

Zeichen, Text und Sinn – Theorie und Praxis der multimodalen Textanalyse

Hartmut STÖCKL (Chemnitz/Leipzig)

Der vorliegende Beitrag versteht sich als ein Überblick über wichtige Trends in der Entwicklung moderner, semiotisch orientierter Texttheorien und Analyseinstrumentarien. Ausgehend von zentralen Konzepten in der Semiotik der letzten Jahre skizziert er vor allem den Stand der Forschung zum multimodalen Text und reflektiert diesen kritisch. Auf diesem Theoriefundament aufbauend, entwirft der Artikel eine Analysemethodik für den multimodalen Text, die semantischen, syntaktischen und pragmatischen Aspekten sowie der Logik inter-modaler Verknüpfungen nachgeht. Anhand von zwei Textbeispielen, einem gedruckten und einem ausschließlich auditiv verfassten Werbetext, demonstriert der Aufsatz die Handhabbarkeit der vorgeschlagenen Textanalyse-Methodik und thematisiert abschließend zukünftige Entwicklungspotenziale.

Inhalt:

1. Einführung – Was ist Textsemiotik?
2. Richtungen des Wandels in Textlinguistik und Semiotik
 - 2.1 Semiotische Ressourcen und Bedeutungspotenziale
 - 2.2 Konnotation – ‚provenance‘ – ‚experimental meaning potential‘
 - 2.3 Codes – Struktur und Funktion
 - 2.4 Message – Text – Genre – Discourse
 - 2.5 Medium und Materialität
3. Zentrale Elemente einer theoretischen Konzeption des multimodalen Texts
 - 3.1 Deutungspotenzial der Modalitäten und Synergien
 - 3.2 Grammatikalisierung der Modalitäten
 - 3.3 Inter-modale Bezüge
4. Praktische Textanalyse – Beispiele
5. Ausblick
6. Bibliographie

1. Einführung – Was ist Textsemiotik?

Textsemiotik ist alles andere als ein eindeutiger Begriff. Dies rührt vordergründig daher, dass seine beiden Bestandteile Text und Semiotik gleichermaßen schwierige Konzepte darstellen. Während sich Textlinguisten darüber einig sein dürften, dass das Phänomen Text aufgrund seiner Formenvielfalt und funktionalen Flexibilität schwer zu definieren ist und sich daher wohl am besten als eine prototypisch strukturierte Kategorie mit variablen Konfigurationen und Gewichtungen von Textmerkmalen beschreiben lässt (Sandig 2000), ist die Semiotik gar „a science still in the process of being defined“ (Martin/Ringham 2000: 6). Pro-

blematisierend kommt hinzu, dass Semiotik als „all forms of formation and exchange of meaning on the basis of phenomena which have been coded as signs“ (Johansen/Larsen 2002: 3) notwendigerweise auf Texte bezogen sein muss, andererseits Texte aber immer aus strukturierten Ansammlungen von Zeichen bestehen. Diesem Tautologieverdacht zum Trotz meine ich, dass der Gebrauch des Begriffs ‚Textsemiotik‘¹ als Disziplin der Beschreibung von Zeichenstrukturen zum Zwecke der Bedeutungs-(re-)konstruktion in konkreten Texten und Textgenres dennoch sinnvoll und berechtigt ist, vorausgesetzt man beachtet zumindest folgende Akzentuierung:

- (i) Semiotik des Textes meint – etwa in Abgrenzung zur linguistischen Semantik – die Architektur und soziale Bedeutung aller möglichen verschiedenen Arten von Zeichen. ‚Textsemiotik‘ betont daher die multimodale Natur von Textprodukten und zeigt damit eine Affinität zu Medientexten.
- (ii) ‚Textsemiotik‘ suggeriert die Dynamik der Bedeutungskonstruktion durch Textproduzent und Rezipient in einem soziokulturellen Kontext. Bedeutungen werden also erst in einem aktiven Prozess der Sinnstiftung generiert, sie sind nicht a priori vorhanden – „objects do not signify by themselves“ (Martin/Ringham 2000: 7).
- (iii) So wie die semiotischen Ressourcen zum Texte Machen vielgestaltig, flexibel einzusetzen und kreativ erweiterbar sind, so bedarf es bei der Analyse gegenwärtiger Medientexte eines methodischen Nebeneinanders, das vom Ziel einer umfassenden Synthese einzelner Beobachtungen auf vielen Bedeutungsschichten geleitet wird.

Im vorliegenden, zwischen Theorie und Praxis vermittelnden Beitrag soll ein breiter methodischer Rahmen für die praktische Analyse und Interpretation multimodaler Texte skizziert werden. Dabei möchte ich zum einen zeigen, wie facettenreich die analytische Arbeit am Text sein kann. Bewusst wähle ich zwei recht verschiedene Medientexttypen: 1) die Werbepostkarte als junges, experimentelles und pragmatisch anspruchsvolles Genre und 2) den Radiofunkspot als flüchtigen, ‚time-based‘ Text, der im Sinne eines „Kino für das Ohr“ auditive Zeichenmodalitäten (Geräusche, Musik) mit gesprochener Sprache verknüpft. Zum anderen geht es mir darum, wichtige Konzepte einer multimodalen Texttheorie schlaglichtartig herauszuheben und damit den aktuellen Stand der Diskussion um den multimodalen Text zu umreißen und ihr Richtung zu geben.

¹ Hodge/Kress (1988: 2) fordern von einer praxiszugewandten Textsemiotik die folgenden Qualitäten: „...some kind of semiotics that must provide this possibility of analytic practice, for the many people in different disciplines who deal with different problems of social meaning and need ways of describing and explaining processes and structures through which meaning is constituted.“

2. Richtungen des Wandels in Textlinguistik und Semiotik

Anfänglich verband sich mit dem Begriff Text die Vorstellung einer über dem Satz anzusiedelnden Beschreibungsebene sprachlicher Formen. Die rein grammatische Auffassung vom Text als komplexe Sprachstruktur wird in ihrer engen Begrenztheit erst in funktionalistischen Theorien von Sprache und Kommunikation aufgegeben. Hier sind Texte Mittel zur Lösung kommunikativer Aufgaben und die linguistische Sichtweise kehrt sich um: ausgehend von soziokulturell zentralen Textfunktionen sowie stereotypisierten Situationen, Kontexten und Konstellationen der Kommunikation blickt man auf damit gekoppelte sprachliche Gebrauchsformen. Der Wandel von einer strukturellen hin zu einer soziokulturell funktional betriebenen Textlinguistik wird auch vom Kognitivismus befördert, denn fragt man nach den Mechanismen des Verstehens und der Bedeutungskonstruktion im Text, so erlangen Aspekte der kommunikativen Situation und des Wissens der Kommunizierenden größere Bedeutung.

Ein weiterer Trend in der linguistischen Beschäftigung mit Texten hängt mit der zunehmenden Technisierung der Kommunikation bzw. mit einer Diversifizierung möglicher Materialien, Instrumente/Techniken sowie institutionalisierter (Massen-)medien zur Produktion und Distribution von Texten zusammen. War man früher ausschließlich mit der sprachlichen und thematischen Struktur eines Textes befasst, so gehört heute ein Blick auf seine Medialität und Materialität zum analytischen Rüstzeug, weil sie entscheidend zur Bedeutung eines Textes beitragen und damit die Handhabung und Nutzbarkeit eines Textes bestimmen können. In ähnlicher Weise wird die Sicht auf den Text in neueren Theorieentwürfen dynamisiert und vom statischen Produkt auf die Umstände und Prozesse seiner Kreation und Interpretation erweitert. Zumindest zwei Teilaspekte sind hier auszumachen: 1) Produktanalysen werden durch Rezeptionsanalysen und Befragungen der Textproduzenten² ergänzt, um die komplexen Prozesse der Bedeutungs-(re-)konstruktion in Produktion und Interpretation nachzuzeichnen und das Verständnis dafür zu schärfen, dass Semiose im Kontext multiple Lesarten zulässt. 2) Das pragmatische Handeln der Kommunizierenden (z.B. Sprechakte ausführen, präsupponieren, implizieren, inferieren, direkt oder indirekt referieren etc.) wird beschrieben, um die kooperative Textarbeit zwischen Produzent und Rezipient zu erfassen.

Schließlich erscheint der Text heute nicht mehr als einheitliches, homogen strukturiertes sprachliches Phänomen. Zum einen sieht man ihn als Ort und Vehikel der funktionalen und sozialen Differenzierung des Sprachgebrauchs. Texte sind deshalb immer typisierte Kommunikationsprodukte, die ihrer praktischen Funktion und dem sozialen Status der Kommunizierenden Rechnung

² Zu Methoden der Textproduktionsforschung siehe u.a. Perrin (2001, 2003), der vor allem an subjektiven Theorien der Schreiber über ihre Gestaltungsentscheidungen und Schreibstrategien interessiert ist.

tragen. In massenmedialer Textproduktion werden zudem auf äußerst diffizile Weise die vermeintliche soziale Verortung/Orientierung sowie der ästhetische Anspruch und die kognitiven Voraussetzungen des Zielpublikums reflektiert und in der Formulierungsarbeit und dem Design des Texts berücksichtigt (*audience design*, cf. Bell 2001). Zum anderen – und hier ist der Zusatz „semiotisch“ am Platze – hat die Textlinguistik begonnen, sich vom rein sprachlichen Interesse am Text zu emanzipieren. Hier öffnet sich der Blick auf die Gesamtarchitektur des multimodalen Texts, der zwar oft in zentraler Weise sprachlich verfasst bzw. geleitet ist, seine Textbedeutung jedoch aus einer Synthese mehrere Zeichenmodalitäten schöpft. Auf diesem Gebiet liegen die größten Entwicklungspotentiale einer noch jungen linguistischen Teildisziplin so wie auch hier die stärksten praktischen Impulse über die Grenzen der Sprachwissenschaft hinaus in eine textinteressierte und analysewütige Öffentlichkeit möglich wären. Zudem ließe sich auf der Ebene des Textes die semiotische Grundidee gewinnbringend exemplifizieren und erklären, dass jegliche Kommunikation zeichenbasiert und Sprache nur eines der verfügbaren semiotischen Systeme des Menschen unter vielen ist. Semiotische Basisterminologie kann also neben Sprache auf alle möglichen Zeichenmodalitäten übertragen werden und für die Sprache konzipierte Beschreibungsmodelle können auf ihre Tauglichkeit in Anwendung auf nicht-sprachliche Codes überprüft werden.

Im Folgenden will ich an wenigen ausgesuchten neueren Begriffsprägungen demonstrieren, wie sich semiotisches Denken über Text und Kommunikation unter dem Einfluss von Multimodalitätstheorien gewandelt hat.

2.1 Semiotische Ressourcen und Bedeutungspotenziale

Spätestens seit den Arbeiten von de Saussure (2001) und Peirce (Nöth 1985: 39ff.), die uns über Zeicheneigenschaften und -typen aufklären, gilt die Zeichenbasiertheit als konstitutives Kriterium kommunikativer Prozesse. Texte sind im doppelten Sinne semiotisch komplex: einerseits verfügen sie über strukturierte Ansammlungen von Zeichen, die den Regeln eines Codes genügen müssen, andererseits verknüpfen die meisten Texte darüber hinaus mehrere verschiedene Codes miteinander, was die Frage nach den Regeln und Prinzipien dieser Codeverkopplungen auf den Plan ruft. Während sich die traditionelle Textlinguistik letzterem Aspekt bei Beschränkung auf den Sprachcode verweigert hat, ist sie ersterem Aspekt nachgegangen, indem sie das syntaktische Regelwerk auf den Text ausgeweitet hat. Sprachliche Codes erscheinen demnach als festes Repertoire von Regeln – beherrschen sie Produzent und Rezipient, scheint textbasierte Kommunikation nichts weiter als ein mechanisches En- und Dekodieren. Semiotiker – insbesondere unter dem Einfluss der steigenden technisch-(massen-)medialen Vermitteltheit von Kommunikation und der damit zunehmenden Mehrfachkodiertheit von Texten – ziehen es heute vor, statt fester, klar strukturierter Codes von *semiotic resources* zu sprechen (Kress/van Leeuwen 2001, Jewitt/

Oyama 2001: 134ff., van Leeuwen 2005a). Der Begriff betont folgende Aspekte von Zeichensystemen: ihre kulturelle und soziale Gebundenheit und damit auch historische Wandelbarkeit, ihren kontext-sensiblen Gebrauch zu sozial bestimmten Zwecken, ihr kognitives Fundament im Wahrnehmen von Weltausschnitten und dem Kategorisieren von Wissensbeständen, und vor allem die kreative Erweiterbarkeit der Zeichenbestände und deren permanente Umdeutung. Die klassische Vorstellung fester Form-Inhalts-Koppelungen ist einem Konzept gewichen, nach dem eine Zeichenform ein Bedeutungspotenzial, ein *field of possible meaning* (Jewitt/Oyama 2001: 135) eröffnet, das je nach Kommunikationssituation und -erfahrung, Vorwissen und Sozialisierung verschieden aktiviert werden kann. Stötzner (2003: 296) spricht von *semantic scale*.

Diese Akzentuierung des Zeichen- und Codebegriffs hat zur Folge, dass Rezipientenperspektiven, Lesarten und soziale Funktionen in der Textanalyse wichtiger werden. Eine eventuelle Begleiterscheinung könnte, aber sollte nicht sein, dass man – vor allem aufgrund der noch schwierigen Handhabbarkeit der nicht-sprachlichen Codes – den Strukturen der Zeichensysteme weniger Aufmerksamkeit schenkt. Man mag die neuen Begriffe als bloßen Etikettentausch sehen, denn hinter ihnen steht ja im Kern immer noch Code und Bedeutung. Ich meine jedoch, dass die Begrifflichkeiten ein Umdenken signalisieren: weg von der Statik und Eindeutigkeit semiotischer Prozesse hin zu ihrer flexiblen Dynamik und einer inhärenten Mehrdeutigkeit und prinzipiellen Formbarkeit durch den Zeichenbenutzer. Diese neue Sichtweise trägt der kommunikativen Praxis eher Rechnung und überwindet die theoriegetriebene Eindimensionalität der Modelle. Auch für den Text selbst hat man konsequent einen neuen Terminus ins Spiel gebracht. *Semiotic product/event* (Kress/van Leeuwen 2001: 21) betont die Rolle von interagierenden Produzenten und Rezipienten, die in einem aktiven Kommunikationsprozess ein semiotisches Artefakt produzieren und damit ihre sozialen Interessen verfolgen. Zugleich umschiffert er die Zögerlichkeit, mit der wir – aufgrund seiner prototypischen Assoziation mit Sprache – den Textbegriff auch auf semiotische Strukturen anderer Art (Bild, Musik, Geräusch) anwenden. Produkt und Ereignis scheinen mit Blick auf multimodale Texte wie Radiofonspots, TV-Spots oder Theaterstücke, Opern und Konzerte eher treffend. Letztlich fußt die gesamte Umorientierung moderner semiotischer Theorien des Texts, d.h. weg von einem statischen Strukturalismus mit festen Regeln und Bedeutungen hin zu einem dynamischen Funktionalismus mit formbaren Ressourcen und flexiblen Bedeutungspotenzialen, in einer Auffassung von Kommunikation, die nicht auf dem automatisierten Versenden bzw. Empfangen fixer Botschaften, sondern auf dem sozialen Verhandeln von semiotischer Artikulation und Interpretation beruht (Kress/van Leeuwen 2001: 8f.). Kommunikation bedarf also immer der aktiven *semiotic work*, d.h. engagierter Semiose und ist in ihrem Ausgang nie restlos bestimmt, sondern stets unsicher.

2.2 Konnotation – ‚provenance‘ – ‚experiential meaning potential‘

Bei der Übertragung semiotischer Modelle auf die verschiedenen Zeichensysteme multimodaler Texte hat sich ein Konzept besonders bewährt: Konnotation³, verstanden im Sinne Barthes (1996) als um den Begriffskern bzw. das Denotat herum gruppierte Bedeutungsassoziationen. Sie kommen dadurch zustande, dass ein Zeichen oberhalb der denotativen Ebene zum Signifikanten eines weiteren, mit dem ersten kontiguitiv verbundenen Signifikats wird. Dadurch entsteht ein neues, sekundäres Zeichen. Diesen semiotischen Prozess, der in kommunikativen Ereignissen massenhaft abläuft, nennt Barthes auch Mythos oder symbolische bzw. ideologische Bedeutung: Konnotative Bedeutungen konkreter Zeichen entfalten sich immer nur im Kontext eines semiotischen Produkts und werden zudem vom Stil ihrer Präsentation und ihrer Materialität, d.h. der ganzheitlichen Anmutung eines Texts bestimmt. So verweisen Mond und Sterne in einem Bild nicht allein auf den Zeitpunkt der dargestellten Szene, sondern können auch Romantik, Stille oder Mystik konnotieren. In einem Werbetext denotiert das Kompositum ‚Hydro-Regulativ-Faktor‘ eine Substanz, die den Feuchtigkeitsgehalt der Haut reduziert, assoziiert zudem aber Technizität und Pseudowissenschaftlichkeit. Englische Schreibschrift kann jede Phonemkombination der deutschen Sprache darstellen, ihre konnotative Bedeutung jedoch verweist auf Eleganz, Feminines und klassische Schönheit. Bedřich Smetanas symphonische Komposition „Die Moldau“ imitiert das Fließen eines Flusses in verschiedenen geographischen Regionen, mag aber zudem eine Vielfalt von Konnotationen vermitteln, wie z.B. Unbeschwertheit, Kraft und Energie. Ein Zischen referiert z.B. auf das Öffnen einer Flasche, kann aber auch solche Konzepte wie Frische oder Kühle assoziieren.

In dem Bemühen, die Bedeutungsbreite von Zeichen jeglicher Modalität besser beschreiben zu können, prägen Kress/van Leeuwen (2001: 10f., 22f.) die Begriffe *provenance* und *experiential meaning potential*. Diese beiden Termini sind letztlich Präzisierungen der Grundidee, dass Zeichen Konnotationen haben. *Provenance* bezieht sich auf den permanenten semiotischen Import aus Genres und Kommunikationsbereichen, in denen Zeichen zuvor in mehr oder weniger stark tradierter Form verwendet worden sind, so dass sie in dem neuen Gebrauchskontext Assoziationen zu den alten herstellen und diese konnotieren können. Als man begann, das Saxophon in symphonischer Musik zu verwenden, transportierte es aufgrund seiner Herkunft aus dem Jazz automatisch Konnotationen dieses Musikgenres, wie Modernität, Ungestüm und Tanzsalons. Frakturschriften signalisierten früher aufgrund ihrer Herkunft aus der Gotik und Reformation eine ehrwürdige Vergangenheit und Solidität, heute sind sie ein Index für die Kriegs- und NS-Zeit oder rechte neonazistische Propaganda. *Experiential*

³ Zu Deutung und Geschichte des Begriffs sowie seiner Verwendung in der Semiotik siehe van Leeuwen (2005a:37ff).

*meaning potential*⁴ hingegen betont den Status metaphorisch-metonymischer oder synästhetischer Bedeutungsübertragungen als zentrales Mittel der Deutung von Zeichen und deren Bedeutungserweiterung und -aufladung. Für Lexeme natürlicher Sprachen ist dieses Phänomen satzungsbekannt: das Adjektiv ‚heiß‘ z.B. kann die Bedeutungen ‚sehr warm‘, ‚heftig‘, ‚leidenschaftlich‘, ‚erregend‘, ‚viel versprechend‘, ‚sehr schnell‘, ‚interessant‘ etc. haben, weil das physikalische Phänomen Hitze in unserer praktischen Erfahrung mit den verschiedensten mental-emotionalen Empfindungen verkoppelt ist. Für Paraverbales und Typographie, Musik, Geräusche und Bilder lässt sich das Konzept ebenfalls nutzbar machen. So z.B. konnotiert Flüstern die Vertraulichkeit einer Botschaft, krakelige und unsaubere Schrift die Sorglosigkeit oder Ungeübtheit des Schreibers, tiefe Töne Dunkelheit oder Traurigkeit, Rattern und Rauschen schnelle Bewegung, und in bildlichen Darstellungen stehen Lenkrad oder Rad für Auto oder eine Brücke für Kontakt oder Kommunikation.

Nun wäre es töricht zu behaupten, die Textanalyse erschöpfe sich im Nachvollziehen von Konnotationen, Zeichenimporten oder metaphorischer, metonymischer und synästhetischer Bedeutungsbezüge. Gegenüber den differenzierten Instrumentarien auf den Ebenen Handlung, Thema, Kohäsion etwa nimmt sich das oben Geschilderte eher dünn aus. Andererseits liegt die Stärke dieser semiotisch-semantischen Herangehensweise an ausgewählte Zeichenkomplexe in ihrer All-gemeingültigkeit und alle Codes umfassenden Anwendbarkeit. Für das Funktionieren vieler massenmedialer Texte sind subtile Konnotationsprozesse zentral. Ihnen nicht nachzugehen bedeutet oft, den wichtigsten Teil der semantischen Architektur eines Textes – trotz detaillierter Analyse auf anderen Ebenen – zu übersehen.

2.3 Codes – Struktur und Funktion

Der Kernbehauptung, multimodale Texte bestünden neben sprachlichen auch aus anderen, d.h. bildlich, musikalisch, textgraphisch und paraverbal oder nonverbal etc. verfassten Textteilen, liegt nicht so sehr die Vorstellung zugrunde, dass es sich hier um Einzelzeichen handelt. Dass Musik, Bild, Graphisches, Stimmlich-Akustisches und Tonales zeichenhaft sind, ist sicherlich unstrittig und daher eher banal. Vielmehr beruht die Idee der Multimodalität auf der These, dass man es bei all diesen Modalitäten mit Codes, d.h. also Zeichensystemen zu tun hat, die regeln, wie Zeichen zusammengesetzt werden müssen, damit sich intendierte Bedeutungen ergeben. Klar ist von vornherein, dass manche Codes eher stark sind, in dem sie über relativ feste Form-Inhalts-Passungen verfügen (Sprache), andere hingegen eher schwächer, weil sich sowohl die Frage nach der Segmentierbarkeit von Einzelzeichen als auch nach der Regelmäßigkeit ihrer Kombi-

⁴ Van Leeuwen (2005a: 29ff.) nennt dies jetzt auch *experiential metaphor*, offenbar in Anlehnung an die konzeptuelle Metaphertheorie der kognitiven Linguistik (Kövecses 2002).

nation stellt (Bild, Typographie, Geräusch, Musik, Paraverbales, Nonverbales). Auf einem recht hohen Niveau der Abstraktion lässt sich der Code-Status verschiedener Zeichenmodalitäten überprüfen, indem man für Sprache postulierte formale und funktionale Kriterien prüft. Dafür kommen Bühlers (1934) und Jakobsons (1971) Sprachfunktionen, Hallidays (1994) 'textuelle Metafunktionen sowie Morris' (1971: 367) Teilung der Zeichendimensionen in Semantik, Syntax und Pragmatik in Frage. Die Anwendung dieser Kriterien bedeutet einen Vergleich verschiedener Zeichensysteme mit Sprache, wobei implizit von der Überlegenheit des sprachlichen Codes ausgegangen wird, weil Sprache alle postulierten Funktionen optimal erfüllt. Wenn es generell auch zutrifft, dass mit Sprache koexistierende Zeichenmodalitäten in einigen Punkten „Schwächen“ offenbaren, so lässt sich dies allerdings auch so interpretieren: Jede Modalität hat ihre spezifischen Codeeigenschaften, die sie in der Kommunikation zu bestimmten Leistungen befähigen. Die Verknüpfung verschiedener Zeichensysteme im multimodalen Text erscheint somit als praktische Notwendigkeit, die durch ein symbiotisches Miteinander unterschiedlicher Zeichentypen gerade die Stärke und kommunikative Flexibilität eines semiotischen Produkts oder Ereignisses garantiert. Die Leistungsfähigkeit verschiedener Codes im Vergleich ist ein weitestgehend un- bzw. unterbelichtetes Feld in Multimodalitätstheorien, obwohl es meines Erachtens gerade ihre Basis und ihren Kern ausmachen sollte. An dieser Stelle begnüge ich mich mit einigen kursorischen Beobachtungen und grundsätzlichen Überlegungen.

Zunächst ist die Bühlersche Darstellungsfunktion⁵, d.h. die Fähigkeit von Zeichen zu referieren, für einige Codes empfindlich einzuschränken. So hat Musik (abgesehen von so genannter Programmmusik) Schwierigkeiten, konkrete Bezüge zu einer außersemiotischen Welt herzustellen. Für Geräusche (realistische und erfundene) gilt, dass Referenz zwar unproblematisch ist, jedoch aufgrund der ausschließlich indexikalischen Funktionsweise und des begrenzten Zeichenrepertoires beschränkt ist. Paraverbales (Stimme etc.) und Nonverbales (Mimik/Gestik) sind ebenso wie Typographie eng an Sprache (gesprochen oder geschrieben) geknüpft – daher wird ihre Darstellungsfunktion von der sprachlichen wesentlich vorbestimmt.⁶ Am ehesten ist Nonverbales zu eigenständiger, wenn auch eingeschränkter Darstellung in der Lage. Paraverbales und Typo- wie Textgraphisches überformen und ergänzen die Darstellungsaspekte der Äußerungen und Texte, an die sie gebunden sind. Ohne sie wäre allerdings auch die Darstellungsfähigkeit

der Sprache unmöglich (ohne Typographie kein Text, ohne Stimmqualitäten keine Äußerung). Zum senderbezogenen Ausdruck sind alle Modalitäten außer Geräusche auf ihre spezifische Weise fähig. Die Indexfunktion eines Geräuschs im Sinne eines Verweises auf die Existenz/Anwesenheit von Objekten, Sachverhalten, Handlungen etc. macht seine Darstellungsfunktion aus, über den Kommunikator kann ein Geräusch kaum etwas aussagen. Bis auf Musik, bei der sich ein empfängerbezogener Appell wohl auf die psychophysikalischen Effekte von Musik (Aktivierung, Puls, Motorik, Emotionen) beschränkt, sind alle anderen Modalitäten – wenn auch im Vergleich mit Sprache in anderer Weise – appellativ funktionsfähig. Aus diesem kurzen Überblick lässt sich verallgemeinern, dass Sprache und Bild im multimodalen Text die Darstellung tragen (bei Abwesenheit des Bildes etwa im Radiotext taugen auch Geräusche punktuell zur Darstellung), während sich Musik eigenständig sowie Para- und Nonverbales im Verbund mit Sprache vor allem zum Ausdruck eignen. Die expliziteste Form des Appells ist zweifelsohne durch Sprache umsetzbar; der Appell wird subtiler, unterschwellig, und dabei nicht weniger wirkungsvoll, wenn er mit anderen Modalitäten realisiert ist.

Jakobson erweitert Bühlers Sprachfunktionen um den phatischen, metasprachlichen und poetischen Aspekt. Dabei ist die phatische Funktion von Zeichensystemen, d.h. die Verwendung eines Codes zum Testen und Aufrechterhalten von Kanal und Medium der Kommunikation, recht diffus und vage. Mir scheint allerdings, dass alle Codes dazu fähig sind – die an Sprache gekoppelten peripheren Modalitäten wie Typographie, Para- und Nonverbales ohnehin; aber auch Bilder, Geräusche und Musik können, wenn sie Darstellung bzw. Ausdruck oder Appell enthoben sind, phatisch wirksam werden. Die Möglichkeit, dass sich ein Code selbst kommentieren und reflektieren kann, scheint mir der Sprache vorbehalten – diese Funktion erfüllt kein anderes Zeichensystem. Poetisch hingegen können vor allem Musik, Bild und Typographie wirken, indem sie mit dem eigenen Codepotenzial spielen und dabei die Form über den Inhalt stellen. Geräusche sind aufgrund ihrer simplen Funktionsweise dazu nicht in der Lage, Para- und Nonverbales aufgrund ihrer Gebundenheit an Sprache nicht in eigenständiger Weise. Jakobsons Blick auf Zeichensysteme zeigt also, dass Sprache das Privileg hat, metakommunikativ zu sein. Für den multimodalen Text bedeutet dies, dass die Reflexion von Form und Inhalt semiotischer Produkte eine Aufgabe der sprachlichen Teiltexthe sein muss. Die poetische Kraft eines multimodalen Texts kann sich potenzieren, weil auch Musik, Bild und Textgraphisches hierzu beitragen können.

Halliday (1994) und in seiner Folge Kress/van Leeuwen (1996: 40) postulieren drei Metafunktionen, die Zeichensystemen zukommen, damit sie im vollen Umfang Aufgaben in der Kommunikation übernehmen können. Die *ideational* oder *representational meta-function* deckt sich weitestgehend mit der Bühlerschen Darstellungsfunktion. Ich habe bereits gezeigt, warum Musik, Geräusche und die an Sprache gebundenen peripheren Modalitäten (Para-, Nonverbales, Typogra-

⁵ Bei Jakobson (1971) heißen die Bühlerschen Funktionen wie folgt: Darstellungsfunktion = referentielle/kognitive Funktion, Ausdrucksfunktion = emotive/expressive Funktion, Appellfunktion = konative.

⁶ Diese Modalitäten habe ich a.a.O. (Stöckl 2004b: 14) als periphere bezeichnet, weil sie „inevitable by-products“ oder „inherent elements of a core mode's specific medial realisation“ sind. Das heißt nicht, dass sie weniger wichtig wären, ihre Funktionsweise ist aber eng mit dem Gebrauch von Sprache verknüpft. Siehe zu nonverbaler Kommunikation Nöth (1985: 321-380) und zu Parasprachlichem Nöth (1985: 275-279).

phie) diesbezüglich eingeschränkte Funktionalitäten besitzen. Da Kress/van Leeuwen (1996: 40) die Repräsentation weiter fassen, d.h. auch als Bezug auf andere semiotische Systeme, würde man wahrscheinlich allen Codes eine konzeptionelle Funktion zuerkennen. Die interpersonelle bzw. interaktive Metafunktion meint, dass semiotische Systeme nur dann kommunikativ wirksam sein können, wenn sie pragmatisch-soziale Bezüge zwischen Zeichenproduzenten und -rezipienten konstruieren. In dieser Rubrik würde ich allen Codes des multimodalen Texts die volle Funktionsfähigkeit zuerkennen, wenn auch hier Besonderheiten zu beachten sind, wie z.B. bei Geräuschen, deren simple Funktionsweise komplexe interaktive Bezüge verbietet. Schließlich sieht die textuelle Metafunktion vor, dass nur die Zeichensysteme uneingeschränkt kommunikativ sein können, die aus Einzelzeichen formal und inhaltlich zusammenhängende komplexe Gestalten komponieren. Auch hier erscheinen Geräusche wiederum problematisch, da sie – außer in rhythmisierten Geräuschkollagen oder in Akustikrätseln (wo sie rein additiv gereiht werden) – kaum eigenständige Texte bilden können. Sonst verfügen alle Modalitäten (ausgenommen wiederum die an Sprache geknüpften peripheren Modalitäten) über jeweils eigene Muster des Zusammenfügens einzelner Zeichen zu Textgebilden. Musik kennt Melodieführung, Phrasen, Motive, Satzformen etc., Bilder verfügen über Kompositionsstrukturen (Zentrum/Rand, Oben/Unten, Links/Rechts) und Typographie prägt sich in „Dispositiven“⁷ aus, d.h. institutionalisierten und konventionalisierten Mustern der Konfiguration typographischer Elemente auf der Seite und im Dokument. Die konkreten Arten der Nutzung solcher Vertextungsprinzipien bestimmen in entscheidender Weise das Genre bzw. die Textsorte des Textemplars. Mit Hilfe der allgemein gehaltenen Hallidayschen Metafunktionen lässt sich am ehesten der Beweis antreten, dass es sich bei den möglichen Zeichenmodalitäten des multimodalen Texts in der Tat um mehr oder weniger vollwertige Codesysteme handelt. Schließlich können Zeichensysteme auch auf ihre semantische, syntaktische und pragmatische Funktionsfähigkeit hin geprüft werden.⁸ Da alle Codes über mehr oder weniger umfangreiche Zeichenvorräte verfügen, denen Bedeutungen (wie auch immer vage und kontextabhängig diese sind) zukommen, kann man sie alle als semantisch funktionsfähig bezeichnen.⁹ Einschränkungen gibt es bei den Geräuschen hinsichtlich der Menge verfügbarer Zeichen und bei Musik wie auch beim Bild hinsichtlich der Segmentierbarkeit minimaler Zeichen (d.h. z.B. Töne

⁷ Zum Begriff des typographischen Dispositivs siehe (Wehde 2000: 119), Stöckl (2004c: 42ff.) und Stöckl (2005: 208). Jede Zeitung z.B. verfügt über ein typographisches Dispositiv – dies sind alle typ- und textgraphischen Festlegungen von Schriftart und -größe bis hin zu Layout, Papier und Druck. Es ist einleuchtend, dass sich Zeitungen hier gravierend unterscheiden (z.B. FAZ vs. BILD).

⁸ Zu diesen drei grundlegenden Dimensionen der Semiotik siehe Nöth (1985: 48-52).

⁹ Für Sprache und Bild habe ich a.a.O. (Stöckl 2004a: 246f.) in der Gegenüberstellung gezeigt, dass sich die Semantik von Zeichen in den drei Dimensionen Zeichenstruktur, kognitive Anforderungen und Bedeutungspotenzial untersuchen lässt.

oder Bildpunkte). Eine Syntaktisierung, d.h. lineare oder räumliche Zusammenfügung von kleineren Zeichen zu größeren Einheiten ist generell in allen Zeichenmodalitäten möglich – wenn auch mit unterschiedlichem Anspruch an Komplexität und Regelbindung. Bei Geräuschen und Nonverbalem z.B. scheint lediglich eine additive Verbindung mehrerer Geräusche (wie im Akustikrätsel) oder Gesten möglich. Das pragmatische Funktionieren aller Zeichenmodalitäten ist jenseits jeglichen Zweifels – hier ist klar, dass Textproduzenten bestimmte Kombinationen von Zeichen in bestimmten Situationen einsetzen, um die Rezipienten zu kognitiven, psychischen oder praktischen Handlungen zu bewegen. Auch mit dem Raster Semantik – Syntax – Pragmatik lässt sich also deutlich zeigen, dass die unterschiedlichsten semiotischen Modalitäten in der Tat Codestatus reklamieren können.

Zusammenfassend empfiehlt sich wahrscheinlich am ehesten eine von Prototypizität geleitete Sichtweise, d.h. die Codes, die alle „Anforderungen“ an ihre Systematizität optimal erfüllen, gelten als zentrale Vertreter von Zeichensystemen. Hier ist es nicht von ungefähr, dass Sprache am besten abschneidet, sind doch unsere Kriterien für die Codehaftigkeit von Zeichen an ihr modelliert. Geräusche z.B. sind eher randständige Zeichensysteme, ebenso wie periphere, an Sprache gebundene Modalitäten (Nonverbales, Paraverbales) keine zentralen Codes sein werden. Dabei ist allerdings erneut zu betonen, dass hier eine logozentrische Sicht zugrunde liegt. Konzentriert man sich eher auf die recht unterschiedlichen Funktionsweise und Leistungen der verschiedenen Zeichensysteme, so gelangt man eher zu einer Vorstellung von gleichberechtigter und kooperativer Koexistenz – dies ist eine nützliche Sichtweise in multimodaler Textanalyse und -theorie. Mir scheint zudem die Beobachtung wichtig, dass Codes ineinander übergehen können und sich strikte Grenzen mitunter nicht ziehen lassen, wie z.B. im Falle von Musik und Geräusch (Radiofunkspot) oder Bild, Sprache und Typographie (Werbeanzeige) (Stöckl 2004b: 18f.). Dies weist auf zugrunde liegende Gemeinsamkeiten der beteiligten Zeichensysteme und auf die Flexibilität von Zeichen in den Prozessen ihrer Verwendung hin.

2.4 Message – Text – Genre – Discourse

Dem Textbegriff liegt inhärent die Idee eines gewissen Grads an Komplexität zugrunde. Das bedeutet, es muss unterhalb der Textebene inhaltlich-formale Elemente geben, die sich zu einer Textstruktur fügen. Diese einfache Überlegung hat die Textlinguistik von Beginn an beschäftigt und ein Großteil der Arbeit ist hier investiert worden: (Makro-)Propositionen, Textteile, logisch-semantische und Handlungsstrukturen, Kohäsions- und Kohärenzmittel etc. sind einige Termini und Konzepte, die diesbezüglich kursieren. Auf den multimodalen Text bezogen ergeben sich hier zwei Dimensionen. Zum einen müssen in unterschiedlichen Zeichenmodalitäten kodierte Textteile miteinander verknüpft werden, so dass man von inter-modalen Bezügen, d.h. „grammatisch“-semantischen Bezieh-

ungen zwischen Elementen, also multimodaler Kohäsion sprechen könnte. Zum anderen muss auch der in anderen als der sprachlichen Modalität kodierte Text(-teil) (Akustikrätsel, Symphonie, Plakat, bildliche Instruktion etc.) gliederbar sein, d.h. es müssen sich Konfigurationen von Elementen zu einem Ganzen ausmachen lassen. Die Vorstellung, dass man z.B. in Bildern einzelne visuelle Elemente, in Musikstücken Melodiebögen oder Motive und in einem Dokument textgraphische Elemente unterscheiden kann, ist von Kress/van Leeuwen (1996: 183-185, 214-218, 2001: 2f.) mit dem Begriff *framing* bezeichnet worden. Dieser besagt, dass in Gestaltung und Wahrnehmung „Rahmen“ konstruiert werden, in denen jeweils kleinere Elemente zusammengefasst und von anderen Gestalten abgetrennt sind. Das *framing* wird jeweils durch codeeigene Mittel realisiert, wie z.B. Farben, Formen, Leerräume im Bild, Rhythmus- oder Tonartwechsel in der Musik oder Blockbildung im textgraphischen Design. Für minimale Textteile, die ein multimodales Produkt formal-semantisch gliedern, empfiehlt sich meines Erachtens der Begriff *message*, den Hodge/Kress (1988: 5) als „the smallest semiotic form that has concrete existence“ definieren. Ein Text wäre demnach „a structure of messages or message traces which has socially ascribed unity“ (Hodge/Kress 1988: 6).

Textexemplare sind trotz ihrer oft vorhandenen ‚Unikalisation‘ (Sandig 1986: 147ff.) immer Vertreter bestimmter Textsorten oder Genres, deren Muster und Merkmale sie mehr oder weniger prototypisch exemplifizieren. Hodge/Kress (1988: 7) bieten eine bündige Klärung des in der Textlinguistik viel diskutierten Genre-Begriffs: „typical forms of text which link kinds of producer, consumer, topic, medium, manner and occasion. [...] [they] control the behaviour of producers of such texts, and the expectations of potential consumers“. Auf den multimodalen Text angewendet, ergibt sich die berechtigte Frage, ob nicht das Vorhandensein bestimmter Zeichenmodalitäten und deren spezifische Kopplung im Text als ein zentrales Genremerkmal gelten müssen. Ich betrachte diese Sichtweise als dringend geboten: das prototypische Werbeplakat ist ein Sprache-Bild-Text, der prototypische Radiofunkspot ist ein Sprache-Musik-Geräusch-Text etc. Hier bietet sich eine sinnvolle Erweiterung des Textsortenstudiums an, das sich vor allem auf die komplexen Verknüpfungsmuster zwischen den Modalitäten fokussieren ließe.¹⁰

Schließlich scheint neben der Trennung von *message* und Text sowie von Textexemplar und Genre die Unterscheidung von Text und Diskurs sinnvoll, die in der Linguistik verschiedenartig ausgelegt worden ist. Semiotische Theorien betrachten den Text hier als konkretes materielles Zeichenartefakt und verstehen Diskurs dann als „the social process in which texts are embedded“ (Hodge/Kress 1988: 6) oder auch als „socially constructed knowledges of (some aspect of) reality“ (Kress/van Leeuwen 2001: 4). Diese Sicht hat den Vorteil, Texte als

¹⁰ Kopplungsmuster von Modalitäten sind auch für die diachrone Textsortenforschung relevant (cf. Eckkrammer 2004, Gieszinger 2000).

Orte der zeichenhaften Materialisierung von Wissen sehen zu können und Diskurse als die Gesamtheit der Vorstellungen und Konzepte zu einem Thema. Diskurse zeigen sich also letztlich nur in Texten – Texte wiederum sind ohne die „dahinter“ liegenden Wissensbestände undenkbar. Für den multimodalen Text scheint mir die Überlegung relevant, dass bestimmte Wissens- und Informationstypen für ihre Darstellung im Text die jeweils optimal geeignete Zeichenmodalität erfordern. So z.B. lassen sich Zahlen gut in Diagrammen, Wegbeschreibungen in Skizzen oder Karten, merkmalsreiche Gegenstände in Bildern, eine zeitliche Abfolge von Ereignissen in Sprache und Gefühlszustände in Musik kodieren. Der unzureichend verstandenen Beziehung zwischen den Arten von Diskursen und den zu ihrer textlichen Darstellung vorzugsweise genutzten Zeichensystemen gebührt in der multimodalen Texttheorie ein besonderer Platz.

2.5 Medium und Materialität

Es waren die rasante Entwicklung technischer Kommunikationsmedien und die gestiegenen Anforderungen an eine effiziente Übermittlung großer Informationsmengen, die dazu geführt haben, dass Zahl und Vielfalt multimodaler Texte angewachsen sind.¹¹ Die Natur der verfügbaren semiotischen Ressourcen und die Qualität ihrer Nutzung hängen dabei letztlich von der Leistungsfähigkeit der jeweiligen Medien¹² ab. Hier sind enorme Verbesserungen in zweierlei Hinsicht erzielt worden: zum einen sind die Gestaltungspotenziale einzelner technischer Medien wie Druck und Fotografie/Bildbearbeitung revolutioniert worden, zum anderen ist die Verkoppelung der durch die verschiedenen Medien produzierten Zeichensysteme erleichtert und verfeinert worden. Die praktische Folge dieser Entwicklungen ist meines Erachtens in einer stärker wahrnehmbaren Materialität bzw. sinnlichen Wahrnehmbarkeit der Texte zu sehen. Für eine Theorie des multimodalen Texts bedeutete dies, sich Fragen der Materialität von semiotischen Produkten und Ereignissen stärker zuzuwenden (Ormerod/Ivanic 2002, Iadema 2003). Ohne hier Details ausbreiten zu können, scheint mir die Materialität eines konkreten Texts dadurch beschreibbar, dass man die verwendeten Zeichen tragenden Materialien sowie die Praktiken des Signierens (Stötzner 2003: 298) auf ihre Funktionen in Textstruktur, -semantik und -pragmatik hin untersucht.

¹¹ Kress (1998: 58) bezeichnet diese und andere Veränderungen „in the semiotic landscape“ als „deep shift taking place in the system of media and modes of representation...“ und sieht ihre Hauptrichtung von Sprache zum Bild verlaufend, d.h. als „a move from narrative to display“ (Kress 1998: 72). Monomodalität erscheint mir zudem ohnehin eine theoretische Fiktion zu sein, denn auch Sprache allein kommt nie ohne Para- und Nonverbales aus.

¹² Zum diffizilen Begriff des Mediums aus semiotischer Sicht siehe Nöth (1985: 130) und Kress/van Leeuwen (2001: 22).

3. Zentrale Elemente einer theoretischen Konzeption des multimodalen Texts

Nach diesem gedrängten Überblick über aus meiner Sicht wichtige Eckpunkte einer zeitgemäßen semiotischen Theorie des Texts will ich im Folgenden kurz skizzieren, auf welchen theoretisch-methodologischen Grundpfeilern eine Beschäftigung mit dem multimodalen Text ruht.

3.1 Deutungspotenzial der Modalitäten und Synergien

Zunächst ist davon auszugehen, dass jeder Modalität ein eigenes, allgemeines Darstellungs- und Deutungspotenzial zukommt. Das bedeutet, die semiotisch-kognitiven Eigenschaften eines Codes prägen seine Verwendbarkeit, sie erlauben die Übermittlung bestimmter Botschaften und eignen sich für bestimmte mentale Operationen der Rezipienten. Für die Zeichenmodalitäten Sprache, Bild, Typographie, Musik und Geräusch ergibt sich dabei folgendes Bild (s. Tab. 1).¹³

Weil die eine Zeichenmodalität in der textuellen Verwendung Funktionen übernehmen kann, für die eine andere nicht geeignet ist, muss man das Verhältnis der Modalitäten untereinander generell als ein symbiotisches bzw. synergistisches bezeichnen. Das bedeutet, dass sich die semantisch-pragmatischen Möglichkeiten der einzelnen Codes wechselseitig ergänzen und jedes Zeichensystem seinen Platz im gemischten Gesamttext einnimmt. Deshalb sind auch Vorstellungen naiv, nach denen man etwa eine in der einen Modalität kodierte Botschaft einfach in eine andere übersetzen könnte. Zwar sind solche ‚transductions‘ (Kress 1998: 76) mitunter möglich – man kann z.B. sprachliche Botschaften visualisieren, Musik sprachlich beschreiben etc. Dabei verändern sich aber die Art der Botschaft und ihr Ausdruckspotenzial grundlegend. Prinzipiell werden in der Herstellung eines Textes, je nach Inhalt und Intention, konzeptuelle Einheiten jeweils in die Zeichenmodalitäten gebracht, für die sie sich eignen. Verkompliziert kommt hinzu, dass die saubere Trennbarkeit der Modalitäten und ihrer Darstellungs- und Deutungspotenziale wie in Tab. 1 eine Vereinfachung ist. In Wirklichkeit verfügt eine Modalität über mehrere verschiedene Funktionalitäten – hier sind jeweils die prototypischen kommunikativen Aufgaben betont worden. Darüber hinaus ist auch die Abgrenzung der Modalitäten selbst nicht immer klar. Rhythmisiert man ein Geräusch und gibt ihm Tempo und Metrum, rückt es in die Nähe von Musik; Typographie kann zu Bildlichem mutieren; es lassen sich

¹³ Außer Typographie habe ich hier die so genannten peripheren Modalitäten, wie Nonverbales und Paraverbales ausgeklammert, weil ihre Funktionsweise semiotisch weitestgehend ungeklärt ist. Der Stellenwert von Typographie ist in den letzten Jahrzehnten gewachsen (van Leeuwen 2005b: 142), sie hat sich vom Rand der Semiosphäre in ihr Zentrum verlagert. Deswegen räume ich Typographischem hier auch einen gleichwertigen Platz neben geschriebener Sprache und Bild ein.

sprachliche Bilder formulieren; und gesprochene Sprache hat automatisch musikalische Qualitäten etc. In ihrem Gebrauch also zeigen die verschiedenen Modalitäten eine gewisse Wandelbarkeit, die ich a.a.O. (Stöckl 2004: 18f.) auch als ‚mode overlapping‘ bezeichnet habe. Wichtig bleibt, dass sich die unterschiedlichen Zeichenmodalitäten semantisch, formal und vor allem funktional ergänzen – dies ist die Basis für die Existenz und den kognitiven wie spielerischen Nutzen und Reiz multimodaler Texte.

Abb.1:

MODALITÄT	PROTOTYPISCHES DARSTELLUNGS-/DEUTUNGSPOTENZIAL	CODESPEZIFIKA
1 Sprache	Handlungen in ihrer zeitlichen Dimension sowie gesamtes Spektrum klar umrissener Sprechakte	linearisierte Zeichen – lexikalisch wie illokutiv explizit
2 Bild	visuell merkmalsreiche Objekte, Verortung von Elementen – räumliche Bezüge von Objekten zueinander	ganzheitlich-simultaner Code – vorwiegend ikonisch
3 Typographie	sprachliche und inhaltliche Text- und Dokumentstrukturen, Schreibereigenschaften und -einstellungen,	größtenteils an Sprache gebunden – sonst bildlich
4 Musik	emotionale Befindlichkeit, psychophysikalische Appelle, Form der eigenen Struktur (selbstreferentiell)	direkte Verbindung ins limbische System (Emotionen)
5 Geräusch	Verweis auf die Existenz von Objekten bzw. das Ablaufen einer Handlung oder eines Geschehens	Einzelzeichen in ausschließlich indexikalischer Funktion

Tab. 1: Modalitäten und ihre generellen Darstellungs- und Deutungspotenziale

3.2 Grammatikalisierung der Modalitäten

Im Kern einer Theorie des multimodalen Texts steht die Aufgabe, zu zeigen, welche „Bausteine“ das „grammatische“ Funktionieren der einzelnen Zeichenmodalitäten gewährleisten. In der stark funktionalgrammatisch geprägten Sozialsemiotik beschreitet man den Weg ausgehend von wichtigen Ausdruckspotenzialen, die Modalitäten in Texten verfügbar machen, hin zu den formalen Mitteln, mit denen diese sozial relevanten Bedeutungen realisiert werden. In Anlehnung daran habe ich anderswo (Stöckl 2004b: 11-16) eine einfache Grammatikalisierung von Modalitäten vorgeschlagen, die wichtige Dimensionen eines Codes als Sub-Modalitäten herausstreicht. Diese wiederum besitzen Merkmale, die je nach Bedeutungsabsicht eingestellt werden können. Für die fünf oben (s. Tab. 1) behandelten Modalitäten exemplifiziere ich mein Modell auszugswise (s. Tab. 2).

Abb. 2:

SINNESKANAL	MODALITÄT	SUB-MODALITÄTEN	MERKMALE	BEDEUTUNGSPOTENZIAL
visuell-auditiv	1 Sprache	Sprechakte	direktiv vs. kommissiv	Instruktion vs. Garantie
visuell	2 Bild	Kamerawinkel vertikal	von oben vs. von unten	Betrachter beherrscht visuell Dargestelltes vs. Dargestelltes beherrscht Betrachter
visuell	3 Typographie	Schriftfarbe	schwarz vs. farbig	faktische/subsidiäre Information vs. emotionale/zentrale
auditiv	4 Musik	Instrumentierung	großes Orchester vs. Streichquartett	Dramatik/Öffentlichkeit vs. Stille/Intimität
auditiv	5 Geräusch	Status	realistisch vs. erfunden	Handlungs- oder Objektdarstellung vs. Illustration

Tab. 2: Modalitäten, Submodalitäten und ihre Merkmale

Natürlich verfügt jede Modalität über ein ganzes Spektrum von Sub-Modalitäten (s. Abb. 4) und innerhalb dieser Codebausteine gibt es nicht nur binär einstellbare Merkmale, sondern oft graduelle Abstufungen und damit verbundene Skalen von Bedeutungen. Dies macht das System einerseits vage im Sinne interpretierbarer Bedeutungspotenziale, andererseits aber auch recht flexibel. Gebrauch und Deutung der semiotischen Ressourcen aller Modalitäten lassen sich adäquat nur auf dem Hintergrund einer konkreten Textsorte bzw. eines Textemplars beurteilen – Tab. 2 zeigt in der letzten Spalte eine mögliche Interpretation der Bedeutungspotenziale für den Werbetext. Werden Modalitäten im Gesamttext verkoppelt, so ist für die semantische und formale Stringenz des Texts oft nur *ein* konkretes Merkmal einer Submodalität ausschlaggebend – es steht im Vordergrund der Wahrnehmung, weil es Bezüge zu anderen Modalitäten herstellt. So z.B. kann sich der Phraseologismus „blau machen“ mit der blauen Farbe eines Bildelements oder einer blauen Schriftfarbe koppeln, die Moll-Tonart eines musikalischen Motivs mit den dunklen Farben einer Bildsequenz oder ein schleppendes Sprechtempo mit dem langsamen Tempo einer musikalischen Phrase. Darüber sollte man aber nicht vergessen, dass zum vollwertigen Ausdruck einer Modalität stets das ganze Spektrum von Submodalitäten eingesetzt werden muss, auch wenn sie in der Kommunikation nicht alle fokussiert sind. Komplexe

Zeichengestalten bestehen immer aus ihren inhärenten Elementen – in der Wahrnehmung ist man sich des kumulativ-komplexen Zustandekommens eines Gestalteindrucks nicht bewusst. Es mag sein, dass diese Art semiotische Analyse, d.h. die penible Grammatikalisierung einzelner Modalitäten aufwendig ist; ein einfacherer Weg daran vorbei ist allerdings kaum vorstellbar.¹⁴ Zum Teil sind die Grammatiken einzelner Zeichenmodalitäten bereits geschrieben, vor allem ihr Zusammenspiel aber harrt noch der Erkundung.

3.3 Inter-modale Bezüge

Sobald ein Text aus mehr als einer Zeichenmodalität besteht, treten die unterschiedlich kodierten Textteile zwangsläufig in eine Beziehung. Diese kann formaler Natur, d.h. an der materiell wahrnehmbaren Gestalt der Zeichen orientiert, oder semantischer Natur, d.h. an den durch die Zeichen transportierten Konzepten orientiert sein. Für die Funktionsweise multimodaler Texte sind diese intermodalen Brücken oder Kontaktstellen essentiell – sie zu analysieren bedeutet, der Struktur der Texte und ihrer pragmatischen Funktionsweise habhaft zu werden (Royce 1998). Formale Bezüge können nur entstehen, wenn die verkoppelten Codes Sub-Modalitäten teilen bzw. wenn sich Sub-Modalitäten in irgendeiner Weise kompatibel zeigen. So z.B. verfügen die zeit-basierten Modalitäten, Sprache, Musik und Geräusch alle über Rhythmus, so dass hier eine formale Verknüpfung möglich ist. In ähnlicher Weise teilen sich Bild und Typographie die Sub-Modalität Farbe (z.B. grüne Schrift – Bild eines Waldes). Durch solche formalen Brücken entsteht vor allem inter-modale Kohäsivität. Eine inhaltlich-konzeptuelle Verknüpfung von Modalitäten ist auf vielfältige Weise möglich. Die kognitive Grundlage dafür sind Assoziationen (z.B. schnelles Tempo – Eile), Synästhesien (z.B. tiefer Ton – dunkel), metaphorisch-metonymische Übertragungen (z.B. Sturm – Zorn, Glocken – Kirche) sowie paradigmatische Sinnbezüge (Antonymie, Hyperonymie) und logische Relationen (Kausalität, Finalität). Inter-modale Bezüge sind bisher weniger untersucht worden als die Funktionsweise einzelner Modalitäten, weil sie eine komplexe Materie bilden. Verkompliziert wird deren Studium durch die Textsortengebundenheit und Kontextsensibilität solcher Bezüge und generell dadurch, dass sich die Vielfalt möglicher Bezüge bei der Verkoppelung von mehr als zwei Modalitäten potenziert. Andererseits entstehen in den gängigen massenmedialen Textsorten mehr oder weniger konventionelle und fixe Gebrauchsmuster von Modalitätskoppelungen, die die

¹⁴ Neuere Entwicklungen in der Sozialesemiotik (van Leeuwen 2005a) erwecken den Eindruck, das Schwergewicht der Beschäftigung mit Zeichen liegt auf den in Texten häufig vorkommenden Funktionen und Ausdruckspotenzialen von Modalitäten, während man weniger gewillt ist, sich mit der inneren Struktur der Zeichensysteme zu beschäftigen. „Grammatik“ wird zwar oft, zumeist jedoch im metaphorischen und daher verallgemeinerten Sinn verwendet. Hier liegt eine gewisse Gefahr.

Beschreibung erleichtern.¹⁵ Generell wird das Verhältnis zweier Modalitäten den in Tab. 3 dargestellten pauschalisierten Konstellationen folgen, die zumindest ein grobes Ordnungsraster bilden. Die Praxis freilich ist komplexer und subtiler.

Abb. 3:

ADDITIV	Zeichenmodalitäten verstärken sich gegenseitig in Ausdruck und Wirkung
KOMPLEMENTÄR	Zeichenmodalitäten ergänzen sich wechselseitig im Verstehen
HIERARCHISCH	Zeichenmodalitäten ordnen sich in einer Rangfolge der Dominanz und Wichtigkeit
DIVERGENT	Zeichenmodalitäten kreieren auseinander laufende Lesarten
KONFLIKTIV	Zeichenmodalitäten laufen einander zuwider und kollidieren semantisch

Tab. 3: Arten der Integration von Zeichenmodalitäten

4. Praktische Textanalyse – Beispiele

Im Folgenden möchte ich an zwei Beispielen zeigen, wie man multimodale Produkte textsemiotisch untersuchen kann. Dazu werde ich zunächst einen schematischen Überblick über die beteiligten Zeichensysteme und deren Sub-Modalitäten geben (s. Abb. 4). Im Anschluss betrachte ich jeweils nur die Ausprägung der Sub-Modalitäten, die für die Bedeutungskonstitution in den Texten ausschlaggebend sind, und gehe vorhandenen inter-modalen Bezügen nach. Meine Beispiele stammen beide aus der Werbung: zum einen analysiere ich eine Werbepostkarte (s. Abb. 5), die neben Sprache über Typographie¹⁶ und Bild¹⁷ verfügt, zum anderen einen Radiofunkspot (s. Abb. 6), der Musik und Geräusche¹⁸ an die Seite der Sprache stellt.

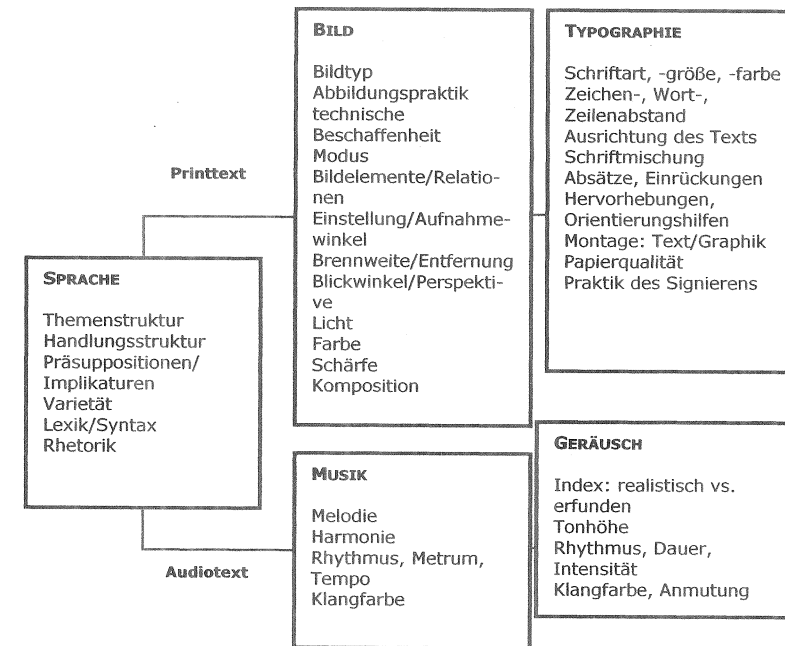
¹⁵ Solche „textuellen Gebrauchsmuster“ von Sprache-Bild-Bezügen sind detailliert in Stöckl (2004a: 253-300) beschrieben.

¹⁶ Zu den Sub-Modalitäten von Typographie siehe Stöckl (2004c, 2005).

¹⁷ Zu einigen wichtigen Sub-Modalitäten von Bildern und deren praktischen Analyse siehe Stöckl (2004d: 111ff).

¹⁸ Zu wichtigen Sub-Modalitäten von Musik siehe Ziegenrucker (1997: 16) sowie zur Semiotik von Musik allgemein siehe Nöth (1985: 390-400). Zu einer funktional-semiotischen Sicht auf Auditives generell, d.h. auf Musik, Geräusch und auf die paraverbalen Aspekte gesprochener Sprache siehe van Leeuwen 1999.

Abb. 4:



Modalitäten in Print- und Audiotexten mit dazugehörigen Sub-Modalitäten

Die Werbepostkarte besteht aus vier Textteilen, denen unterschiedliche pragmatische Funktionen zukommen. Die Vorderseite, die primär bildlich verfasst ist, weckt als eine Art *teaser* die Aufmerksamkeit/Neugier des kursorischen Betrachters. Die Rückseite (linke Innenseite) ähnelt der einer traditionellen Postkarte, sie bietet den konventionellen Platz für die Briefmarke und die Empfängeradresse – wie für Werbepostkarten typisch, ist die linke obere Ecke eine kurze Zusammenfassung der Gesamtbotschaft des Kommunikats. Auf der rechten Innenseite findet sich dann der eigentliche Werbetexte (mit *Headline* und *Body Copy*), während die Rückseite wiederum eine traditionelle Postkartenfunktion hat, die hier auch Absenderdaten verlangt und diese in einen Antwortcoupon zur Teilnahme an einem Gewinnspiel integriert. Diese pragmatisch motivierte Textgliederung wird im Wesentlichen text- und typographisch kommuniziert.

Abb. 5:

ICH HABE DEN ERNST DES LEBENS KENNEN GELEART.
ER IST EIN FURCHTBAR LANGWEILIGER TYP.

Du nicht so erwachsen.

MODUS.DE

BEFREI DICH VOM ALLTAG!

DER RENAULT MODUS BEFREI DICH VOM ALLTAGSSTRESS: SAG UNS, WAS DICH AM MEISTEN NERVT, UND VIELLEICHT BIST DU EINER VON ZEHN GEWINNERN, DIE BALD EINE SORGE WENIGER HABEN, WIE RASEN MÄHEN, WOHNUMG PUTZEN, GETRÄNKE WISSTEN TRAGEN, EINKÄUFEN GEBEN. AUS DEN ORIGINALSTEN ZUSCHRIFTEN WERDEN ZEHN NERVIGE ALLTAGSSORGEN AUSGEWÄHLT UND DEN EINSENDERN DANKT. JE NACH PROBLEM Z. B. FÜR EIN JAHR, ABGENOMMEN. EINFACH AUF DEN COUPON DER POSTKARTE SCHREIBEN, WAS DICH NERVT, DEINEN NAMEN UND DEINE ADRESSE EINTRAGEN UND DEN COUPON IN DIE BOX WERFEN – ODER DIE POSTKARTE IN DEN BRIEFKASTEN. SCHON NIMMST DU AM GEWINNSPIEL TEIL.

WAS MICH SCHON IMMER GENERVT HAT:

SCHÖN, DASS DER RENAULT MODUS FÜR DAS ABNIMMIT!

Werbepostkarte Renault Modus „Befrei Dich vom Alltag“ – Vorderseite & rechte Innenseite

Die Vorderseite koppelt eine personalisierte sprachliche Äußerung mit einem Bild. Dabei sind vor allem die Personalpronomina von Bedeutung (Myers 1994: 77ff.): *ich* bietet den potenziellen Adressaten die Möglichkeit, sich zu identifizieren und als Autor der Äußerung zu projizieren; *er* hingegen verweist (Textdeixis zwischen sprachlichen und bildlichen Elementen) eindeutig auf die schematisch dargestellte Person, die damit als *furchtbar langweilig* kategorisiert wird. Zum Ausdruck kommt also hier eine sozial stereotypisierte Situation, in der eine Frau mit einem Mann (hier eventuell stellvertretend für das Leben insgesamt) unzufrieden ist. Der Fließtext auf der zweiten Innenseite bietet für dieses „Problem“ dann eine „Lösung“. Das Bild auf der Vorderseite macht verschiedene Elemente konzeptuell verfügbar. Wichtig ist es dabei zu erkennen, dass es sich um zwei Bildebenen handelt, die quasi übereinander liegen. Das Format wird von der schematisierten, gesichtslosen Männergestalt beherrscht, dessen Pullovermuster nahtlos in den Hintergrund übergeht. Die Farbe grau steht hier zentral als metaphorisches Lexem für *langweilig*, ebenso das kantige Muster, das biedere Nüchternheit und faden Geschmack signalisiert. Antithetisch hierzu ist die linke untere Ecke des Formats gestaltet. Hier häufen sich farbige Kreiselemente und Farbspritzer, was der an sich statischen Komposition Dynamik verleiht. Dies kommt auch darin zum Ausdruck, dass die graue „Tapete“ an der Stelle aufgerissen ist: das Auto könnte hier als Auslöser dieses Prozesses dienen. Obwohl also das Bild an sich eher konzeptuell ist (Kress/van Leeuwen 1996: 79, Jewitt/Oyama 2001: 143ff.), hat es dadurch auch ein narratives Element (Kress/van Leeuwen 1996: 43, Jewitt/Oyama 2001: 141ff.) – es bringt die Veränderung eines Zustandes durch die Wirkung eines graphischen Vektors (Auto) zum Ausdruck. Für die interaktive Wirkung dieses Bildes zeichnet vor allem der gesichtslose Kopf in Verbindung mit dem frontalen Aufnahmewinkel verantwortlich. Gesichter in der frontalen Perspektive haben normalerweise einen direkt auffordernden Charakter (Kress/van Leeuwen 1996: 121ff.) – dieser kann sich hier aber nicht entfalten, weil die Augen als aus dem Bild weisende Blickvektoren fehlen. Der Sprache-Bild-Text der Vorderseite bringt also das Auto als Problemlöser für ein gelangweiltes Leben einer Frau ins Spiel, indem die Antithese grau/eckig vs. bunt/rund visuell aufgebaut und sprachlich gedeutet wird. In dem eher unspektakulären Fließtext auf der rechten Innenseite sind die Submodalitäten Handlungs- und Themenstruktur vordergründig. Einer Aufforderung (*Befrei Dich vom Alltag!*) folgt die noch vage gehaltene Werbe-Behauptung bzw. ein Versprechen (*Der Renault Modus befreit Dich vom Alltagsstress.*), das dann expliziert und illustriert wird. Der Rest instruiert den Leser bezüglich des Gewinnspiels und erneuert das Versprechen. Lexikalisch zentral ist die Nennung verschiedener stressiger Alltagsaufgaben, die hier wiederum die weibliche Zielgruppenorientierung durchscheinen lassen und deren Übernahme für ein Jahr die Firma in ihrem Gewinnspiel verspricht. Ebenso relevant für die Konstruktion eines sozialen Bezugs zum Leser ist die Wahl der zweiten Person Singular (*Dich, Du, Deine*).

Die Funktion der text- und typographischen Gestaltung besteht einerseits darin, Inhalt und Pragmatik der sprachlichen Äußerung zu unterstützen, andererseits darin, die Struktur des Gesamttexts zu verdeutlichen. Wesentlich ist zunächst der Kontrast zwischen der Schriftart der Vorderseite und der Innenseite. *Ernst* und *furchtbar langweilig* werden dabei durch eine fette, eckig-kantige, gradlinige und letztlich nicht leicht lesbare Schrift realisiert – hier liegt eine inter-modale Parallele zwischen Typographie und Bild. Im Gegensatz dazu ist die Schriftart des Fließtexts dünn, gerundet, und in ihrer vertikalen Achsenführung unstet. Sie weist deutlich in Richtung Handschrift, obwohl es sich natürlich um einen computergenerierten Schriftsatz handelt. Die Schrift wirkt leicht, schwingvoll-dynamisch und ist gut lesbar, ohne normativ zu sein. Die individualisierte Note konnotiert Alltag, passt aber auch zum Kreativen des beworbenen Autos. Zu dieser Schriftart-Antithese gesellt sich eine Gegensatz-Beziehung zwischen lebendig-unruhiger Typographie und dem statisch-technokratischen Kästchenmuster der Schriftfläche. Umgekehrt entsteht allerdings wiederum eine inter-modale Parallele zwischen Typographischem und Bildlichem: das Kästchenpapier passt zur Pullovermustertapete der Kartenvorderseite. Die übrige typographische Gestaltung steht im Dienste der Textgliederung – unterschiedliche Schriftgrößen, Schriftarten und -richtungen in Zusammenhang mit Papierqualitäten (glatt vs. gestrichen) signalisieren verschiedene funktionale Textteile.

Zusammenfassend ist festzustellen, dass die Printwerbung alle drei Zeichenmodalitäten strategisch einsetzt und miteinander koppelt. Wenn wir auch geneigt sind, Sprache eine dominante Rolle zuzusprechen, weil sie aufgrund der expliziten Referenz und der pragmatischen Deutlichkeit die Hauptbotschaft trägt, so muss jedoch daran erinnert werden, wie stark sich die Botschaft, pragmatische Funktionsweise und stilistische Anmutung des Texts ändern würden, wenn Bild und Typographie ausgeblendet wären.

In dem Vergleich einer Werbepostkarte mit einem Radiofunkspot sollte deutlich werden, wie sich die mediale Realisierung eines Genres in den Textstrukturen niederschlägt. Dies ist in erster Linie die Frage nach der Art der Zeichenmodalitäten und ihrer Verknüpfung. Zunächst eignet sich gesprochene Sprache zur Wiedergabe dialogischer Texte, die mit Hilfe paraverbaler Mittel (Stimmqualitäten etc.) eindrucksvoll und differenziert gestaltet werden können. In unserem Beispiel sind die Sprecher als Vater und Kind stereotypisiert („Familie“) und aufgrund der Geräusche ist auch die Pragmatik der Kommunikationssituation leicht zu identifizieren: im Stau per Auto auf einer Reise. Der szenische Dialog dient hier primär der Zielgruppenorientierung und inhaltlich-argumentativ der Schilderung des „Problems“, für das der folgende Textteil des Off-Sprechers eine „Lösung“ anbietet. Logisch-semantisch betrachtet stellt der Off-Sprecher-Teil eine werbliche Behauptung auf, die durch den szenischen Dialog vorwegnehmend bewiesen wird. Die beiden Textteile stehen aber auch antithetisch zueinander; dies zeigt ein Blick auf die Lexik: im ersten Teil informell knapp, im zweiten Teil technisch (*Mitfahrerrabatt*) und faktenüberladen. Sprachlich ähneln sich

beide Textteile durch ihren stark elliptischen Charakter – einmal konnotiert dies die umgangssprachliche Natürlichkeit, ein anderes Mal die Dringlichkeit und nüchterne Faktizität der Botschaften. Allenfalls kommt im zweiten Textteil mit der zeugmatischen¹⁹ Verwendung der Präpositionen *für* und *mit* ein rhetorisch-spielerisches Element zur Geltung, das die Faktenreihung auflockert und „rhythmisiert“.

Abb. 6

ZEIT/SEK.	1-9	10-22	23-25
SPRACHE	A: Papi, is es noch weit? B: Nur noch fünf Kilometer? A: Und wann sind wir endlich da? B: In drei Stunden.	C: Familien fahren besser Bahn. Und Kinder unter fünfzehn kostenlos. Leipzig-München und zurück für nicht mal 89 €. Für zwei Kinder und zwei Erwachsene. Mit Bahncard. Mit Plan- und Spar 40 und Mitfahrerrabatt. Genaue Bedingungen unter www.bahn.de	oder den DB-Verkaufsstellen. Die Bahn.
STIMME	A = genervte Kinderstimme B = angespannte Männerstimme	C = Off-Sprecher; junge, sonore Stimme	
GERÄUSCH	im Hintergrund des Gesprächs: Verkehrslärm, Verkehrsfunkdurchsage, gereiztes Hupen	vorbeirauschender Zug	
MUSIK			im Bahntakt rhythmisierter gleich bleibender Ton, gegen Ende mit Hall ausgeblendet
SPRACHE	DIALOG		MONOLOG
GERÄUSCH	VERKEHRSLÄRM-VERKEHRSDURCHSAGE-HUPEN		VORBEIRAUSCHENDER ZUG
MUSIK			RHYTHMISIERTER TON

Audio-Transkript des Radiofunkspots für Die Bahn „Wann sind wir endlich da?“

Die Geräusche übernehmen in bildloser Audiowerbung eine unersetzliche Funktion. Sie verweisen im ersten Teil auf Ort und Umstände der fiktiven Unterhal-

¹⁹ Hier liegt zweimal ein inkongruentes Zeugma vor (Plett 1991: 57f.), bei dem die Complements/Objekte der Präpositionen semantisch ungleichartig sind: für + Preis/für + Kunden, mit + Bahncard/mit + Tarif.

tung und fungieren so als einfache Indexzeichen. Im zweiten Teil interpretiert man das Rauschen und Rattern des Zuges nicht als den Ort des Sprechers. Vielmehr illustriert das Geräusch den Inhalt des Textteils und suggeriert solche Konnotationen wie Geschwindigkeit und Vorankommen. In der Modalität Geräusch entsteht, wie bereits für Sprache beobachtet, eine auditive Antithese: der erste Geräuschtext signalisiert den Erfahrungsbereich ‚Auto‘ und assoziiert ihn mit lähmendem Stillstand, der zweite Geräuschtext signalisiert ‚Zug‘ und konnotiert Schnelligkeit und Dynamik. Die minimale Art Musik wie sie in dem Bahnspot zur Anwendung kommt dient vor allem der Strukturierung des Gesamttexts – sie setzt als Jingle quasi einen Schlusspunkt. Andererseits entwickelt sie sich harmonisch aus dem Geräusch (Rhythmus und Klangfarbe), so dass inter-modale Kohäsion entsteht. Inhaltlich passt die Musik zum Thema Zug und Bahn fahren, weil sie in ähnlicher Weise wie das Geräusch schnelle Bewegung und aufgrund des Halleffekts Kraft und Intensität vermittelt. So signalisiert die Musik auch die Konstanz des Textthemas.

Die funktional-semiotische Analyse der beiden medial-kontrastiven und genre-gleichen Texte sollte zeigen, wie man Beobachtungen zu den jeweils relevant gesetzten Sub-Modalitäten zusammenfügen und miteinander zu einer schlüssigen Interpretation vernetzen muss, um die Verfasstheit eines Texts in Struktur, Funktionsweise und Effekten nachzuzeichnen.

5. Ausblick

Die Beispieltex-te waren bewusst recht einfach gehalten, um die herausgearbeiteten Grundpositionen und Konzepte zur multimodalen Textsemiotik und –analyse verständlich illustrieren zu können. Vor allem die Musik, aber auch sprachliche Texte und Geräusche sind mitunter weitaus komplexer. Prinzipiell jedoch lässt sich allen Arten von modernen massenmedialen Texten mit der hier vorgeschlagenen Theorie und Methodologie beikommen – dies gilt auch für diffizile humoristische, polyvalente und selbstreferentielle Effekte wie sie in einigen Medientextsorten und künstlerischen semiotischen Produkten häufig vorkommen. In mindestens drei Bereichen besteht Nachholbedarf. Die „Grammatik“ einzelner Codes weist noch Lücken auf, die es zu schließen gilt. Vor allem ist unsere durch logozentrische Konditionierung geprägte Vorstellung vom Text noch nicht sensibel und integrativ genug, um die wahre Leistung und den eigentlichen Wert der Zeichensysteme neben Sprache adäquat wahrnehmen zu können. Hier benötigen unsere Sinne mehr Training und unsere Lehrbücher größere außersprachliche Expertise. Außerdem ist die Transkription multi-modaler Texte noch nicht ausgereift, wenngleich das Bewusstsein für deren Wichtigkeit gelegentlich geschärft worden ist.²⁰ Ziel muss es hier sein, Formate zu finden, die

²⁰ Zur Praxis multimodaler Texttranskription siehe Thibault 2000 und Norris 2004.

einfach und übersichtlich bleiben, ohne wichtige Merkmale eines Texts auszublen-den. Vor allem den nonverbalen und paraverbalen Codes gebührt hier mehr Aufmerksamkeit.²¹ Gleiches gilt für die Interaktion zwischen Bild und Sprache sowie zwischen Musik, Geräusch und Bild. Schließlich sollten die Produktions- und Rezeptionsprozesse eines multimodalen Texts eingehend studiert werden, um die in der Produktanalyse nahe gelegte Lesart anzureichern, zu verifizieren oder zu modifizieren.

6. Bibliographie

- Barthes, Roland (1996): Connotation. In: Cobley, P. (ed.), *The communication theory reader*. London: Routledge, 129-133.
- Bell, Alan (2001): Back in style: reworking audience design. In: Eckert, P./Rickford, J.R. (eds.), *Style and sociolinguistic variation*. Cambridge: Cambridge University Press, 139-169.
- Bühler, Karl (1934): *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Jena: Fischer.
- De Saussure, Ferdinand (2001): *Grundfragen der Allgemeinen Sprachwissenschaft*. Berlin: de Gruyter.
- Eckkrammer, Eva Martha (2004): Drawing on theories of inter-semiotic layering to analyse multimodality in medical self-counselling texts and hypertexts. In: Ventola, E./Charles, C./Kaltenbacher, M. (eds.), *Perspectives on multimodality*. Amsterdam: Benjamins, 211-226.
- Gieszinger, Sabine (2000): Two hundred years of advertising in *The Times*: The development of text type markers. In: Ungerer, F. (ed.), *English media texts past and present*. Amsterdam: Benjamins, 85-109.
- Halliday, Michael A.K. (1994): *An Introduction to Functional Grammar*. London: Arnold.
- Hodge, Robert/Kress, Gunther (1988): *Social Semiotics*. Cambridge/Oxford: Polity Press.
- Idema, Rick (2003): Multimodality, resemiotization: extending the analysis of discourse as multi-semiotic practice. In: *Visual Communication* vol. 2, n° 1, 29-57.
- Jakobson, Roman (1971): *Fundamentals of Language*. The Hague: Mouton.
- Jewitt, Carey/Oyama, Rumiko (2001): Visual meaning: A social semiotic account. In: van Leeuwen, T./Jewitt, C. (eds.), *Handbook of visual analysis*. London: Sage, 134-156.
- Johansen, Jørgen Dines/Larsen, Svend Erik (2002): *Signs in use. An introduction to semiotics*. London: Routledge.
- Kövecses, Zoltán (2002): *Metaphor. A practical introduction*. Oxford: Oxford University Press.
- Kress, Gunther (1998): Visual and verbal modes of representation in electronically mediated communication: the potentials of new forms of text. In: Snyder, I. (ed.), *Page to screen. Taking literacy into the electronic era*. London: Routledge, 53-79.
- Kress, Gunther/van Leeuwen, Theo (1996): *Reading images. The grammar of visual design*. London: Routledge.
- Kress, Gunther/ van Leeuwen, Theo (2001): *Multimodal discourse. The modes and media of contemporary communication*. London: Arnold.
- Martin, Bronwen/Ringham, Felizitas (2000): *Dictionary of Semiotics*. London: Cassell.
- Morris, Charles W. (1971): *Writings on the general theory of signs*. Den Haag: Mouton.
- Myers, Greg (1994): *Words in ads*. London: Arnold.

²¹ Diese Zeichenmodalitäten sind hier ausgeblendet worden. In der gängigen gesprächsanalytischen Praxis wird man ihnen bereits recht gut habhaft.

- Norris, Sigrid (2004): Analyzing multimodal interaction. A methodological framework. London: Routledge.
- Nöth, Winfried (1985): Handbuch der Semiotik. Stuttgart: Metzler.
- Ormerod, Fiona/Ivanic, Roz (2002): Materiality in children's meaning-making practices. In: *Visual Communication* vol. 1, n° 1, 65-96.
- Perrin, Daniel (2001): Wie Journalisten schreiben. Ergebnisse angewandter Schreibprozessforschung. Konstanz: UVK.
- Perrin, Daniel (2003): Progression analysis (PA): Investigating writing strategies at the workplace. In: *Journal of Pragmatics* 35(6), 907-921.
- Plett, Heinrich F. (1991): Einführung in die rhetorische Textanalyse. Hamburg: Buske.
- Royce, Terry (1998): Synergy on the page: Exploring intersemiotic complementarity in page-based multimodal text. In: *JASFL Occasional Papers* Vol. 1 n°. 1, 25-48.
- Sandig, Barbara (1986): Stilistik der deutschen Sprache. Berlin: de Gruyter.
- Sandig, Barbara (2000): Text als prototypisches Konzept. In: Mangasser-Wahl, M. (ed.), *Prototypentheorie in der Linguistik. Anwendungsbeispiele – Methodenreflexion – Perspektiven*. Tübingen: Stauffenburg, 93-112.
- Stöckl, Hartmut (2004a): Die Sprache im Bild – Das Bild in der Sprache: Zur Verknüpfung von Sprache und Bild im massenmedialen Text. *Konzepte, Theorien, Analysemethoden*. Berlin: de Gruyter.
- Stöckl, Hartmut (2004b): In between modes. Language and image in printed media. In: Ventola, E./Charles, C./Kaltenbacher, M. (eds.), *Perspectives on multimodality*. Amsterdam: Benjamins, 9-30.
- Stöckl, Hartmut (2004c): Typographie: Körper und Gewand des Textes. Linguistische Überlegungen zu typographischer Gestaltung. In: *Zeitschrift für Angewandte Linguistik (ZfAL)* 41/2004, 5-48.
- Stöckl, Hartmut (2004d): Bilder - Konstitutive Teile sprachlicher Texte und Bausteine zum Textstil. In: *Mitteilungen des Deutschen Germanistenverbandes* Jg. 51, Heft 2/2004 (= Themenheft: Sprache und Bild II, herausgegeben von Holly, W./Hoppe, A./Schmitz, U.), 102-120.
- Stöckl, Hartmut (2005): Typography: body and dress of a text – a signing mode between language and image. In: *Visual Communication. Special Issue "The New Typography"*, vol. 4, n° 2, 204-214.
- Stötzner, Andreas (2003): Signography as a subject in its own right. translated from German by Hartmut Stöckl. In: *Visual Communication* vol. 2 (3), 285-302.
- Thibault, Paul (2000): The multimodal transcription of a television advertisement: theory and practice. In: Baldry, A. (ed.), *Multimodality and multimediality in the distance learning age*. Campobasso: Palladio Editore, 311-385.
- Van Leeuwen, Theo (1999): *Speech, music, sound*. London: Routledge.
- Van Leeuwen, Theo (2005a): *Introducing social semiotics*. London: Routledge.
- Van Leeuwen, Theo (2005b): Typographic meaning. In: *Visual Communication. Special Issue "The New Typography"*, vol. 4, n° 2, 137-143.
- Wehde, Susanne (2000): *Typographische Kultur: eine zeichentheoretische und kulturgeschichtliche Studie zur Typographie und ihrer Entwicklung (= Studien und Texte zur Sozialgeschichte der Literatur 69)*. Tübingen: Niemeyer.
- Ziegenrucker, Wieland (1997): *ABC Musik. Allgemeine Musiklehre*. Wiesbaden/Leipzig/Paris: Breitkopf & Härtel.