hrsg.): *Textdesign und Multimodalität. Zur Semantik und Pragmatik medialer Gestaltungsformen*. Konstanz [UVK] 2007, S. 49-76
- BUCHER, HANS-JÜRGEN; PETER SCHUMACHER: *Interaktionale Rezeptionsforschung. Theorie und Methode der Blickaufzeichnung in der Medienforschung*. Wiesbaden [Springer VS] 2011
- CLARK, AUSTEN: Cross Modal Links and Selective Attention. In: MACPHERSON, FIONA (Hrsg.): *The Senses: Classic and Contemporary Philosophical Perspectives*. New York [Oxford UP] 2011, S. 375–395
- COHN, NEIL: *The Visual Language of Comics. Introduction to the Structure and Cognition of Sequential Images*. London [Bloomsbury] 2013
- DIEKMANNSHENKE, HANS-JOACHIM: ›Schlagbilder‹. Diskursanalyse politischer Schlüsselbilder. In: DIEKMANNSHENKE, HANS-JOACHIM; MICHAEL KLEMM; HARTMUT STÖCKL (Hrsg.): *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. Berlin [ESV] 2011, S. 161-184
- ECKKRAMMER, EVA-MARIA; GUDRUN HELD (Hrsg.): *Textsemiotik. Studien zu multimodalen Texten*. Frankfurt/M. [Lang] 2006
- FORCEVILLE, CHARLES: Book Review: Multimodal Transcription and Text Analysis: A Multimedia Toolkit and Coursebook by Anthony Baldry and Paul J. Thibault. In: *Journal of Pragmatics*, 39(6), 2007, S. 1235–1238
- FORCEVILLE, CHARLES: Book Review: The Routledge Handbook of Multimodal Analysis. In: *Journal of Pragmatics*, 42, 2010, S. 2604–2608
- FORCEVILLE, CHARLES: Pictorial Runes in Tintin and the Picaros. In: *Journal of Pragmatics*, 43(3), 2011, S. 875–890
- FRAAS, CLAUDIA; STEFAN MEIER; CHRISTIAN PENTZOLD (Hrsg.): *Online-Diskurse. Theorien und Methoden transmedialer Online-Diskursforschung*. Köln [Herbert von Halem] 2013
- FRICKE, ELLEN: *Grammatik multimodal. Wie Wörter und Gesten zusammenwirken*. Berlin [De Gruyter] 2012
- GOODMAN, NELSON: *Languages of Art. An Approach to a Theory of Symbols*. London [Hackett] 1967
- GREG, MICHEL: *Achille Talon aggrave son cas. Tombe II*. Paris [Dargaud] 1967
- GRICE, HERBERT PAUL: Logic and Conversation. In: COLE, PETER; JERRY L. MORGAN (Hrsg.): *Syntax and Semantics*. New York [Academic Press] 1975, S. 41-58
- HALLIDAY, MICHAEL A. K.: *Language as Social Semiotic*. London [U Park P] 1978

- HALLIDAY, MICHAEL A.K.: *An Introduction to Functional Grammar*. London [Hodder Education] 1994
- HARRIS, ZELIG: Discourse Analysis. In: *Language*, 28, 1952, S. 1-30
- HARTMANN, KNUT: *Text-Bild-Beziehungen in multimedialen Dokumenten: Eine Analyse aus Sicht von Wissenrepräsentation, Textstruktur und Visualisierung*. Aachen [Shaker Verlag] 2002
- HJLEMSLEV, LOUIS: *Prolegomena to a Theory of Language*. Madison [U of Wisconsin P] 1961
- HORSTKOTTE, SILKE; PEDRI NANCY: Focalization in Graphic Narrative. In: *Narrative*, 19(3), 2011, S. 330-357
- JEWITT, CAREY (Hrsg.): *The Routledge Handbook of Multimodal Analysis*. London [Routledge] 2006
- JEWITT, CAREY (Hrsg.): *The Routledge Handbook of Multimodal Analysis*. London [Routledge] 2014
- JEWITT, CAREY; JEFF BEZEMER; KAY O'HALLORAN: *Introducing Multimodality*. London [Routledge] 2016
- KAMP, HANS: A Theory of Truth and Semantic Representation. In: GROENENDIJK, JEROEN A.G.; THEO M.V. JANSSEN; MARTIN B.J. STOKHOF (Hrsg.): *Formal Methods in the Study of Language*. Part 1. Amsterdam [Mathematisch Centrum] 1981, S. 277-322
- KAMP, HANS: Script Semantics II: Syllabus Semantics II. <http://www.ims.uni-stuttgart.de/~hans/Teaching/SemII.Scr.Prtl.1.21.12.07.pdf> [letzter Zugriff: 01.10.2015]
- KAMP, HANS; UWE REYLE: *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht [Springer] 1993
- KESSELHEIM, WOLFGANG: Sprachliche Oberflächen: Musterhinweise. In: HABSCHEID, STEPHAN (Hrsg.): *Textsorten, Handlungsmuster, Oberflächen. Linguistische Typologien der Kommunikation*. Berlin [De Gruyter] 2011, S. 337-366
- KLUG, NINA-MARIA; HARTMUT STÖCKL: Sprache im multimodalen Kontext. In: FELDER, EKKEHARD; ANDREAS GARDT (Hrsg.): *Handbuch Sprache und Wissen*. Berlin [De Gruyter] 2015, S. 242-264
- KLUG, NINA-MARIA; HARTMUT STÖCKL (Hrsg.): *Handbuch Sprache im multimodalen Kontext*. Berlin [De Gruyter] 2016
- KLUSS, THORSTEN; NICLAS SCHULT; KERSTIN SCHILL; MANFRED FAHLE; CHRISTOPH ZETZSCHE: Investigating the In-Between: Multisensory Integration of Auditory and Visual Motion Streams. In: *Seeing and Perceiving*, 25(1), 2012, S. 45-69
- KOCH, JONAS: *Erklären und Verstehen fiktionaler Filme: Semantische und ontologische Aspekte*. Münster [mentis] 2015
- KRAFFT, ULRICH: *Comics lesen. Untersuchungen zur Textualität des Comics*. Stuttgart [Klett-Cotta] 1978

- KRESS, GUNTHER: Commentary. Media Discourse – Extensions, Mixes, and Hybrids: Some Comments on Pressing Issues. In: *Text*, 24(3), 2004, S. 443-446
- KRESS, GUNTHER: *Multimodality. A Social Semiotic Approach to Contemporary Communication*. London [Routledge] 2010
- KRESS, GUNTHER; CAREY JEWITT; JON OGBORN; CHARALAMPOS TSATSARELIS: *Multimodal Teaching and Learning*. London [Bloomsbury] 2000
- KRESS, GUNTHER; THEO VAN LEEUWEN: Visual and Verbal Modes of Representation in Electronically Mediated Communication: The Potentials of New Forms of Text. In: SNYDER, ILANA (Hrsg.): *Page to Screen. Taking Literacy into the Electronic Area*. London [Routledge] 1998, S. 53-79
- KRESS GUNTHER; THEO VAN LEEUWEN: *Multimodal Discourse. The Modes and Media of Contemporary Communication*. London [Hodder Arnold Publication] 2001
- KRESS, GUNTHER; THEO VAN LEEUWEN: *Reading Images. The Grammar of Visual Design*. London [Taylor & Francis] 2006
- LUGINBÜHL, MARTIN: Medienkultur und Medienlinguistik. Komparative Textsortengeschichte(n) der amerikanischen »CBS Evening News« und der Schweizer »Tagesschau«. Bern [Peter Lang] 2014
- MACHIN, DAVID: *Introduction to Multimodal Analysis*. London [Hodder Arnold Publication] 2007
- MAIORANI, ARIANNA; CHRIS CHRISTIE (Hrsg.): *Multimodal Epistemologies: Towards an Integrated Framework*. London [Routledge] 2014
- MANN, WILLIAM C.; SANDRA A. THOMPSON: Rhetorical Structure Theory: A Theory of Text Organization. In: *Text*, 8(3), 1987, S. 243-281
- MARTIN, JAMES R: Process and Text: Two Aspects of Human Semiosis. In: BENSON, JAMES D.; WILLIAM S. GREAVE (Hrsg.): *Systemic Perspectives on Discourse. Volume 1: Selected Theoretical Papers from the Ninth International Systemic Workshop*. Nordwood [Ablex Publishing] 1985, S. 248-274
- MARTIN, JAMES R.: *English Text: Systems and Structures*. Amsterdam [John Benjamins] 1992
- MASSEY, IRVING: Words and Images: Harmony und Dissonance. In: *Georgia Review*, 34, 1980, S. 375-395
- MCCLOUD, SCOTT: *Understanding Comics: The Invisible Art*. New York [William Morrow] 1994
- MEIER, STEFAN: *Visuelle Stile. Zur Sozialesemiotik visueller Medienkultur und konvergenter Design-Praxis*. Bielefeld [transcript] 2014
- MITCHELL, W.J.T.: *Picture Theory. Essays on Verbal and Visual Representation*. Chicago [U of Chicago P] 1995
- MÖLLER-NASS, KARL-DIETMAR: *Filmsprache. Eine kritische Theoriegeschichte*. Münster [MAKS] 1986
- MONACO, JAMES: *How to Read a Film. Movies, Media, Multimedia*. Oxford [Oxford UP] 2000

- MONTAGUE, RICHARD: *Formal Philosophy: Selected Papers of Richard Montague*. Edited and with an Introduction by Richmond H. Thomason. New Haven [Yale UP] 1974
- MÜLLER, CORNELIA: Gesture and Language. In: *The Routledge Linguistics Encyclopedia*. London [Routledge] 2009, S. 214–217
- MÜLLER, CORNELIA, ALAN CIENKI: Words, Gestures, and Beyond: Forms of Multimodal Metaphor in the Use of Spoken Language. In: FORCEVILLE, CHARLES; EDUARDO URIOS-APARISI (Hrsg.): *Multimodal Metaphor*. Berlin [De Gruyter] 2009, S. 297–328
- MÜLLER, CORNELIA; ALAN CIENKI; ELLEN FRICKE; SILVIA LADEWIG; DAVID MCNEILL; SEDINHA TESSENDORF (Hrsg.): *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*. Volume 1. Berlin [De Gruyter] 2013
- MÜLLER, CORNELIA; ALAN CIENKI; ELLEN FRICKE; SILVIA LADEWIG; DAVID MCNEILL; JANA BRESSEM (Hrsg.): *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*. Volume 2. Berlin [De Gruyter] 2014
- MÜLLER, MARKUS: ›Halt's Maul Averell!‹ - Die Inszenierung multimodaler Interaktion im Comic. In: PIETRINI, DANIELA (Hrsg.): *Die Sprache(n) der Comics*. München [Martin Meidenbauer] 2012, S. 75-90
- NICHOLS, BILL: Style, grammar and the movies. In: *Film Quarterly*, 28(3), 1975, S. 33-49
- NÖTH, WINFRIED: Zur Komplementarität von Sprache und Bild aus semiotischer Sicht. In: *Mitteilungen des Deutschen Germanistenverbandes*, 51, 2004, S. 8-21
- NÖTH, WINFRIED: Verbal-visuelle Semiotik. In: KLUG, NINA-MARIA; HARTMUT STÖCKL (Hrsg.): *Handbuch Sprache im multimodalen Kontext*. Berlin [De Gruyter], S. 190-216
- NORRIS, SIGRID; CARMEN DANIELA MAIER (Hrsg.): *Interactions, Images and Texts. A Reader in Multimodality*. Boston [De Gruyter] 2014
- NORRIS, SIGRID (Hrsg.): *Multimodality*. London [Routledge] 2015
- PEIRCE, CHARLES SANDERS: *The Collected Papers of Charles Sanders Peirce*. Edited by Charles Harthshorne and Paul Weiss. Cambridge [Belknap] 1931-1958
- PERRIN, DANIEL: *Medienlinguistik*. Konstanz [UVK] 2011
- POSNER, ROLAND: Zur Systematik der Beschreibung verbaler und nonverbaler Kommunikation. In: BOSSHARDT, HANS-GEORG (Hrsg.): *Perspektiven auf Sprache: Interdisziplinäre Beiträge zum Gedenken an Hans Hermann*. Berlin [De Gruyter] 1986, S. 267-313
- ROYCE, TERRY D.; WENDY L. BOWCHER (Hrsg.): *New Directions in the Analysis of Multimodal Discourse*. Mahwah [Psychology Press] 2006
- RUSSEL, STUART J.; PETER NORVIG; F. KIRCHNER: *Künstliche Intelligenz: Ein moderner Ansatz*. München [Pearson] 2012

- RYAN, MARIE-LAURE: Introduction. In: RYAN, MARIE LAURE (Hrsg.): *Narrative across Media: The Languages of Storytelling*. Lincoln [U of Nebraska P] 2004, S. 1-40
- SACHS-HOMBACH, KLAUS: Prädikative und modale Bildtheorie. In: DIEKMANN-SHENKE, HANS-JOACHIM; MICHAEL KLEMM; HARTMUT STÖCKL (Hrsg.): *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. Berlin [Erich Schmidt Verlag] 2011, S. 97-120
- SAUSSURE, FERDINAND DE: *Course in General Linguistics*. Edited by Charles Bally and Albert Sechehaye. Translated by Wade Baskin. London [McGraw-Hill] 1959
- SCHNEIDER, JÜRGEN G.; HARTMUT STÖCKL (Hrsg.): *Medientheorien und Multimodalität. Ein TV-Werbespot --- Sieben methodische Beschreibungsansätze*. Köln [Herbert von Halem] 2011
- SEELEY, WILLIAM P.: Hearing How Smooth It Looks: Selective Attention and Crossmodal Perception in the Arts. In: *Essays in Philosophy*, 13(2), 2012, S. 498–517
- SEIZOV, OGNJAN; JANINA WILDFEUER: *New Studies in Multimodality: Conceptual and Methodological Elaborations*. London [Bloomsbury] 2017
- SPILLNER, BERND: Stilanalyse semiotisch komplexer Texte. Zum Verhältnis von sprachlicher und bildlicher Information in Werbeanzeigen. In: *Kodikas/Code*, 4/5, 1982, S. 91-106
- SPITZMÜLLER, JÜRGEN; INGO H. WARNKE: *Diskurslinguistik. Eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse*. Berlin [De Gruyter] 2011
- STAM, ROBERT: Film and Language: From Metz to Bakhtin. In: PALMER, ROBERT BARTON (Hrsg.): *The Cinematic Text: Methods and Approaches*. Georgia [Ams Press Inc] 1989, S. 277-302
- STÖCKL, HARTMUT: Inbetween Modes. Language and Image in Printed Media. In: VENTOLA, EIJA; CASSILY CHARLES; MARTIN KALTENBACHER (Hrsg.): *Perspectives on Multimodality*. Amsterdam [John Benjamins] 2004, S. 9-30
- STÖCKL, HARTMUT: Multimodalität – Semiotische und textlinguistische Grundlagen. In: KLUG, NINA-MARIA; HARTMUT STÖCKL (Hrsg.): *Handbuch Sprache im multimodalen Kontext*. Berlin [De Gruyter] 2016, S. 3-35
- THOMPSON, ROY; CHRISTOPHER BOWEN: *Grammar of the shot*. London 2009
- TITZMANN, MICHAEL: Theoretisch-methodologische Probleme einer Semiotik der Text-Bild-Relationen. In: HARMS, WOLFGANG (Hrsg.): *Text und Bild, Bild und Text: DFG-Symposium 1988*. Weimar [Metzler] 1990, S. 368-384
- TSENG, CHIAO-I; JOHN A. BATEMAN: Multimodal Narrative Construction in Christopher Nolan's *Memento*: A Description of Method. In: *Journal of Visual Communication*, 11(1), 2012, S. 91-119

- TSENG, CHIAO-I; JOCHEN LAUBROCK; JANA PFLAEGING: Character Developments in Comics and Graphic Novels: A Systematic Analytical Scheme. In: DUNST, ALEXANDER; JOCHEN LAUBROCK; JANINA WILDFEUER (Hrsg.): *Empirical Comics Research: Digital, Multimodal and Cognitive Methods*. London [Routledge] 2018, S. 154-175
- VAN LEEUWEN, THEO: *Introducing Social Semiotics*. London [Taylor & Francis] 2005
- VAN LEEUWEN, THEO: Multimodality and Multimodal Research. In: MARGOLIS, ERIC; LUC PAUWELS (Hrsg.): *The SAGE Handbook of Visual Research Methods*. London [SAGE Publishing] 2011, S. 549-569
- VAN LEEUWEN, THEO: About Images and Multimodality: A Personal Account. In: NORRIS, SIGRID; CARMEN DANIELA MAIER (Hrsg.): *Interactions, Images and Texts*. Boston [De Gruyter] 2014, S. 19-24
- VENTOLA, EIJA; CASSILY CHARLES; MARTIN KALTENBACHER (Hrsg.): *Perspectives on Multimodality*. Amsterdam [John Benjamins] 2004
- WILDFEUER, JANINA: Intersemiosis in Film. Towards a New Organisation of Semiotic Resources in Multimodal Filmic Text. In: *Multimodal Communication*, 1(3), 2012, S. 276-304
- WILDFEUER, JANINA: *Film Discourse Interpretation. Towards a New Paradigm for Multimodal Film Analysis*. London [Routledge], 2014a
- WILDFEUER, JANINA: Formale Zugänge zur Diskursanalyse. In: *Zeitschrift für Semiotik*, 35(3-4), 2014b, S. 393-417
- WILDFEUER, JANINA: Coherence in Film: Analysing the Logical Form of Multimodal Discourse. In: MAIORANI, ARIANNA; CHRISTINE CHRISTIE (Hrsg.): *Multimodal Epistemologies. Towards an Integrated Framework*. London 2014 [Routledge], S. 260-274
- WILDFEUER, JANINA (Hrsg.): *Building Bridges for Multimodal Research. International Perspectives on Theories and Practices of Multimodal Analysis*. Frankfurt/M. [Peter Lang] 2015
- WILDFEUER, JANINA: Diskurssemiotik = Diskurssemantik + multimodaler Text. In: KÄMPER, HEIDRUN; ERNEST W.B. HESS-LÜTTICH; MARTIN REISIGL; INGO H. WARNKE (Hrsg.): *Diskurs semiotisch. Sammelband zur vierten Netzwerktagung Diskurs interdisziplinär*. Berlin [De Gruyter] 2017a, S. 198-210
- WILDFEUER, JANINA: From Text to Performance: Discourse Analytical Thoughts on New Forms of Performances in Social Media. In: SINDONI, MARIA GRAZIA; JANINA WILDFEUER; KAY O'HALLORAN (Hrsg.): *Mapping Multimodal Performance Studies*. London [Routledge] 2017b, S. 180-200
- WILDFEUER, JANINA: Diskurslinguistik und Text. In: Warnke, Ingo H. (Hrsg.): *Handbuch Diskurs*. Berlin [De Gruyter] 2018
- WILDFEUER, JANINA; JOHN A. BATEMAN: Zwischen gutter und closure. Zur Interpretation der Leerstelle im Comic durch Inferenzen und dynamische Diskursinterpretation. In: *Closure. Kieler E-Journal für Comicforschung*, 1, 2014, S. 3-24

WILDFEUER, JANINA; MARTIN W. SCHNELL; CHRISTIAN SCHULZ: Talking about Dying and Death: On New Discursive Constructions of a Formerly Postulated Taboo. In: *Discourse & Society*, 26(3), 2015

WINKLER, HARTMUT: Zeichenmaschinen: oder warum die semiotische Dimension für eine Definition der Medien unerlässlich ist. In: MÜNKER, STEFAN; ALEXANDER ROESLER (Hrsg.): *Was ist ein Medium*. Frankfurt/M.[Suhrkamp] 2008, S. 211-222