

# Multimodalität

---

Wibke Weber

Zeitschriften sind multimodal. In ihnen verschmelzen Texte, Fotos, Infografiken, Datenvisualisierungen, Illustrationen und Design zu einem sinnvollen Kommunikationsangebot. Dabei wird die Botschaft eines Berichts, einer Reportage, nicht allein durch Text oder Bild vermittelt, sondern immer auch durch das multimodale Zusammenspiel der verschiedenen Elemente. Diese Elemente werden in der Multimodalitätsforschung *Modi* genannt. Der multimodale Ansatz ermöglicht es, Zeitschriften ganzheitlich zu analysieren unter Einbeziehung der verschiedenen Ebenen von Modi, Genre und Diskurs. Dieser Beitrag erläutert Multimodalität und relevante Schlüsselbegriffe, verortet Multimodalität im theoretischen Kontext der Sozialsemiotik und zeigt schließlich, wie sich der Multimodalitätsansatz für die Analyse von Zeitschriften operationalisieren lässt.

## 1. Sämtliche Kommunikation ist multimodal

Neue Medientechnologien führen zu neuen Ausdrucksformen. Vor allem die Digitalisierung und das Internet haben die Art und Weise, wie wir uns informieren und wie wir miteinander kommunizieren, enorm verändert. Ein Phänomen der digitalen Medienkultur ist Medienkonvergenz: das Annähern oder Verschmelzen von Einzelmedien oder Kanälen und deren Funktionalitäten, etwa wenn wir mit dem Smartphone telefonieren, im Internet browsen, Videos anschauen, Nachrichten verschicken und fotografieren. Die technologische Konvergenz wiederum hat Auswirkungen auf Medienformate, Genres, Erzähltechniken und Design und führt zu einer veränderten Medienästhetik. Eine digitale Zeitschrift, als App fürs Tablet oder Smartphone konzipiert, wie etwa das *Time Magazine*, kombiniert Text und Bild anders als eine gedruckte Zeitschrift. Digital beginnt eine Story bspw. mit einem großen Aufmacherbild und entfaltet sich dann, während man nach unten scrollt (Scrollytelling), in einem Wechsel aus Text, Foto und Infografik. Bilder, die in der Printausgabe statisch sind, wandeln sich in der digitalen Version zu Videos oder gar immersiven Bildgeschichten in Virtual Reality (VR). Statt die Texte zu lesen, kann man sie auch hören. Das Layout ist für den kleinen Bildschirm optimiert und individuell anpassbar. Gerade digitale Zeitschriften nutzen das Repertoire an semiotischen Modi in hohem Maße: verbal in Form von gesprochener und geschriebenen Sprache, vi-

suell als Einzelbild, Bewegtbild sowie durch Layout und Typografie, auditiv in Form von Sound und Musik, haptisch durch Touchscreen-Gesten wie *Scrollen* und *Wischen*. All diese Modi tragen zur Bedeutungskonstruktion bei.

Die Beschäftigung mit Multimodalität findet ihren Anfang in dem sog. *pictorial turn* (Mitchell 1992) oder *iconic turn* (Boehm 1994): einer verstärkten Beschäftigung mit visuellen Ausdrucksformen und wie sie Sinn erzeugen. Kress beschreibt diese Entwicklung als »move from telling the world to showing the world« (2003: 140). Da aber weder Bilder noch Texte allein vorkommen in der medienvermittelten Kommunikation, sondern immer in Kombination mit anderen Modi, sprechen einige Wissenschaftler:innen vom *multimodal turn* (Bucher 2011; Jewitt 2014: 4). Ein Leitgedanke in der Multimodalitätsforschung lautet daher: sämtliche Kommunikation ist multimodal (Bateman/Wildfeuer/Hiippala 2017; Bucher 2019; Kress/van Leeuwen 1996; Machin 2007; Stöckl 2016). Was genau Multimodalität bedeutet und was Modi sind, soll im folgenden Abschnitt erläutert werden.

## 1.1 Was ist Multimodalität?

Der Begriff Multimodalität wird je nach Kontext unterschiedlich definiert. In der Kognitionsforschung und Lernpsychologie wird unter Multimodalität die Aktivierung der verschiedenen Sinnesmodalitäten verstanden, mit denen Rezipient:innen ein Medienangebot wahrnehmen. Danach sind solche Medienangebote multimodal, die bei Nutzer:innen verschiedene Sinnesmodalitäten wie Sehen, Hören, Tasten ansprechen (Mayer 1997; Weidenmann 1997). Im Forschungskontext von Human-Computer Interaction erfährt der Begriff Multimodalität eine andere Bedeutung. Zwar bezieht man sich auch hier auf die Sinnesmodalitäten; als Modalitäten werden aber bestimmte Formen der Interaktion zwischen Mensch und Computer bezeichnet. Es wird dabei unterschieden zwischen den verschiedenen Eingabeoptionen seitens der Nutzer:innen (z.B. durch Sprache, durch motorische Steuerung von Touchscreen oder Körperbewegung) und den Ausgabeoptionen seitens des Gerätes (z.B. durch Sprache, Geräusche, Klänge und haptisches Feedback). Werden mehrere Modalitäten verwendet, spricht man von multimodaler Interaktion.

Eine wieder andere Deutung erfährt der Begriff Multimodalität in der Linguistik und Medienwissenschaft. Ausgehend von der Sozialsemiotik entwerfen Kress und van Leeuwen eine Theorie der multimodalen Kommunikation, die über das Zeichensystem Sprache hinausgeht (Kress/van Leeuwen 1996; Kress 2010) und weitere Modi einbezieht: »*Image, writing, layout, music, gesture, speech, moving image, soundtrack and 3D objects are examples of modes used in representation and communication*« (Kress 2010: 79; Hervorh. i. Orig.). Multimodale Texte sind also semiotische Einheiten, die aus dem sinnvollen Zusammenspiel verschiedener Modi resultieren und so eine bestimmte Botschaft vermitteln. In diesem Sinn werden auch Zeitschriften als multimodale Texte bzw. »multimodale Kommunikationsformen« betrachtet (Pfurtscheller 2017).

## 1.2 Was sind Modi?

Ein multimodaler Text ist bspw. eine Reportage, die aus geschriebener Sprache und Bildern besteht. Ein Vortrag ist multimodal, da er gesprochene Sprache und Gesten kombiniert. Zeitschriftenseiten zeigen sich als äußerst komplexe multimodale Texte, da in ihnen die Modi Bild, geschriebene Sprache und Design zu einem bedeutungsvollen Ganzen orchestriert werden (siehe Abb. 1a-d).

Abbildung 1: Zeitschriften als multimodale Texte: Beispielseiten aus dem Bereich Corporate Publishing. a) Impact 45/2019, S. 36-37; b) Impact 46/2019, S. 34-35.



c) Impact 47/2019, S. 22-23; d) Impact 50/2020, S. 24-25.



Das Hochschulmagazin *Impact-ZHAW* ist journalistisch geprägt. Es widmet jede Ausgabe einem bestimmten Thema; diese Themen-Dossiers folgen einer bestimmten Struktur und bestehen aus Infografiken und Datenvizualisierungen, Fotostrecken, Interviews, Berichten, Porträts über Forschende, Reportagen, Kommentaren und Meinungsbeiträgen.

Quelle: Hochschulmagazin Impact-ZHAW Zürcher Hochschule für Angewandte Wissenschaften.

Der Begriff Modus (*mode*) lässt sich am besten über seine semiotischen Ressourcen (*semiotic resources*) erklären. Jeder Modus basiert auf einer Reihe semiotischer Ressourcen, die sozial und kulturell geprägt sind. Modus ist also »a set of socially and culturally shaped resources for making meaning« (MODE 2012; siehe Kress 2010: 79; vgl. auch Jewitt/Henriksen 2016). Ein geschriebener Text zeichnet sich zum einen durch seine sprachlichen Ressourcen aus: Wörter und Sätze, die nach grammatischen Regeln linear und kohärent organisiert sind. Zum anderen kommen die semiotischen Ressourcen der Typografie zum Tragen: z.B. Schriftart, Schriftgröße, Schriftschnitt, Zeilenabstand oder Leerzeichen. Bei einem Foto sind die semiotischen Ressourcen Farbe, Bildschnitt, Perspektive oder Format; bei einer Datenvizualisierung sind es Linie, Punkt, Farbe, Form oder Größe. Verschiedene Modi können aus den gleichen semiotischen Ressourcen bestehen: gesprochene Sprache, der Soundtrack eines Films, ein Musikstück, eine Nachrichtentrommel – sie alle basieren auf dem Material von Klang, der sich wiederum über Parameter wie Klanghöhe, Klangtempo, Lautstärke, Modulation und Rhythmus charakterisieren lässt (siehe Kress 2010: 79ff).

Die semiotischen Ressourcen – und damit auch die Modi – verfügen über spezifische Potenziale, Limitationen und innere Logiken, um etwas auszudrücken, was in der Sozialsemiotik mit *Affordance* (Angebotscharakter) bezeichnet wird. Sprache eignet sich, um Sachverhalte zu beschreiben oder um Aussagen zu einer Argumentationskette zu verknüpfen. Sprache ist also dort gut, wo logisches Denken und Kausalität gefordert sind. Bilder dagegen sind nicht-linear gestaltet. Sie können Dinge auf einen Blick sichtbar machen und wirken wie »schnelle Schüsse ins Gehirn« (Kroeber-Riel 1996: 53); zudem können sie eine starke emotionale Wirkung entfalten. Bilder vermögen es, die Oberfläche eines Gegenstandes viel detaillierter wiederzugeben als eine Reihe von Wörtern. Ein Bild kann aber nicht sagen, wie viel der abgebildete Gegenstand wiegt oder wem er gehört; dazu bedarf es wiederum der Sprache. Abstrakte Bilder, z.B. Datenvizualisierungen, eignen sich, um Daten, Strukturen oder Muster anschaulich darzustellen. Design wiederum hat eine integrierende Funktion und organisiert die einzelnen Modi zu einem sinnvollen Ganzen, akzentuiert Inhalte und unterstützt die Leser:innenführung. Ist man sich der Potenziale und Limitationen der Modi bewusst, kann man die verschiedenen Modi gezielt für das jeweilige Kommunikationsziel einsetzen.

Zusammenfassend lässt sich sagen, dass semiotische Ressourcen über bestimmte Bedeutungspotenziale verfügen, wobei immer auch der jeweilige soziale Nutzungskontext, die Kultur und der historische Kontext zur Bedeutungskonstruktion beitragen. Wir gehen heute anders mit Fotos um als unsere Vorfahr:innen im späten 19. Jahrhundert. Die Porträtfotos des 19. Jahrhunderts spiegeln eine Ästhetik wider, die noch an die Porträtmalerei angelehnt ist: eine Person oder eine Gruppe von Personen, die vor einem schlichten Hintergrund in einem Atelier sitzt. Zu dieser Zeit war der Gang in ein Fotostudio die einzige Möglichkeit für ein Porträtfoto. Heute können wir mit Smartphones überall und in jeder Situation fotografieren, machen Selfies und bearbeiten sie, um ein möglichst ästhetisches Resultat zu erzielen. Wir publizieren unsere Fotos auf Social Media und teilen sie mit Familie und Freundeskreis, der Community oder der ganzen Welt. Bilder übernehmen damit eine kommunikative Funktion.

## 2. Von der Semiotik zur Sozialsemiotik

Den theoretischen Hintergrund zur Multimodalität bildet die Sozialsemiotik (*social semiotics*). Die Sozialsemiotik hat zum einen ihre Wurzeln in der Semiotik; zum anderen nimmt sie Bezug auf die systemisch-funktionale Linguistik. Semiotik ist die Lehre von den Zeichen. Sie untersucht die Beziehung zwischen den Zeichen, ihren Bedeutungen und den Dingen in der Welt. Der Semiotiker Charles Sanders Peirce entwickelte einen umfassenden Zeichenbegriff (Peirce 1931-1958). Er klassifizierte Zeichen danach, in welcher Beziehung ein Zeichen zu dem steht, was es bezeichnet. Bei einem *ikonischen* Zeichen besteht eine gewisse Ähnlichkeit zwischen dem Zeichen und dem, wofür es steht (Beispiel: Foto). Bei einem *symbolischen* Zeichen dagegen basiert diese Beziehung auf Konventionen, d.h. sie ist arbiträr (Beispiel: Sprache). Bei einem *indexikalischen* Zeichen besteht ein kausaler Zusammenhang zwischen dem Zeichen und dem Bezeichnetem (Beispiel: Rauch als Zeichen für Feuer). Zudem betrachtet Peirce die Zeichen als etwas Dynamisches und führte den Begriff der Semiose ein. Damit meint er den unendlichen Prozess des Zeichengebrauchs, »in dem Zeichen durch Verweise auf andere Zeichen mit Zeichen verbunden sind und aus Zeichen (Repräsentamen) neue Zeichen (Interpretanten) entstehen« (Nöth 2000: 227).

Genau hier setzt die Sozialsemiotik an. Während sich die Semiotik mit den Zeichen beschäftigt, rückt die Sozialsemiotik den Zeichengebrauch und damit die Zeichenbenutzer:innen (*sign-makers*) in den Mittelpunkt. Es geht zum einen darum, wie Menschen in bestimmten sozialen Situationen mit verschiedenen Mitteln kommunizieren, um bestimmte Ziele zu erreichen, z.B. politische, wirtschaftliche oder ideologische Ziele; zum anderen geht es darum, wie Bedeutungen – und damit auch Machtverhältnisse – zwischen verschiedenen Akteur:innen ausgehandelt werden. Die Sozialsemiotik geht zurück auf die systemisch-funktionalen Linguistik, die von M.A.K. Halliday (1978, 1985) entwickelt wurde. Er beschrieb Sprache als ein System aus Bedeutungspotenzialen (*meaning potentials*) und war vor allem an den Funktionen von Sprache interessiert (*meta-functions*). Kress und van Leeuwen (1996) haben Hallidays Metafunktionen in ihren sozialsemiotischen Ansatz zur Multimodalität integriert und auf andere semiotische Modi übertragen, vor allem auf Bilder und visuelles Design.

### 2.1 Die Metafunktionen

Eine zentrale Idee in der Sozialsemiotik ist, dass jede bedeutungstragende Einheit, also jeder Modus, drei Metafunktionen erfüllt. Die erste Metafunktion (*representational/ideational*) fragt danach, was repräsentiert wird; hier geht es um die Darstellung von Personen, kulturellen Erfahrungen, Ereignissen, Handlungen, Objekten oder einem Zustand. Die zweite Metafunktion (*interpersonal/interactional*) adressiert die soziale Beziehung zwischen Sender:innen und Rezipient:innen, d.h. die Rollen, die die Kommunizierenden einnehmen, z.B. die Bildproduzent:innen, die Grafiker:innen, die Journalist:innen, die Leser:innen, die Betrachter:innen eines Bildes. Die dritte Metafunktion (*compositional/textual<sup>1</sup>*) betrifft die Komposition des multimodalen Textes zu einer sinn-

<sup>1</sup> Textual im Sinne von *weber* (*textere*).

vollen Gesamteinheit. Diese drei Metafunktionen greifen im Kommunikationsprozess ineinander. Im Folgenden werden die drei Metafunktionen näher beschrieben, um sie als Analyseinstrument besser greifbar zu machen (siehe Engebretsen/Weber 2022).

### 2.1.1 Die repräsentative Metafunktion

Mögliche Analysefragen, die die repräsentative (*ideational*) Metafunktion betreffen, sind:

- i. Welche Aussagen werden über die Welt gemacht?
- ii. Welche Personen, Akteur:innen werden dargestellt?
- iii. Welche Themen, Motive, Ereignisse, Situationen werden abgebildet?
- iv. Welche Aktionen werden beschrieben, wer ist wie daran beteiligt?
- v. Welche Begriffe, Konzepte (auch metaphorische), Denkweisen liegen hinter den getroffenen Aussagen?
- vi. Wessen Welt oder Weltsicht wird abgebildet?

Handelt es sich beim visuellen Modus um ein Foto, lässt sich z.B. fragen: Wer oder was wird visuell dargestellt? Was können wir über die Eigenschaften (Attribute) der Akteur:innen sagen, über Gesten, Körperhaltung und Gesichtsausdruck? Wie werden die Beteiligten kategorisiert, z.B. als Individuen, Gruppen oder generische Typen? Wichtig ist auch zu beschreiben, was fehlt, also eine Art negative Analyse durchzuführen: Was wird nicht gezeigt oder wurde absichtlich weggelassen, indem ein bestimmter Bildausschnitt gewählt wurde? Bezuglich des verbalen Modus (geschriebener Text) lässt sich untersuchen, wer oder was im Text erwähnt wird, z.B. ob die Akteur:innen, Ereignisse, Dinge mit Namen bezeichnet werden oder ob durch die Verwendung des Passivs absichtlich das Subjekt/die Akteur:innen nicht genannt werden. Kriterien sind auch hier wieder soziale Kategorien, Metaphern, Attribute oder Adjektive. Welche Konnotationen schwingen bei den jeweiligen Begriffen mit? Mit der repräsentativen Funktion wird auch deutlich, ob sich die Modi (z.B. Bild und Bildunterschrift) in ihren Aussagen gegenseitig stützen, ergänzen oder sich widersprechen.

### 2.1.2 Die interpersonale Metafunktion

Die interpersonale Metafunktion legt den Fokus auf die sozialen Beziehungen zwischen Sender:innen/Produktionsteam/Kommunikator:innen und der Zielgruppe. Interpersonal meint, wie die Zielgruppe adressiert wird, inwieweit sie involviert wird und welche Macht- und Kontrollverhältnisse im multimodalen Text etabliert werden. Fragen, die diese Aspekte aufgreifen, sind bspw.:

- i. Welche Identitäten und Rollen werden in Text und/oder Bild hergestellt?
- ii. Wie treten Sender:innen in Erscheinung, wie werden ihre kommunikativen Absichten deutlich?
- iii. Wie wird die Beziehung zwischen Sender:in und Rezipient:in dargestellt?
- iv. Welche Interaktionen zwischen Sender:in und Rezipient:in gibt es?

- v. Welche Beziehung wird zwischen den Rezipient:innen und den im Text oder Bild repräsentierten Personen erzeugt?
- vi. Wessen Agency (soziales Handeln) lässt sich im multimodalen Text erkennen?

Auch wenn die Sender:innen nicht direkt in Erscheinung treten, lassen sich über sie Aussagen machen: im verbalen Modus etwa aufgrund der Erzählperspektive, der verwendeten rhetorischen Stilmittel, des Sprachstils (z.B. werden Fachbegriffe erklärt?) und des Kommunikationszwecks (z.B. erzählen, beschreiben, erklären, argumentieren). Weitere Hinweise sind die Einbeziehung der Zielgruppe durch direkte Ansprache (z.B. die Verwendung von Personalpronomen Sie, du, ihr), die Haltung, in der Inhalte präsentiert werden (objektiver vs. subjektiver Sprachstil) oder das Offenlegen des methodischen Vorgehens – Stichwort Transparenz, z.B. woher die Daten zum Klimawandel stammen oder wie die regionalen Daten zur Pandemie erhoben wurden. Transparenz wiederum sorgt für Glaubwürdigkeit. Auch Zitate von Expert:innen, Statistiken, Aussagen von Augenzeug:innen oder überprüfbare Quellenangaben unterstreichen die Glaubwürdigkeit eines Textes – ein wichtiges Kriterium in Zeiten von fake news, alternative facts und gefühlter Wahrheit.

Bei Fotos (visueller Modus) lässt sich das interpersonale Bedeutungspotenzial an Kriterien festmachen wie Kameraeinstellung, Perspektive, Positionierung, Bildausschnitt oder Blickrichtung. Eine abgebildete Person kann aus der Nähe oder der Ferne abgebildet werden (Nahaufnahme, Totale), perspektivisch auf Augenhöhe mit den Betrachtenden stehen oder von oben auf sie herabschauen (Vogelperspektive). Letzteres hat den Effekt, dass den Betrachtenden eine unterlegene und damit schwächere Position zugewiesen wird, während der abgebildeten Person oder dem Objekt Überlegenheit, Stärke und damit Macht zugeschrieben wird. Eine Nahaufnahme vermittelt ein Gefühl der Intimität, während eine Totale soziale Distanz schafft. Eine Person auf einem Foto kann entweder die Betrachtenden direkt anschauen und durch den direkten Blickkontakt etwas fordern; oder der Blick der abgebildeten Person ist abgewendet und nicht an die Betrachtenden gerichtet. Kress und van Leeuwen nennen diese Arten von Blickkontakt »demand« and »offer« (1996: 121-130).

### 2.1.3 Die kompositionelle Metafunktion

In einem dritten Schritt wird das Zusammenspiel der Modi untersucht. Es geht um die Gesamtkomposition: »the meaning of composition« (Kress/van Leeuwen 1996: 181-229). Nun rücken Layout und Design-Elemente in den Mittelpunkt, die idealerweise die einzelnen Einheiten zu einem kohärenten, sinnvollen Ganzen orchestrieren (siehe Voelker im Band). Zentrale Fragen sind hier:

- i. Wie wird ein Zusammenhang hergestellt zwischen den Modi, z.B. durch Farben, Linien, Positionierung?
- ii. Welche Dynamiken lassen sich erkennen, z.B. durch Linienführung oder Pfeile?
- iii. Welche Spannungen im multimodalen Text lassen sich ausmachen? Passt der Layoutstil zur Aussage des Textes?
- iv. Worauf wird die Aufmerksamkeit der Leser:innen gelenkt?

Wichtige Begriffe im Zusammenhang mit der kompositionellen Metafunktion sind *information value*, *salience* und *framing* (Kress/van Leeuwen 1996: 186–218).

- i. *Information value* bezieht sich auf die Position, die einem Element im Layout einer Seite zugeteilt wird (Kress/van Leeuwen 1996: 186–211). Mit dieser Position sind eine bestimmte Bedeutung und Wertigkeit verknüpft. Aufgrund von Lesegewohnheiten schreiben wir Elementen, die oben auf einer Seite stehen, einen höheren Informationswert zu als Elementen, die unten am Rand platziert sind. Kress und van Leeuwen nennen drei Kategorien von »information value«: »given–new« (links–rechts), »ideal–real« (oben–unten) und »center–margin« (Mitte–Rand) (1996: 208). Es bleibt jedoch fraglich, ob diese mit der Position verbundene Informationswertigkeit im modernen und digitalen Zeitschriftendesign noch ihre Gültigkeit hat. Vor allem das Smartphone mit seinem kleinen Bildschirm erfordert ein spezielles Web- und Interfacedesign, das mit traditionellen Layoutkonzepten und Lesarten bricht.
- ii. *Salience* betrifft die Art und Weise, wie die einzelnen Modi ausgestaltet sind, um die Aufmerksamkeit der Leser:innen auf sich zu ziehen. Welches Element fällt besonders auf durch Farbe, Größe, Position oder Anordnung? Tritt ein Modus dominanter auf als die anderen?
- iii. *Framing* beschreibt das Prinzip von Kohärenz und Abgrenzung, nämlich wie die einzelnen Elemente miteinander verbunden oder voneinander getrennt sind, so dass man sie als Einheiten wahrnimmt. Semiotische Ressourcen sind hier bspw. Linien, Weißraum, Farben, Kontraste oder Ausrichtung. *Framing* schafft Orientierung in einer Zeitschrift und schlägt Leserichtungen vor. *Framing* knüpft an Erkenntnisse aus der Gestaltpsychologie und den so genannten Gestaltgesetzen an. Die Gestaltgesetze sind eine Reihe von Organisationsprinzipien, die beschreiben, wie Menschen visuelle Elemente wahrnehmen, gruppieren und zueinander in Beziehung setzen. Bekannte Gestaltgesetze sind das Gesetz der Nähe, Ähnlichkeit, Kontinuität, Symmetrie, Geschlossenheit, Figur-Grund (siehe Ware 2000: 203–213).

## 2.2 Intermodale Relationen

Intermodale Relationen (*information linking*) sind Beziehungen zwischen den verschiedenen Modi, vor allem zwischen Text und Bild (siehe Ruchatz im Band). Van Leeuwen (2005: 219–227) unterscheidet zwei Grundmuster von Text-Bild-Relationen: Elaboration und Extension. Diese Grundmuster ähneln Roland Barthes' Konzept (1977) von *ancrage* (*anchorage*) und *relais* (*relay*). Während bei einer Elaboration der Text das Bild erklärt oder umgekehrt, bezeichnet die Extension die Erweiterung des Textinhaltes durch das Bild oder umgekehrt. Wenn eine Bildunterschrift mehr oder weniger das sagt, was man auf dem Foto sieht, handelt es sich um eine Elaboration. Als Elaboration gelten auch Spezifizierung, Exemplifizierung und Zusammenfassung. Liefert aber die Bildunterschrift zusätzliche Informationen, z.B. wann etwas passiert ist, wo, wie oder warum, dann liegt eine Extension vor. Extension kann in Form einer zeitlichen, räumlichen oder logischen Erweiterung (Begründung, Ursache–Wirkung, Wenn-Dann-Beziehung) auftreten oder auch als rhetorisches Stilmittel wie Pars pro Toto, Analogie, Kontrast, Humor oder Ironie (vgl. auch Stöckl 2011). So kann der Text die Ursache benennen und das Bild die Wirkung

zeigen. Text-Bild-Relationen, die dem Muster der Elaboration folgen, können zu einer besseren Verständlichkeit führen; sie können aber auch redundant oder gar langweilig wirken. Text-Bild-Relationen, die dem Muster der Extension folgen, können interessant und attraktiv wirken, weil sie spannungsvoll gestaltet sind; sie können aber auch irritieren und Verständnisfragen aufwerfen. Das Analysekriterium *information linking* fragt also nach den wechselseitigen semantischen Bezügen und danach, ob ein Modus den anderen dominiert.

### 3. Multimodale Analyse

Wie sich die Metafunktionen mit den entsprechenden Analysekriterien methodisch umsetzen lassen, soll im Folgenden skizziert werden.<sup>2</sup> Multimodale Analyse bedeutet immer Produktanalyse. Sie eignet sich für die Analyse von analogen wie digitalen Medienprodukten. Sie eignet sich nicht, um Produktions- und Rezeptionsprozesse zu untersuchen. Die vorgeschlagene Methode ist in drei Phasen organisiert: Zuerst beschreibt man die vorhandenen Modi und was spontan auffällt. In der zweiten Phase werden die verschiedenen Modi und ihre Bedeutungspotenziale analysiert. In der dritten Phase geht es um die semiotischen und sozialen Potenziale des multimodalen Textes in Bezug auf Genre und Diskurs.

#### 3.1 Bestandsaufnahme

Die erste Phase ist deskriptiver Natur und soll dazu dienen, die grundlegenden Merkmale des untersuchten multimodalen Textes bewusster zu erfassen: Welche semiotischen Modi sind am multimodalen Text beteiligt? Was fällt auf bezüglich Anordnung der Elemente, Layout oder Sprachstil? Wird Farbe verwendet, um bestimmte Elemente hervorzuheben? Danach kategorisieren wir die Elemente. Typische Kategorien sind bspw.:

- i. Verbaler Modus: Titel, Lead, Bildunterschrift, Faktenbox
- ii. Visueller Modus (Bild): Porträtfoto, generisches Foto (Symbolbild), Datenvisualisierung (Karte, Balkendiagramm etc.)
- iii. Visueller Modus (Design): Layout-Aspekte, z.B. Anordnung von Text und Bild, Überschriftenhierarchie, Farbkonzept, Weißraum

Für eine solche Bestandsaufnahme spricht, dass sie den Blick schärft für die verschiedenen Elemente und Modi und deren Rolle im multimodalen Text. Wenn eine Doppelseite in einer Zeitschrift mit einem dominanten Titel überschrieben ist, z.B. »Mr. Robotic – Kollege oder Konkurrent« (siehe Abb. 1b), wird uns signalisiert, dass alle weiteren Elemente (Fotos, Text, Infobox) entsprechend diesem Titel interpretiert werden sollen. In der ersten Phase halten wir auch fest, was wir über das mediale Umfeld sagen können. Handelt es sich um ein interaktives Produkt oder ist es statisch? Gibt es Interaktionsmöglichkeiten, z.B. über eine Kommentarfunktion? Hat die Zeitschriftenausgabe ein

<sup>2</sup> Eine ausführliche Beschreibung mit Fallbeispielen findet sich bei Engebretsen und Weber (2022).

übergeordnetes Thema, z.B. »Künstliche Intelligenz« (siehe Abb. 1c)? Folgt das Layout einem vorgegebenen Gestaltungsraster? Was können wir über das Medienunternehmen oder den:die Herausgeber:in sagen, über Zielgruppe(n), Leitbild, Geschichte, Auflage? Welche Art von Zeitschrift liegt vor? Ist die Publikation eine wissenschaftliche Fachzeitschrift, ist sie aus dem Bereich Corporate Publishing oder Journalismus? Gilt die Zeitschrift als journalistisches Leitmedium?

### 3.2 Bedeutungspotenziale analysieren

In der zweiten Phase der Analyse wollen wir die Bedeutungspotenziale des multimodalen Textes aufdecken. Hierfür werden nun die Metafunktionen mit den oben genannten Analysekriterien herangezogen. Ziel ist, am Ende eine Aussage zu treffen über (i) das, was repräsentiert wird, (ii) die im multimodalen Text etablierten sozialen Beziehungen zwischen Sender:innen und Rezipient:innen und (iii) das Zusammenspiel der Modi, das die Gesamtaussage des multimodalen Textes prägt.<sup>3</sup>

### 3.3 Genre und Diskurs

In der dritten Phase der Analyse wird der multimodale Text im Kontext von Genre und Diskurs verortet. Genre bedeutet im traditionellen Sinn Texttyp oder Textsorte. Genres zeichnen sich durch gemeinsame spezifische Merkmale und Muster bezüglich Inhalt, Form oder Stil aus. In der Sozialsemiotik und Multimodalitätstheorie sind Genres konventionelle Muster multimodaler Kombinationen. Genre wird hier eng verknüpft mit den Modi und deren intermodalen Zusammenspiel, aber auch mit den verfügbaren Werkzeugen, dem gewählten Medium, dem sozialen Kontext, dem kommunikativen Zweck und der Zielgruppe (Bateman 2008). Diskurs bezieht sich auf die Art und Weise, wie wir über die Welt denken und sprechen. Indem wir Sprache und andere semiotische Modi verwenden, konstruieren wir Bedeutungen und kommunizieren so eine bestimmte Sicht auf die Welt. Diskurse manifestieren sich in multimodalen Texten und repräsentieren das Wissen und Denken einer bestimmten Zeit, Kultur oder sozialen Gruppe.

Eng mit dem Diskurs verbunden ist der Begriff der sozialen Praxis. Soziale Praktiken sind Handlungsweisen, Handlungsmuster und praktisches Wissen, die verbunden sind mit einem kommunikativen Zweck. Schreiben oder der Umgang mit Bildern (fotografieren, bearbeiten, posten, sharen) sind Beispiele für soziale Praktiken (Pfurtscheller 2019). Soziale Praktiken sind immer auch geprägt von professionellen Normen und Berufsethiken, von Gestaltungskonventionen und Designtrends wie etwa Flat Design oder der Einsatz von einfachen Illustrationen (siehe Abb. 1d). Mit folgenden beispielhaften Fragen lässt sich die Kontextebene von Genre und Diskurs untersuchen:

- i. Welchem Genre gehört der multimodale Text an? Handelt es sich z.B. um ein journalistisches Genre (Bericht, Reportage, Kommentar) oder um ein Genre aus dem Be-

---

<sup>3</sup> Ein Beispiel für eine multimodale Analyse aus dem Zeitungsdesign geben Weber und Rall in dem Band *Periodical Studies Today: Multidisciplinary Analyses* (Ernst/von Hoff/Scheiding 2022).

- reich der Organisationskommunikation (Medienmitteilung, Geschäftsbericht, Advertorial)?
- ii. Wie verhält sich der multimodale Text zu den Erwartungen des Genres? Ist er konventionell oder innovativ gestaltet? Stellt er eine Variation des Genres dar?
  - iii. Inwiefern beeinflussen technologische oder ökonomische Bedingungen die sozialen Praktiken des Schreibens, des Visualisierens, des Gestaltens?
  - iv. Nimmt der multimodale Text implizit Bezug auf andere Texte oder Medienprodukte im Diskurs?
  - v. Wie wird der multimodale Text kommentiert, geliked, geshared? Wie sehen die Reaktionen darauf aus, z.B. auf Social Media?
  - vi. Wer sind die Auftraggeber:innen oder die Beteiligten im Produktionsprozess, wer ist die Zielgruppe? Welche anderen Akteur:innen sind zudem involviert (z.B. politische Interessensgruppen)?
  - vii. Wessen Weltsicht transportiert der multimodale Text, wessen Interesse steckt dahinter?

Phase III ist sowohl durch Analyse und Interpretation wie auch Diskussion und Reflexion gekennzeichnet. Die vorgeschlagene Methode für die multimodale Analyse ist als Werkzeug und Orientierungshilfe gedacht für den Umgang mit der semiotischen und sozialen Komplexität von multimodalen Produkten. Es ist jedoch ebenso wichtig, Intuition, Erfahrung und kritisches Denken einzusetzen, um zu verstehen, was an dem untersuchten Produkt wirklich interessant und von gesellschaftlicher Bedeutung ist.

#### 4. Kritische Schlussbetrachtung

Obwohl Modus (*mode*) in der Multimodalitätsforschung ein zentraler Begriff ist, wird er kontrovers diskutiert und lässt Fragen offen (siehe z.B. Bateman/Wildfeuer/Hippala 2017: 112-123; vgl. Stöckl 2016). Diskutiert wird bspw., wo genau die Grenzlinie zu ziehen ist zwischen Modus und Medium: Ist ein Video ein Modus oder ist es ein Medium, das sich aus verschiedenen Modi (verbal, visuell, auditiv) speist? Wird das Bedeutungspotenzial eines Modus, einer semiotischen Ressource, durch kulturelle Konventionen bestimmt oder vor allem durch die konkrete Nutzungssituation? Musik vermittelt Bedeutung nicht in gleicher Weise wie es Sprache und Bild vermögen. Ist dann Musik immer noch ein Modus? Und kann Farbe auch die Position eines Modus einnehmen oder bleibt Farbe eine semiotische Ressource? Zumindest letztere Frage könnte man so beantworten: In einem Zeitschriftenartikel ist Farbe meist eine semiotische Ressource (schwarzer Text auf weißem Hintergrund). Im Kontext von Werbung oder Corporate Design kann Farbe aber durchaus zu einem eigenen Modus werden; die semiotischen Ressourcen wären dann bspw. Farbton, Helligkeit und Sättigung (vgl. auch Kress/van Leeuwen 2002; van Leeuwen 2010).

Ähnlich wie *mode* werden auch die *metafunctions* kritisch diskutiert (z.B. Bucher 2019; Ledin/Machin 2018a; Bateman et al., 2017). Ein Kritikpunkt ist, inwiefern sich die Mefatfunktionen, die aus der Linguistik abgeleitet sind, als generelles Analyseinstrument für alle Modi eignen, da die Modi in ihren Materialitäten und Affordanzen doch sehr

unterschiedlich sind. Wie lässt sich etwa der Inhalt von Musik oder Farbe beschreiben (*representational metafunction*)? Ist das, was man in einem Bild als *interpersonal meaning* analysiert nicht immer auch eine repräsentative Darstellung (*representational meaning*) zwischenmenschlicher Beziehungen? Diskutiert wird auch, ob ein Modus erst dann als Modus gilt, wenn alle drei Metafunktionen erfüllt sind. Van Leeuwen relativiert dieses Prinzip des gleichzeitigen Vorhandenseins der drei Metafunktionen:

You could say that in multimodal communication we always need the three metafunctions, so that all three are present in any act of multimodal communication, but which metafunction is mostly or solely carried by which kind of mode in the mix may differ. And when looking at modes separately, you may find that some develop the ideational metafunction more than others, and others the interpersonal. Multimodality requires the metafunctions to be rethought and not taken for granted. (Andersen et al. 2015: 106-107; zit.n. Andersen 2016)

Trotz dieser kritischen Erörterungen bieten sich die Metafunktionen als Startpunkt an für eine multimodale Analyse. Allerdings darf man nicht der Versuchung erliegen, die Metafunktionen ausschließlich interpretativ anzugehen. Vielmehr braucht es Analysekriterien, die den drei Metafunktionen zugeordnet sind (wie in diesem Beitrag skizziert) und mit denen sich die damit verbundenen Forschungsfragen untersuchen lassen. Ziel eines solchen systematischen Analyserasters ist es, die multimodale Analyse intersubjektiv nachvollziehbar zu machen. Zudem sollte eine multimodale Analyse über die Ebene der Modi hinausgehen und die Kontexte von Genre und Diskurs berücksichtigen. Einige Vertreter:innen in der Multimodalitätsforschung verbinden ihre Arbeiten mit der kritischen Diskursanalyse (z.B. Ledin/Machin 2018b). Ihr Interesse liegt darin, anhand einer multimodalen kritischen Diskursanalyse multimodale Texte nicht nur als Medienprodukte zu untersuchen, sondern sozialkritisch zu betrachten, um so ideologische Muster aufzudecken in Diskursen über politische und gesellschaftliche Themen.

Zusammenfassend lässt sich sagen: Multimodalität gekoppelt mit Sozialsemiotik kann für die Zeitschriftenforschung äußerst bereichernd sein, denn: (i) Multimodalität befasst sich eben nicht nur mit Text und Bild, sondern auch mit Design, Audios und Videos. Zudem berücksichtigt der multimodale Ansatz die medienadäquate Realisation von Inhalten, was gerade für die Analyse digitaler Zeitschriften wertvoll ist. (ii) Multimodalität untersucht das Zusammenspiel der verschiedenen Modi und wird so auch hybriden Medienformaten gerecht, die sich im Zuge von Medienkonvergenz herausbilden. (iii) Die hier vorgestellte multimodale Analyse geht über eine reine Medienproduktanalyse hinaus, indem sie die Text-/Produktebene mit der Diskusebene verbindet und den multimodalen Text im historischen oder aktuellen Diskurs verankert.

## Literatur

Andersen, Thomas Hæstbæk (2016): Metafunctions. In: Nørgaard, Nina (Hg.): Key Terms in Multimodality: Definitions, Issues, Discussions; <https://www.sdu.dk/multimoda/lkeyterms> (November 2020).

- Andersen, Thomas Hæstbæk/Boeriis, Morten/Maagerø, Eva/Seip Tønnesen, Elise (2015): *Social Semiotics. Key Figures, New Directions*. London: Routledge.
- Barthes, Roland (1977 [1964]): *Rhetoric of the Image*. In: Heath, Steven (Hg.): *Image, Music, Text*. London: Fontana Press. S. 32-51.
- Bateman, John A. (2008): *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. Basingstoke: Palgrave Macmillan.
- Bateman, John A./Wildfeuer, Janina/Hiippala, Tuomo (2017): *Multimodality. Foundations, Research and Analysis. A Problem-Oriented Introduction*. Berlin/Boston: de Gruyter.
- Boehm, Gottfried (1994): *Die Wiederkehr der Bilder*. In: Boehm, Gottfried (Hg.): *Was ist ein Bild?* München: Fink. S. 11-38.
- Bucher, Hans Jürgen (2011): *Multimodales Verstehen oder Rezeption als Interaktion: Theoretische und empirische Grundlagen einer systematischen Analyse der Multimodalität*. In: Diekmannshenke, Hajo/Klemm, Michael/Stöckl, Hartmut (Hg.): *Bildlinguistik: Theorien – Methoden – Fallbeispiele*. Berlin: Erich Schmidt. S. 123-156.
- Bucher, Hans Jürgen (2019): *Multimodalität als Herausforderung für die Visuelle Kommunikationsforschung*. In: Lobinger, Katharina (Hg.): *Handbuch visuelle Kommunikationsforschung*. Wiesbaden: Springer. S. 651-677.
- Engebretsen, M. & Weber, W. (2022): *Multimodal Analysis*. In: Gnach, Aleksandra/Weber, Wibke/Engebresten, Martin/Perrin, Daniel (Hg.): *Digital Communication and Media Linguistics*. Cambridge: Cambridge University Press. S. 174-179.
- Ernst, Jutta/von Hoff, Dagmar/Scheiding, Oliver (Hg.) (2022): *Periodical Studies Today: Multidisciplinary Analyses*. Bd. 1. *Studies in Periodical Cultures*. Leiden/Boston: Brill.
- Halliday, Michael Alexander Kirkwood (1978): *Language as Social Semiotic. The Social Interpretation of Language and Meaning*. London: Edward Arnold.
- Halliday, Michael Alexander Kirkwood (1985): *An Introduction to Functional Grammar*. London/New York: Arnold.
- Jewitt, Carey/Henriksen, Berit (2016): *Social Semiotic Multimodality*. In: Klug, Nina-Maria/Stöckl, Hartmut (Hg.): *Handbuch Sprache im multimodalen Kontext*. Berlin/Boston: de Gruyter. S. 145-164.
- Jewitt, Carey. (Hg.) (2014): *The Routledge Handbook of Multimodal Analysis*. 2nd edition. London: Routledge.
- Kress, Gunther & van Leeuwen Theo (2002): *Colour as a Semiotic Mode: Notes for a Grammar of Colour*. In: *Visual Communication October 1.3. 2002*. S. 343-368.
- Kress, Gunther (2003): *Literacy in the New Media Age*. London: Routledge.
- Kress, Gunther (2010): *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London: Routledge.
- Kress, Gunther/van Leeuwen, Theo (1996): *Reading Images: The Grammar of Visual Design*. London: Routledge.
- Kroeber-Riel, Werner (1996): *Bildkommunikation*. München: Vahlen.
- Ledin, Per/Machin, David (2018a): *Doing Critical Discourse Studies with Multimodality: From Metafunctions to Materiality*. In: *Critical Discourse Studies* 29. April. S. 1-17.

- Ledin, Per/Machin, David (2018b): Multimodal Critical Discourse Analysis. In: Flowerdew, John/Richardson, John E. (Hg.): *Routledge Handbook of Critical Discourse Studies*. London: Routledge. S. 60-76.
- Ledin, Per/Machin, David (2007): *Introduction to Multimodal Analysis*. London: Hodder Arnold.
- Mayer, Richard E. (1997): Multimedia Learning: Are We Asking the Right Questions? In: *Educational Psychologist* 32.1. S. 1-19.
- Mitchell, William John Thomas (1992): The Pictorial Turn. In: *Artforum*, March. S. 89-94.
- MODE (2012): Glossary of multimodal terms; <https://multimodalityglossary.wordpress.com/> (3. November 2020)
- Nöth-Buch, Winfried (2000): *Handbuch der Semiotik*. Stuttgart: Metzler.
- Peirce, Charles Sanders (1931-1958). *Collected Papers of Charles Sanders Peirce*. Hartshorne, Charles/Weiss, Paul (Hg.) (1974-1979). Cambridge, MA: Belknap Press.
- Pfurtscheller, Daniel (2017): *Visuelle Zeitschriftenforschung. Nachrichtenmagazine als multimodale Kommunikationsformen*. Innsbruck: Innsbruck University Press.
- Pfurtscheller, Daniel (2019): *Bilder zwischen Zeichen, Handlungen und Praktiken*. In: Lübinger, Katharina (Hg.): *Handbuch Visuelle Kommunikationsforschung*. Wiesbaden: Springer. S. 23-43.
- Stöckl, Hartmut (2011): Sprache-Bild-Texte lesen. Bausteine zur Methodik einer Grundkompetenz. In: Diekmannshenke, Hajo/Klemm, Michael/Stöckl, Hartmut (Hg.): *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. Berlin: Erich Schmidt. S. 45-70.
- Stöckl, Hartmut (2016): Multimodalität – Semiotische und textlinguistische Grundlagen. In: Klug, Nina-Marie/Stöckl, Hartmut (Hg.): *Handbuch Sprache im multimodalen Kontext*. Berlin/Boston: de Gruyter. S. 3-35.
- van Leeuwen, Theo (2005): *Introducing Social Semiotics*. London: Routledge, 2005.
- van Leeuwen, Theo (2010): *The Language of Colour: An Introduction*. London/New York: Routledge.
- Ware, Colin (2000): *Information Visualization*. Burlington, MA: Morgan Kaufmann.
- Weber, Wibke/Rall, Hans-Martin (2021): The Semiotic Work Design Can Do: A Multimodal Approach to Visual Storytelling. In Ernst, Jutta/von Hoff, Dagmar/Scheiding, Oliver (Hg.): *Periodical Studies Today: Multidisciplinary Analyses*. Bd. 1. *Studies in Periodical Cultures*. Leiden/Boston: Brill, 2022. S. 187-204.
- Weidenmann, Bernd (1997): Multicodierung und Multimodalität im Lernprozeß. In: Ising, Ludwig J./Klimsa, Paul (Hg.): *Information und Lernen mit Multimedia*. Weinheim: Psychologie Verlags Union.